

Force-directed layout algorithms for Python (ID: 7442)

Owner/Supervisor: [Matthew Chalmers](#)

* Suitable for both IT/SD and CS+.

Description:

Python is used widely in data analysis and visualisation, often via tools such as Jupyter Workbench. Many common algorithms are freely available there, but there is a weakness in one type of visualisation algorithm: force-directed layout. Standard force-directed algorithms are weak in terms of the speed of execution and the scale of data set that can be laid out.

In SoCS, we developed some of the fastest force-directed algorithms anywhere. They were originally implemented in Java, but recent projects have re-implemented them in Javascript. It has been observed that they might work more quickly, and appeal to a different set of developers/analysts, if they were also implemented in Python.

The project here focuses on establishing fast and accurate methods, and preparing a solid set of experimental results based on a few different data sets, so as to compare our old Javascript and the new Python implementations done in this project.

Special Requirements :

None.

Game for supporting learning mathematics (ID: 7326)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+.

Description:

We have witnessed an increase in new and effective technology for our everyday use such as tablets, smartphones, games consoles and even smart TV's. The advent of HTML 5 supports cross compatibility between different devices within browsers.

At the same time, youngsters spend considerable amount of time on games. The idea of this project is to exploit this time creatively by developing a game for maths learning.

The project will focus on creating a game using suitable web technologies (HTML5/ javascript/React). The game will deal with several aspects of learning- initiation, reinforcement and reflection. It will be developed on a cloud platform.

Special Requirements :

None.

Lifelog Data Management - Front End (ID: 7335)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+.

Description:

Lifelogging is a process in which a person capture continuously data with live first-person video/images using a wearable camera (e.g., Autographer/ Google lens; Samsung - waistline-tracking smart belt and VR motion controllers).

However, we do not have a server side technology to gather, process and organise such vast and diverse amounts of data to exploit such data for later tasks, such as remembering events. In the school we have developed a number of lifelogging techniques (for

example <http://ieeexplore.ieee.org/document/7500236/?arnumber=7500236>).

The objective of this project is to build a scalable and efficient lifelogging front end. This involves developing modules in Javascript/React scripts; developing visualisation modules for summarising a day; a week; a month or an year; mapping location traces to online maps like Google maps. conducting user studies.

This will suit a person interested in learning popular Javascript/React libraries and developing secure and scalable solutions and conducting user studies.

Special Requirements :

None.

Sarcasm detection on Twitter (ID: 7261)

Owner/Supervisor: [Iulia Popescu](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The aim of this project is to build a sarcasm detection classifier utilising Twitter data. As a starting point, we will look at Bamman & Smith's model (2015) and try to explore the extra-linguistic information (audience, environment, etc) that can be used to infer the presence or absence of sarcasm in the interpersonal interaction.

Project plan:

- Familiarisation with Twitter API.

Understand the role of hashtags in messages.

- Look into related work on sarcasm detection in text (e.g. the SASI algorithm (Davidov et. al. 2010)).

- Build a script which retrieves tweets using the Twitter API.

- Decide which pre-processing techniques need to be applied to the tweets, in

order to facilitate the extraction of linguistic and non-linguistic features.

- Identify the features which are relevant for the detection of sarcasm.
- Determine which type of classifier is necessary and implement it.
- Evaluate the performance of the classifier and draw the appropriate conclusions.

This project will be co-supervised by Dr Nikos Ntarmos.

Special Requirements :

None.

38M pixel retina to emulate human visual performance (ID: 7196)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+.

Description:

We have developed a software retina that models the mammalian retina to compress the input image to a level where it can be processed by a Deep Convolutional Neural Network, DCNN, in a single pass. DCNNs currently must scan a small window over large input images, slowing their performance considerably. The retina preprocessor removes this limitation.

The pattern used to sample the input image and therefore simulate the cone cell sampling layout in the retina is produced by an annealing process that only needs to be executed once to obtain the sampling layout, but is very computationally expensive.

We currently have retina layouts that can sample images just under 1Mpixels and would like to be able to sample much larger input images in order to approach human levels of visual resolution over the field of view.

Therefore, this project aims to adapt or replace the current retina layout method, based on annealing to allow much larger retina layouts to be computed. We are procuring a Sigma Quattro H camera which produces 38M pixel images at full RGB resolution using the Foveon imaging sensor (no Bayer pattern subsampling) to provide input images.

It should be possible to demonstrate the new retina layout running on our retina GPU acceleration software (GTX1080TI based) and perhaps part of the existing retina-deep learning pipeline we have now demonstrated using the new ultra-high resolution retina.

We have already developed:

- * GPU codes contained in Python wrappers to allow the software retina to

compress input images to a size that the DCNN can handle directly,

- * Codes to find features in the cortical image output from the retina that can be used to direct where to look at next when either collecting training images ,or attempting to recognise unknown objects.

- * Basic DCNN model architectures that interface with the retina and have demonstrated learning and recognition.

- * Colour retina cells that sample the red-green & blue-yellow components of objects and their contours or surface texture, to provide colour discrimination.

Special Requirements :

No ethics required. CVMA4 desirable.

A benchmarking platform for depth cameras (ID: 7465)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

- * Suitable for both IT/SD and CS+.

Description:

This project is about building and developing a benchmark platform for depth cameras such as the Microsoft Kinect. The project will consist of interfacing a computer controlled turntable and 4 different depth cameras. The idea is to scan an object in order to characterise and evaluate these cameras. The benchmark evaluation comprise a set of pre-defined vision tests to quantitatively measure the depth accuracy, quality and repeatability of each depth camera with respect to ground truth. In this project, you will integrate all hardware and software components using the Robot Operating System (www.ros.org) and Python.

Special Requirements :

None.

A blockchain simulation environment. (ID: 7503)

Owner/Supervisor: [Inah Omoronyia](#)

- * Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A java application for simulating transactions between nodes in a network based on blockchain

The simulation should be able to model different node properties.

Example passive nodes that only make and forward transactions, and active nodes that are involved in the creation, validation and auditing of blocks

?

Special Requirements :

None.

A context-specific recommender system (ID: 7459)

Owner/Supervisor: [Matthew Chalmers](#)

** Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

Some years ago, I made a recommender system (called Recer) that tracked the histories of web and file activity for a number of people, and used temporal patterns in that activity to make recommendations of files and web pages for each individual: things that the individual had not seen, but that other people with similar recent histories had seen. in this way, it modelled recent/current context, and based recommendations on that -- instead of all of a person's history, as a recommender system like Amazon's might.

One student project used the same method to track Java coding activity inside an editor/IDE (JEdit), track a number of programmers, and recommend potentially useful files to each person based on recent editing activity.

I'd like to try these tools and methods out again, maybe in the same web/file/code setting but applied to new web browsers and editing tools, but perhaps in a different setting (e.g. tracking and recommending spatial locations on/near campus) but that would be something to discuss with the student. I've got the (very) old Java/SQL code, and the student would use this as a resource to help get them started.

Special Requirements :

None.

A Crowd based Fact Checking website- Front end (ID: 7614)

Owner/Supervisor: [Joemon Jose](#)

** Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

The news ecosystem has changed considerably, from a limited reputed and accountable community of publishers to numerous online real-time news sources. Changes in how the public consume news has made reading online news one of the most predominant activities of Internet users, and is clearly demonstrated by evolving statistics: 57% of readers routinely get their news from 2-5 sites; 33% of readers use search engines to find news; and 11% of web queries are related to current news.

Identifying relevant, real-time, unbiased information, from such data streams, is a daily activity of many. However, the increased ability to produce and

distribute information at will comes with a cost in transparency, and truthfulness of information. For example, Brandtzaeg et. al (CACM 2017) reported that false rumours and misinformation are disseminated further and faster than ever before due to social media: in 2012, 49% of US social media users received false breaking news through social media; political analysts continue to debate the influence of misinformation and fake news in online media and its effect on democratic elections. The impact and cost of such actions on society is enormous and ever-lasting: false rumours causing the death of innocent people or changing the trajectory of a democratic processes.

Fact checking sites help to clarify fake news, rumours, hoax. The objective of this project is develop a front end to a crowd engaged fact checking website. The idea is to allow genuine users to register with the fact checking website and allow this to verify the news rumours etc.

Special Requirements :

None.

A deep learning approach to musical effect reproduction (ID: 7151)

Owner/Supervisor: [John Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The application of musical effects such as reverb, distortion and chorus are key components of modern production. There are now tools that can reproduce the reverberation of existing places or synthetic effects using simple linear models. However, they cannot accurately reproduce nonlinear effects such as distortion or frequency modulation.

This project will look at using deep learning to model musical effects. This will produce emulations that sound like the original effect but do not require that hardware/software. For example, the specific sound of an existing tape deck might be precisely replicated so that period-accurate music can be reproduced.

This will involve development with deep learning tools like Keras or PyTorch.

Interested students should have a look at Google's Project Magenta, which has developed deep learning-based synthesis models for music:

<https://magenta.tensorflow.org/datasets/>

Knowledge of musical signal processing would be helpful.

Special Requirements :

None.

A Generic Assembler Generator (ID: 7314)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

An assembler is a program that reads in a source program written in an assembly language, and translates it to the corresponding machine language. A machine language specifies the internal representation of the instructions comprising a program (typically represented as hexadecimal strings, such as "62b1", while an assembly language allows the programmer to write the machine instructions in a readable symbolic form, such as "load R2,x[R2]").

There are many different computer architectures, each with its own machine (and assembly) language. Research in computer architecture involves experimenting with many variations in a machine language.

In order to support such research, it is useful to have an assembler generator. This is a program that reads in two inputs: (1) a specification of the computer architecture and assembly language, and (2) a source program written in that assembly language. The software then outputs the corresponding machine language result. A refinement on this is for the generator to take just one input (the specification of the architecture) and then to output a new computer program, which is the assembler.

The aim of this project is to design and implement an assembler generator. You will be given a notation for writing the architecture specification, and your software needs to read this in and generate the assembler. The software can be written in any suitable programming language, including Haskell, Java, or C.

This project is an example of one of the hottest current research topics in programming languages: you will learn about program generation, domain specific languages, and metaprogramming.

The outcome of the project will be a software tool that reads in an architecture specification and generates an assembler for that architecture.

Special Requirements :

None.

A Map of the Internet (MSc IT+) (ID: 7018)

Owner/Supervisor: [Colin Perkins](#)

*** Not suitable for CS+ students.**

Description:

Network routing researchers use graphs of the Internet constructed from real-world data in their research. These are large, complex graphs, the scale of which renders simple visualisation for easily conveying meaning to the viewer tricky at best. This project involves visualisation of Internet graphs to convey useful high-level information. The aim is to build tools which can generate visualisations similar to some of those hosted at CAIDA

(http://www.caida.org/research/topology/as_core_network/). Of particular interest is showing the extent to which IPv6 is used in the wild.

This project may interest students who wish to work with real-world data sets, or who have an eye for good visualisation techniques, and who are keen to learn more about the grubby low-level details of Internet routing.

Project goals: 1) Learn about the Internet, BGP, and inter-domain routing; 2) Learn and understand one or more graph decomposition algorithms; and 3) build application code to produce visualisations from real input data.

Outputs will be code to download, process, and visualise the data.

Special Requirements :

None.

A modular web-based booking system (IT/SD+) (ID: 7073)

Owner/Supervisor: [Bjorn Jensen](#)

*** Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.**

Description:

In this software development project, we will be looking to develop a modular and web-based system for managing (i.e creating, booking, canceling, notifying) and collaborating on increasing the utilization of a shared physical resource. There are certain temporal, access and regularity constraints to be considered in designing the system.

We will be looking to develop a general framework but the case study will be based on a specific resource such as a piece of (very expensive) lab equipment where several people have both economic and moral right to access the resources at certain times.

The project will be based on Python using standard libraries such as <https://www.djangoproject.com/> and <https://www.datacamp.com/getting->

started?step=2&track=python

The project will be tailored to the successful candidate.

Special Requirements :

None.

A neural network approach to biological cell segmentation (ID: 7076)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+.

Description:

This project will investigate and build a software pipeline for visualising and analysing images of biological cells using newly developed deep neural network-models which can segment, cluster and classify individual cells [3].

Practitioners without much experience with computer science or data analysis often apply computational methods in a suboptimal way [1,2]. This project aims at developing a web-based visualisation and machine learning pipeline which will provide non-CS experts with state of the art solutions for cell analysis based on recent results in machine learning and artificial intelligence.

The project is in collaboration with School of Engineering who will provide requirements and data. The project will be tailored to the successful candidate.

Prerequisites:

Interest in artificial intelligence, machine learning and image analysis.

Experience with Python and web-technologies (e.g. html, JavaScript) is an advantage.

Literature:

[1] cellprofiler.org

[2] Data-analysis strategies for image-based cell profiling, Caicedo et al, 2017, <https://www.nature.com/nmeth/journal/v14/n9/pdf/nmeth.4397.pdf?foxtrotcallback=true>

[3] Mask R-CNN, He et al, 2017, <https://arxiv.org/abs/1703.06870>

[4] Deep learning, <http://www.deeplearningbook.org/>

Special Requirements :

None.

A puzzle generator (ID: 7191)

Owner/Supervisor: [Alice Miller](#)

* Not suitable for CS+ students.

Description:

This project will build a system to generate puzzles. The system should ideally be web based and will generate puzzles such as Sudoku (<http://www.sudoku.com/>) or the 24 puzzle (Using four numbers and the usual arithmetic operations plus, minus, multiply and divide, you have to make exactly 24. No tricks are allowed like powers, cube roots, just pure arithmetic), or cross-reference puzzles (Cross Reference is considered to be the most difficult of the popular types of word-based puzzles around. It is similar to the Word Fit puzzles but you do not get the luxury of being given the list of words. Each letter of the alphabet is given a number (from 1 to 26). From these numbers in the grid you must try to find out which letter corresponds to a number. You complete the game once you discover the number to all the letters used in the grid).

You might like to consider other types of puzzles too.

Your website will not only generate and display these puzzles but also display hints to help the puzzler.

Special Requirements :

None.

A web application for the Robotics Foundations course (ID: 7467)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

* Suitable for both IT/SD and CS+.

Description:

The Robotics Foundations course is an introduction to the computational and mathematical concepts, information processing and software frameworks of robotic systems. The practical labs of the course make extensive use of the Robot Operating System (ROS; www.ros.org) and ROS' simulation and visualisation facilities. These facilities allow students to gain practical experience on interfacing robot modules together, implementing data information streams, and ultimately, how to program the Baxter robot. Currently, the labs are distributed a Docker containers running ubuntu, ROS and Jupiter notebooks.

The aim of this project is to develop a web application to provide students with a simple to use front-end for the labs. The specifications for this web app are:

- * Use of a wide range of internet technologies (i.e. Django, JQuery, AJAX, CSS, etc)
- * Interface with ROS and its facilities.
- * This front-end should be inside a single Docker container
- * Students should be able to access the app's functionality through responsive CSS navigation tabs

This project will be suitable to a student who enjoyed WAD2/ITECH. The project will be based on Django and Python and use the technologies covered in WAD2/ITECH plus learn about ROS; although the student will not develop software for ROS but use the existing labs to demonstrate the web app. The Docker component is optional if time permits.

Special Requirements :

None.

A web application to support the WAD2 course (ID: 6966)

Owner/Supervisor: [David F Manlove](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The Web Application Development 2 (WAD2) course covers a range of technologies that arise in the development of web applications, most notably the web application framework Django. The WAD2 assessment involves the development of an application called Rango, using Python and Django, which is marked using automated testing. Students can also carry out automated testing themselves in order to obtain early feedback on the development of their application.

The WAD2 course is supported by a web app that handles a range of functionality concerned with course administration. This app was developed by a Level 4 student in session 2017-18 and is already in use for the WAD2 course this session.

The purpose of this project is to augment the existing WAD2 web app so that automated testing functionality can be added for course administrators and students alike. Individual students should be able to initiate the testing and obtain feedback on how many they passed, while course administrators should be able to begin a batch run of testing for the whole class after certain assessment deadlines have passed.

It is likely that this application will make use of many of the technologies covered in the ITECH course, and hence this project is most suitable for a

student who enjoyed ITECH and who has a firm grasp of Python, Django, CSS, HTML, XML, Javascript, jQuery and AJAX.

Special Requirements :

None.

Acronym translator (ID: 7006)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Many academic papers are full of acronyms, making them difficult for people not familiar with the field to read. The goal of this project is to create an automatic acronym decoder that provides predictions in a particular paper (based on the paper's content) of the meaning of the acronyms present. User's can then hover over a particular acronym and see the predicted meaning. As an extension, the user should be able to re-learn as the user identifies errors in the translations.

Special Requirements :

None.

Active (machine) learning in biology (ID: 7263)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

In this project, we will develop a software system (with web-based user interfaces and machine learning components) to speed up biological cell annotation and analysis via so-called active (machine) learning [1].

Active (machine) learning [1] seeks to query an expert user about uncertain examples, e.g. an image of a cell to quickly learn about the phenomenon being studied (e.g. nucleus detection, cell segmentation or classification).

The machine learning component can be based on existing deep neural network or be developed from scratch.

We will look to demonstrate the approach on data from a recent Kaggle competition: <https://www.kaggle.com/c/data-science-bowl-2018>

The project will be tailored to the successful candidate.

[1] <http://burrsettles.com/pub/settles.activelearning.pdf>

Special Requirements :

None.

Adapting a GPU-accelerated weather simulation to allow larger simulations (ID: 7316)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Many physical models (e.g. weather simulations) have a very large problem space (e.g. the portion of the atmosphere being simulated). Frequently the problem size is limited by the available memory, esp. on accelerators such as GPUs.

The aim of this project is to investigate partitioning of the data in memory so that it becomes possible to run much larger problems by consecutively running part of the data in memory. As a practical example the project will use an OpenCL weather simulator. The goal is to develop a suitable compression mechanism and assess its performance.

This project requires you to know or learn OpenCL (i.e. C) and Fortran (only to be able to understand the host source code).

Special Requirements :

None.

Adaptive feature selection for Internet of Things environments (ID: 7470)

Owner/Supervisor: [Natascha Harth](#)

* Suitable for both IT/SD and CS+.

Description:

This master project is intent to identify feature selection methods that can be adopt in an Internet of Things (IoT) environment. The focus on these networks is to run highly efficient algorithms with low computation power as these networks are constrained with limited power resources and computational abilities. The outcome of this project is to identify different feature selection mechanism that can be adopted in a streaming environment and that is adaptive to the changing data distribution coming from IoT data. The student should implement one or many light weighted feature selection methods using python. Also the aim is to transfer the implemented code into a RaspberryPi network. This should give the student the opportunity to test the properties and

abilities of the proposed feature selection methodology in a real IoT network.

This project will be co-supervised by Dr. Christos Anagnostopoulos

Special Requirements :

None.

Agent Based Modelling of Social and Spatial Processes (ID: 7400)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

Agent based modelling means representing individuals as distinct objects in a computer virtual environment and observing how they react. A number of interesting results have been obtained, and I would like this project to check the results of two of the studies reported at the 1997 conference at the Sante Fe Institute. The first study modelled primate societies and came up with some surprisingly simple theories to explain dominance behaviour, while the second tried to explain the genetic basis of altruism and cooperation.

The computer system developed will have the following components:

(i) A simple model of the spatial environment, together with a pleasing visualization.

(ii) A simple model of individual 'agents' and their movement strategies.

(iii) A simple and flexible model of interactions between two agents.

There are a number of interesting software engineering aspects to this project.

(i) The system will support two different simulations, and so issues such as code sharing and reuse verses application specific code will need to be addressed.

(ii) Each simulation will involve a number of different agents acting simultaneously. Is the best way to engineer this to use a different thread for each agent, each accessing a shared resource, the environment? Alternatively, should our code control any interactions through a synchronizing Agenda object?

(iii) Each agent will be controlled by its own internal 'program'. Should this be hard coded, or should more flexibility be provided by using controller objects to encapsulate the internal 'program'? If so, what is the basic behaviour that each controller object must conform to? The controller program can potentially be quite complex. Is it best to invent a new program language to describe it, or should we stick with a well know computer science programming language?

This is essentially a software engineering project. There are a number of requirements to be determined, followed by a risk assessment. A number of

experiments will need to be defined, leading to prototypes that help to answer the questions raised above. Finally, the two simulations need to be designed and built.

This project does not require any knowledge of sociology and related topics, although an interest in the area of modelling behaviour would be useful.

Special Requirements :

None.

Algorithm Animator (ID: 7322)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

For teaching purposes it is useful to be able to animate algorithms and produce a visual representation of how they work. The basic idea is to use a diagrammatic representation of a data structure, for example an array or a tree, and illustrate the algorithm step by step, showing how the data structure is accessed and changed. The aim of this project is to design and implement a system for animating algorithms. There are at least two possible approaches. One is to design and implement a simple programming language in such a way that all programs are animated while being executed. Another is to design and implement an API for animations, so that an existing program (in Java, for example) can be animated by inserting calls to your library. The system should be as general as possible in the sense of supporting a range of styles of algorithm, and should be demonstrated by producing a range of animations of standard algorithms. It would also be useful to be able to capture the animation in a form that can be viewed independently of your system, for example as a sequence of HTML pages or a Flash animation.

Special Requirements :

None.

Algorithm Animator (ID: 6956)

Owner/Supervisor: [Gethin Norman](#)

* Suitable for both IT/SD and CS+.

Description:

This aim of this project will be to build a system which can be used to animate/visualize the steps of an algorithm. Potential algorithms include those for text compression or graph-based algorithms, but I am open to suggestions.

Special Requirements :

None.

An applicaiton for collaboratively improving utilisation of shared physical resources (ID: 7072)

Owner/Supervisor: [Bjorn Jensen](#)

* Not suitable for CS+ students. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

In this software development project, we will be looking to develop a modular and web-based system for managing (i.e creating, booking, canceling, notifying) and collaborating on increasing the utilization of a shared physical resource. There are certain temporal, access and regularity constraints to be considered in designing the system.

We will be looking to develop a general framework but the case study will be based on a specific resource such as a piece of (very expensive) lab equipment where several people have both economic and moral right to access the resources at certain times.

The system will ideally include a piece of advance scheduling software mitigating different conflicting interests in a "rational" manner.

The project will be based on Python using standard libraries such as <https://www.djangoproject.com/> and <https://www.datacamp.com/getting-started?step=2&track=python>

The project will be tailored to the successful candidate.

Special Requirements :

None.

An execution environment for lambda-calculus (ID: 7318)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+.

Description:

Lambda-calculus is the theoretical foundation of functional programming, studied for example in the Theory of Computation course. For educational purposes, it is useful to have a way of defining and evaluating lambda-calculus expressions. Although it is possible to do this within standard programming languages, for example Python, there are some drawbacks - especially, that it's not possible to choose different reduction strategies (eager, lazy, call-by-need).

The aim of this project is to implement a system in which lambda-calculus expressions can be defined and evaluated, with a choice of reduction strategy. It should be possible to execute step-by-step, or run all the way to a normal form. It should be possible to visualise expressions by expanding or collapsing parts of them. Ideally the system should be implemented as a web application.

The product will be an execution/simulation system for lambda-calculus.

Special Requirements :

None.

An execution environment for pi-calculus (ID: 7317)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+.

Description:

Pi-calculus is a modelling language for concurrent communicating systems, used to study the theory of concurrency (for example in the Theory of Computation course). It is based on the idea of sending messages on point-to-point communication channels. This could be done by translating it into Go, which has the same communication mechanism, or by building on some other language's concurrency mechanism, or by implementing something completely new. The result will be a system that can be used for teaching pi-calculus, for example in the new honours course Theory of Computation. Ideally it will be implemented as a web application.

As well as implementing an execution environment for pi-calculus, you can implement a typechecker for pi-calculus, to check that channels are used for messages of the correct type. There is a large amount of research that uses pi-calculus as a vehicle for developing and studying type systems for concurrent programming, so it will be possible to implement some of these type systems.

Find out more about pi-calculus at <https://en.wikipedia.org/wiki/?-calculus>

The product will be a typechecker and execution system for pi-calculus.

Special Requirements :

None.

An inverted approximate subgraph index (ID: 7223)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

The goal of this project is to produce a system that will index graphs in a graph database/dataset, in order to allow for them to be retrieved based on their similarity/distance to a graph-structured query. The student will first be briefed on the type of graphs this work will be carried on. Then the student will (a) review relevant literature on graph distance/similarity metrics and on approximate subgraph matching, (b) create a simple data generator to produce test graphs and queries, (c) design and implement an indexing structure to expedite these queries, and (d) evaluate the system's performance from both a qualitative and quantitative point of view.

Special Requirements :

Excellent programming skills. Good understanding of data structures and algorithms.

Analysing network protocol standards (MSc IT+) (ID: 7019)

Owner/Supervisor: [Colin Perkins](#)

* Not suitable for CS+ students.

Description:

The Internet Engineering Task Force (IETF) develops technical network protocol standards for the Internet. These standards, and the working drafts that lead up to them, are all publicly and freely available at <https://www.ietf.org/> as a series of text files, giving specifications for the protocols (a simple example is <https://datatracker.ietf.org/doc/rfc1939/>).

The goal of this project is to develop analysis tools and visualisations of the standards. Download and parse the standards and working drafts, extract information about the authors, citations, updates/obsoletes, and other metrics of interest, then (a) extract this into a machine readable format for future processing; and (b) visualise the data as a series of webpages, figures, etc.

To start, this would be able to generate high-level visualisations and statistics about the standards like those at <http://www.arkko.com/tools/docstats.html> (but generated by tools written in a maintainable manner, rather than using a series of bash scripts). Perhaps

more interesting is then to look at the interdependencies, cross-references, and development of the standards over time, and to develop ways of showing this information. Finally, a stretch goal might be to detect and extract particular information from the standards (e.g., to find and extract packet format diagrams, sample code, etc.).

This project would suit a student interested in how network protocol standards are developed. Implementation would likely be in Python. There's no network programming in this project: it's about understanding the way IETF standards are developed.

Special Requirements :

None.

Animating a Max Clique Solver (ID: 7213)

Owner/Supervisor: [Patrick Prosser](#)

* Suitable for both IT/SD and CS+.

Description:

Given a graph $G=(V,E)$, a clique is a subset of the vertices in V , such that all vertices in that subset are pair-wise adjacent. A max clique is a largest clique in G . In this project we will take an existing algorithm and animate it, so that we can see how it goes

Special Requirements :

None.

Animating geometric algorithms (ID: 6939)

Owner/Supervisor: [David F Manlove](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A useful way to learn about how an algorithm works is to visually demonstrate its execution as applied to a problem instance. This can be done by illustrating the decisions that an algorithm makes at various points during its execution via a user interface.

This project is concerned with animating various geometric algorithms, including algorithms for constructing simple polygons, convex hulls, and finding pairs of points that are closest and furthest apart.

The challenges relating to this project include allowing the user to input points interactively via the mouse, reading instance data from a file, implementing the various algorithms and animating their executions.

This project would suit a student who has taken a course in data structures and algorithms either during their Masters programme or during their undergraduate degree.

For more information about geometric algorithms, see M. Goodrich and R. Tamassia, Algorithm Design: Foundations, Analysis, and Internet Examples, Wiley, 2002, Chapter 13.

Special Requirements :

None.

Animating maximum matching algorithms (ID: 6965)

Owner/Supervisor: [David F Manlove](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A useful way to learn about how an algorithm works is to visually demonstrate its execution as applied to a problem instance. This can be done by illustrating the decisions that an algorithm makes at various points during its execution via a user interface.

This project is concerned with animating algorithms for finding maximum matchings in bipartite graphs. These algorithms have applications to matching students to projects, or junior doctors to hospitals. It is not necessary to have studied these algorithms before in order to do this project, as relevant background will be given.

The challenges relating to this project include allowing the user to input graphs interactively via the mouse, reading instances from a file, implementing the various algorithms and animating their execution.

This project would suit a student who has taken a course in data structures and algorithms either during their Masters programme or during their undergraduate degree.

For more information about matching algorithms, see C. H. Papadimitriou and K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, 1982, pp. 218-225.

Special Requirements :

None.

Animation of Digital Circuit Simulation (ID: 7323)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

A good way to understand the operation of a digital circuit is to watch an animation of its behaviour as signals propagate through the logic gates, and flip flops are updated at the clock ticks. A software tool that implements circuit animation is useful for digital circuit designers, as it helps to test and debug circuits, as well as to assess their performance.

The aim of this project is to design and implement an interactive program that shows graphically the events that occur as a digital circuit is executing. The input to the program is a specification of the structure of the circuit, in the form of a netlist. The program reads and parses the circuit netlist, draws the circuit diagram on screen, and then simulates the circuit, showing the values of signals as they become valid.

The program will be implemented in either Java or Haskell. It will require the development of a circuit simulation algorithm, as well as graphics programming.

Product: a program that simulates a digital circuit, using graphics to show signals as they become valid, and the progress of clock cycles.

Special Requirements :

None.

Animation of the type-checking and code-generation phases of a compiler (ID: 7324)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+.

Description:

For teaching purposes, for example in the Programming Languages course, it would be useful to have a system to animate and illustrate how a type-checker and code-generator traverse an abstract syntax tree during the compilation process.

You can use the Fun language that's the basis for the PL course, and develop an implementation that illustrates the work of the type-checking and code-generation phases by displaying the syntax tree and showing how each node is analysed.

Ideally this should be implemented as a web application, using any language or framework that you find convenient.

The product will be the implementation of the animating compiler.

Special Requirements :

None.

Annotated Picture Editor (ID: 7401)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

This project will build a tool that allows users to construct and manipulate a fairly complex picture, adding comments as they go. The process of constructing the picture will be saved, not just the final picture. This will let the user replay the process, reading the comments as they go. The user will also be able to edit the construction process, cutting out dead ends and long pauses while they thought of what to do next. This will produce a concise description of the picture construction process.

The motivation for doing this is in designing visual tools for scientific computing, where the scientist is conducting a virtual experiment and wishes to be able to recall it later on, editing out his mistakes and so on.

Special Requirements :

None.

Assessing and improving the UMAP visualisation algorithm (ID: 7453)

Owner/Supervisor: [Matthew Chalmers](#)

* Suitable for both IT/SD and CS+.

Description:

Recently, a new visualisation algorithm for high-dimensional data was announced. An reference implementation (in Python) is here:

<https://github.com/lmcinnes/umap>

and also a (quite mathematical) paper from Feb 2018 is here:

<https://arxiv.org/abs/1802.03426>

It claims to be faster and better than the most favoured algorithm of its type, tSNE, which is part of the established data analytics toolkit for Python, scikit-learn.

This project will begin by getting to understand the implementation and the paper, and comparing UMAP against tSNE for some different data sets.

(We've often found that algorithms like this look scarily complex when described mathematically, but when you look closer at the code then they are not that bad.) Then, the project will look for ways to speed up or otherwise improve the algorithm, implement them, and do experiments to show their costs and benefits.

Special Requirements :

None.

Augmented VR/AR keyboard (ID: 7441)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Virtual and Augmented Reality (VR/AR) headsets have potential to be able to augment our physical keyboards. For example, keys could be highlighted, shortcut keys could be visualized, and per-application keyboard layouts provided, with each key instead rendering an image of what function it performs. This project will explore what augmentations are possible, when these augmentations should be presented, and what benefit they provide over existing keyboards without visual augmentation

Special Requirements :

None.

Augmented VR/AR touchpad (ID: 7448)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Given a multi-touch, pressure sensitive touchpad, this project will investigate potential augmentations of this touchpad in Virtual and Augmented Reality (VR/AR). For example, can touchpad gestures be made more accessible by rendering options on the touchpad surface; when looking at the touchpad, should it continue to control a mouse cursor, or transition into a different mode? And if so, what might this mode be e.g. rendering shortcuts, touch overlays specific to each application, or turning into a direct-touch version of the current application? This project would be expected to pick one of these possibilities and develop and evaluate this interaction technique using a VR headset.

Special Requirements :

None.

Avoiding Unintentional Gestures (ID: 7349)

Owner/Supervisor: [Euan Freeman](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Mid-air gestures are hand or finger movements used to control a user interface. False-positive gesture recognition occurs when a user interface mistakenly thinks a person has performed a gesture. Incidental hand movements, like reaching to pick up an object, may be incorrectly treated as a gesture and could unintentionally perform a function of the user interface. This is undesirable, so many solutions have been investigated for reducing the likelihood of false-positive gesture recognition. For example, using a 'clutch' gesture to explicitly initiate interaction and performing gestures repeatedly to show that they are intentional. These solutions require more effort from users.

This project will consider a different approach, where false-positive recognition is allowed to happen and it is up to the user to decide if the action should be cancelled. If the system thinks the user has gestured unintentionally, it will give them a chance to 'cancel' the gesture: if the user cancels, no further action is taken; if the user does not cancel, the action is executed. A key benefit of this approach is that gesture recognition constraints can be relaxed, because false-positive recognition is the user's responsibility, not the system's responsibility. This could also help users develop a better understanding of how the gesture system works, as they learn which behaviour can be recognised by the system, even though they might not intend to interact.

There are two design goals for this project:

- (1) design and implement a method for informing users that a possible false-positive gesture has been recognised; and
- (2) design and implement a mechanism for cancelling the gesture.

The first part could be something simple like a count-down timer that shows a gesture has been recognised and informs the user that they have a short period of time to cancel its action. The second part could be something simple like waving the hand or holding it still over the device for a couple of seconds. A Leap Motion sensor will be provided for implementation; this is a hand and finger tracker with a good SDK for most programming languages.

A user evaluation is required for this project. The evaluation will see if the feedback is informative and effective, and if the cancel gesture is easy to perform.

Special Requirements :

None.

Baby Steps : Visualizing Software Development by Novice Programmers (ID: 7531)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

At Glasgow, we run a massive open online course (MOOC) to teach people about functional programming in Haskell. As part of this MOOC, learners have to work through a scaffolded set of one-line Haskell expressions. Some learners enter the code we tell them too, whereas other (more adventurous) learners go off-piste and try to evaluate alternative expressions.

This project is about programmatically generating novel visualizations of the routes different learners take through our Haskell tutorial sessions. You might start with something as simple as a WordCloud, and then progress onto more advanced flow-charts or similar. The end-result would be a system that reads in a set of learner log files, and outputs a set of image files summarizing the learner interactions.

Skills required:

- - graphical design
- - Unix scripting
- - interest in Functional Programming - ideally taking FP(H) course

You can read a little more about analysis of MOOC Haskell learners by reading about Functional Baby Talk at the [TFPIE conference website](#).

Special Requirements :

None.

BOINC for Raspberry Pi Clusters (ID: 7524)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+.

Description:

BOINC [\[link\]](#) is a volunteer computing platform.

It allows users to donate CPU cycles to solve complex scientific problems.

There is a BOINC client for Raspberry Pi - see [here](#) for more details on how to set this up.

At Glasgow, we have several large clusters of Raspberry Pi nodes. The aim of this project is to create a software management tool to allow a RasPi cluster administrator to donate the entire cluster to solve BOINC tasks. Effectively, you will be creating a high-level task orchestration framework - either as a scriptable API, or a GUI, or both.

We will provide access to a Raspberry Pi cluster. There will be opportunities to demo this project as part of the FRuIT testbed activities - see [website](#)

Skills required:

- - Linux experience
- - Python, C or Java programming

Special Requirements :

None.

Building a Questionnaire delivery application (ID: 7463)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Existing digital Likert-type questionnaire scales do not take into account a) your previous responses b) the difficulty in making relative comparisons when responses are close together (e.g. both responses are "highest", but is there any relative difference between them - is one higher still?) c) being constrained by previous responses at the extremes of a scale (e.g. previous response was "highest", but next response should be much higher still). This project would involve designing and implementing a new digital questionnaire application that offers solutions to these problems and allows richer questionnaire scales to be presented.

Special Requirements :

None.

CcubeFlow: TensorFlow based implementation of clustering cancer cell fractions (ID: 7095)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Clustering cancer cell fraction (Ccube) is a powerful probabilistic model for determining the clonality from whole genome sequencing data. Currently, Ccube is prototyped as an R package (<https://github.com/keyuan/ccube>), limiting its scalability and flexibility. In this project, we will rewrite Ccube in Python with the popular TensorFlow engine from Google.

The complete software package will have three components:

1. Parser and preprocessing for standard genetic variants formats
2. Core Ccube algorithm
3. Postprocessing for quality control and result formats

Special Requirements :

None.

collaborative Web Search back-end (ID: 7310)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+.

Description:

Web search interfaces are mainly designed to support a single searcher at a time. However, more and more search activities are collaborative in nature, examples include, a couple searching for a Hi-Fi specification, a group looking for a holiday information. Concurrent search is a type of collaborative search where the members of a team share a period of time to achieve a goal for example, a holiday planning where a group of people try to find travel information together. With the advent of Web 2.0 and other technologies, there is a need to design and implement effective collaborative search interfaces.

There are some research activities in designing collaborative search interfaces, for example, SearchTogether system from Microsoft and VideoSearch system from FX Pal. In this project the design and development of a concurrent collaborative search interface is proposed. This will be based on our earlier work in this area (Comparing collaborative and independent search in a recall-oriented task to be presented at the IiX 2008 conference.). The activities include: development of system specification (mainly be analyzing there research publications and investigating the Web 2.0 technologies) and creating a design document; implementation of the system and if possible a user-evaluation.

In this project we will mainly focus on the back end/ server side technology/cloud technology. We will use a cloud server – Amazon or Google or Microsoft Azure for the development.

In addition to designing database modules consider attention needed to

designing secure access systems. It is expected that pairs of users log in to the server and hence we need effective and secure modules for “user id” management. Often users need to invite partners to search together and hence mechanisms to invite co-searchers needed.

Other component is coordination/ communication tools for the searchers. This may take the form of messaging system/commenting system;

Since a prototype is meant for practical use and user-experimentation a good quality software development is needed. In addition, the requirement capture involve some level of literature survey and a bit of interest in research is required. Given these activities, I would rate this project as challenging and hence suits candidates who can handle advanced software engineering/system building and are very serious about their project.

Special Requirements :

None.

Comparative study of Bayesian filtering methods in state space point process model (ID: 7097)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Background:

Particle filter (Doucet et al, 2001) and Gibbs sampler (Geweke and Tanizaki, 2001) are two

classical Bayesian inference methods to solve filtering problem in time-series data. In this

project, we compare the property and performance of the two methods in state space

model with point process observations (Smith and Brown, 2003). The results will shed light

on how to choose between machine learning methods designed to solve the same

problem but using different principles.

Details of the project:

1. Implementation of particle filter for SSPP.
2. Implementation of single-site Gibbs sampler for SSPP.
3. Find metric and carried out comparative study between the two inference strategies.

References:

1. Doucet, de Freitas, and Gordon (2001). Sequential Monte Carlo methods in practice. New York: Springer-Verlag.

2. Geweke and Tanizaki (2001). Bayesian estimation of state-space models using the

Metropolis-Hasting algorithm within Gibbs sampling. Computational Statistics and

Data Analysis, 37(2), 151–170.

3. Smith and Brown (2003). Estimating a state-space model from point process

observations. Neural Computation, 15(5), 965–991.

Special Requirements :

None.

Completing an on-line system to help learners develop code comprehension skills (ID: 7580)

Owner/Supervisor: [Quintin I Cutts](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The School has pioneered a method for helping novice programmers to understand how simple programs execute. This has been initially trialled as a pencil-and-paper exercise. An earlier student project has started to convert this to an on-line system - preserving the best of the off-line manual system, while opening up the huge potential for a system that can be shared easily on-line.

The earlier project has only got half way to full deployment. The work to date

is excellent, and will be left in a good state to take to completion. You would be expected to trial the completed system with teachers and maybe also pupils. The project is probably better suited to those with stronger programming skills.

Special Requirements :

None.

Constructing and analysing quantitative scheduling problems (ID: 6953)

Owner/Supervisor: [Gethin Norman](#)

* **Suitable for both IT/SD and CS+.**

Description:

Scheduling problems concern how to determine the schedule the execution of a number of tasks or jobs on a number of processors or machines such that some properties (e.g. the time or energy required to complete all tasks is minimised). Such problems occur naturally in many applications domains including multi-processor, manufacturing and mobile systems.

There are a number of variants, for example, different tasks can require different resources and some tasks cannot be started until other have been completed. In addition, different processors/machines can require different times to perform the same task and may not have fixed execution times or may fail.

The probabilistic model checker PRISM can be used to construct optimal schedulers in the presence of randomised delays and failures. However, low level understanding of both the input to the tool and output is required. The aim of this project is to construct a translator from a high level input to the PRISM input code and a corresponding translator back from the low level output. This will allow users unfamiliar with PRISM to solve

scheduling problems with PRISM as well as being a useful tool to those already familiar with PRISM.

Only a limited understanding of the PRISM tool and its language will be required, with the focus will be on how to develop an interface to construct instances of scheduling problems and view the returned optimal schedulers.

Special Requirements :

None.

Continuous visual perception for object recognition (ID: 7477)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

* Suitable for both IT/SD and CS+.

Description:

We, humans, are able to continuously see the environment, yet robots have to see then think what to do with the acquired information. The aim of this project is to investigate a continuous visual perception approach to recognise objects within a robot's environment. For this project, you will use the Robot Operating System, Python and Scipy to develop the software modules for Baxter.

The idea is to use Gaussian processes techniques to aggregate probability predictions through time to successfully identify and locate an object on a tabletop. The later will run concurrently as Baxter attempts to grasp the detected object.

If time permits, it might be possible to explore advance machine learning techniques for spatio-temporal recognition.

SciPy: [http://scikit-](http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier)

[learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier](http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessClassifier.html#sklearn.gaussian_process.GaussianProcessClassifier)

ROS: <http://www.ros.org>

Special Requirements :

None.

Conversational interaction with Amazon Alexa Skills Kit (ID: 7122)

Owner/Supervisor: [Mary Ellen Foster](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The goal of this project is to develop a conversational dialogue system using the Amazon Alexa Skills Kit (<https://developer.amazon.com/alexa-skills-kit>). You will develop and test a new interface that supports conversational interaction. For development, you will mainly use the web-based tools provided with Alexa. However, for final demos and user evaluation, you will have the opportunity to use an Amazon Echo device.

The student should have a good background in web-based programming, ideally along with an interest in dialogue systems and natural language processing. They will also need to carry out a user study with human participants at the end of the project and to analyse the resulting data for the final report.

Special Requirements :

None.

Conveying Emotion with Thermal Interfaces (ID: 7469)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Despite the development of many types of interfaces e.g. visual, audio, haptic etc., it can still hard to convey various types of information in an expressive manner. In an attempt to address this issue this project involves investigating the use of thermal interfaces, i.e. devices that heat up and cool down. Thermal stimulation is a rich, emotive and salient feedback channel that is well suited to HCI, but one that is yet to be fully investigated.

The aim of this project is to investigate augmenting physical objects with thermal feedback to add extra affective richness to the interaction experience. You will develop a set of prototypes to evaluate the benefits of thermal feedback. In conjunction with this you will also conduct usability evaluations to investigate the usefulness and impact of adding the thermal feedback.

Final Product: A suite of demonstrators that show adding thermal stimulation to interaction. A set of design guidelines based on the evaluation of those demonstrators.

Special Requirements :

will need to use our thermal boards

Creating pseudo hand drawn scientific surface plots (ID: 7152)

Owner/Supervisor: [John Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

This project will look at augmenting 3D rendering software to produce high-quality scientific 3D surfaces with pseudo hand drawn appearance. The project will focus on extending the Blender [blender.org] renderer to produce these graphs, most likely via Python scripting.

Traditional hand drawn plots (see below) are able to reveal structure in 3D surfaces that is often lost in modern renders. Although modern renders have accurate light transport models, they are designed for photo realism rather than to reveal the structure of surfaces. This is particularly relevant when producing figures for reproduction, which must be clear and might only be monochrome.

Traditional artists developed techniques to reveal shading and surface features. The aim of this project will be to develop a system to produce high quality images *automatically*, according to a specification provided by a user (e.g. line-only, highlight creases, no lighting).

Experience with 3D rendering is helpful (particularly with Blender).

Examples:

The following are some example images of hand drawn 3D surfaces:

http://4.bp.blogspot.com/-ZAoh2R_Gi8E/UXVhs22EeYI/AAAAAAAAABYA/9JL1mKqrfdI/s640/image_epigenetic.jpg

<https://faculty.math.illinois.edu/~gfrancis/miami/penchalkink.html>

http://www.amoebe.com/san/uploaded_images/Francis_topology_90_94_sm-726109.jpg

<https://i.stack.imgur.com/jHCnc.jpg>

<https://i.stack.imgur.com/V1mGT.png>

<https://i.stack.imgur.com/7iWoA.png>

<https://i.stack.imgur.com/zRvTI.png>

<https://i.stack.imgur.com/X9u5O.png>

<https://i.stack.imgur.com/b1php.png>

Special Requirements :

None.

Data collection and visualisation tool for Smart Pebble sensing software (ID: 7134)

Owner/Supervisor: [Alice Miller](#)

* Suitable for both IT/SD and CS+.

Description:

This project is to build software for Georgios Maniatis, a Research fellow in the School of Geographical and Earth Sciences (GES).

The Smart Pebble is designed for sediment transport investigation. It is essentially a synthetic pebble made of metal designed to be the same density as a rock, with a waterproof pocket inside for sensors. Inertial Motion Unit (IMU) sensors can provide high frequency data on accelerations and angular velocities (hence forces and energy) experienced by grains during entrainment and motion. In parallel to the development of the sensing equipment it is crucial to develop a user-friendly data collection and visualisation GUI which will enhance the interpretation and the post-processing of the measurements. The same software can provide the base for future advancements in the communication stack of the system, helping us to recognise abrupt riverbed changes and areas that are highly susceptible to the risk of flooding.

Special Requirements :

None.

Data Management Tools for Retina-Based Vision Systems (ID: 7268)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

We have developed a vision system based on a software model of an artificial retina and now have many sources of data produced by this retina, such as

camera glasses, robot vision systems, eye-tracking camera glasses, cameras integrated within the robot's gripper or on its wrist and even a retina implementation on an iPhone.

A common characteristic of all of these sources of data is the need to collect multiple views of each object the system is to recognise which have been sampled by the retina. There can be potentially thousands of images of each object - in some cases we have been capturing as many as 7,500 images samples of each object we gather image data about.

The objective of this proposal is to continue work we have initiated to build a data management tool that allows us to open files of the retina image sampled data for each object, review these samples to check for "bad data" and re-allocate different class labels to the samples as required and then export the data in a form suitable to train a deep net.

This data management task will involve implementing some of the image processing function that takes place in the retina to convert the sample data into a "cortical image" that can be used for deep net training.

If time permits the system should be demonstrated training a standard deep net architecture.

Special Requirements :

No ethical implications.

Database for go games (ID: 7333)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

Go games are recorded using a standard file format called sgf, which contains metadata about the game, the sequence of moves, and also analyses of alternative sequences. Many sgf game records are available.

The aim of this project is to develop a program that takes a set of sgf game records and finds games that match various search criteria, including particular positions and move sequences. In order to display these, a graphical user interface is needed.

It will be necessary to learn the basic rules of go, but you do not need to know the game in advance to do this project, and it will not be necessary to learn to play well. The program can be implemented in Java.

Special Requirements :

None.

Decomposing smCOGS (ID: 6987)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Parts of the genome known as Biosynthetic Gene Clusters (BGCs) can be represented as a binary string indicating the presence and absence of particular smCOGS. Decomposing this data to find conserved smCOG families would be a useful advance. Here you will develop a matrix factorisation model that is able to do this, assuming that each BGC is generated via a combination of factors, each providing a probability of presence or absence for a particular smCOG. You will evaluate your approach on data from bacterial gene sequences (already processed into smCOG space). If time permits, it would be useful to extend the work to visualising the result of the decomposition.

Special Requirements :

None.

Deep Gaussian Processes vs Bayesian Networks for Human Motion Capture - an empirical evaluation (ID: 7212)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+.

Description:

Both Deep neural networks placed in a (parametric) Bayesian framework and Deep Gaussian process [1] supports a principled representation and dissemination of uncertainty.

This project will look to compare Bayesian neural networks and Deep Gaussian processes focusing on their ability to represent calibrated uncertainties for visualization and interpretation, for example, using human motion capture data.

Pre-requisites: An interest in calculus, linear algebra and probability is an advantage.

The project will be tailored to the successful candidate.

[1] <http://proceedings.mlr.press/v31/damianou13a.pdf>

Special Requirements :

None.

Deep Sensor Fusion for Inertial Sensors (ID: 7154)

Owner/Supervisor: [John Williamson](#)

*** Not suitable for CS+ students.**

Description:

This project will investigate minimal calibration approaches for inertial measurement units (IMUs), which measure motion and orientation via gyroscopes, accelerometers and magnetometers. Traditionally, these sensors are fused using probabilistic filters and need extensive calibration.

This project will investigate applying deep learning to obtain fused, reliable estimates of orientation from IMU devices. Ideally, this will be able to recover device orientation for human motion tracking purposes with minimal calibration.

Implementation will be in Python with standard deep learning libraries (Keras/Tensorflow or Pytorch)

Special Requirements :

None.

Designing an intuitive interface for Collaborative Search System (ID: 7304)

Owner/Supervisor: [Joemon Jose](#)

*** Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.**

Description:

Web search interfaces are mainly designed to support a single searcher at a time. However, more and more search activities are collaborative in nature, examples include, a couple searching for a Hi-Fi specification, a group looking for a holiday information.

Concurrent (Synchronous) search is a type of collaborative search where the members of a team share a period of time to achieve a goal for example, a holiday planning where a group of people try to find travel information together. It is a challenge to design systems that allow people to search at the same time but physically separated.

In asynchronous search, searchers look for information at different time intervals from different IP addresses. Imagine a scenario where a person searching from Glasgow and a partner searching from Los Angeles, USA at different times of the day.

With the advent of Web 2.0 and other technologies, it is possible to design and implement effective collaborative search interfaces. There are some research activities in designing collaborative search interfaces, for example,

SearchTogether system from Microsoft and another VideoSearch system from FX Pal. In addition my research group has developed a number of collaborative search systems as well.

In this project the design and development of a collaborative search interface is proposed. This will be based on our earlier work in this area (Comparing collaborative and independent search in a recall-oriented task presented at the IliX 2008, ECIR 2009, AIRS 2012 conferences.).

The activities include: refine the system specification (mainly by analyzing their research publications and investigating the Web 2.0 technologies) and creating a design document; implementation and evaluation of the system.

Since a prototype is meant for practical use and user-experimentation a good quality software development is needed. Given these activities, I would rate this project as very difficult and hence suits people who can handle advanced software engineering/system building and are very serious about their project. The research and development component can be appropriately structured based on the interest of the student. This specification can be adapted for various MSc cohorts.

Special Requirements :

None.

Developing a spatial skills training app (ID: 7582)

Owner/Supervisor: [Quintin I Cutts](#)

** Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

Spatial skills have been shown to be linked to success in STEM subjects, and can be developed with training. The skills development consists of a number of exercise formats, some of which are multiple choice based, and some of which require drawing isometric and orthographic projections of shapes. The purpose of this project is to develop a web app to support the setting and attempting of skills development exercises of these kinds. This app could be ground-breaking across the University and more widely as a fast way of developing these important skills.

Special Requirements :

None.

Development and validation of an analytical cost model for NoSQL database operations (ID: 7221)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

In the last few years we have witnessed a proliferation of new database management systems, collectively referred to as NoSQL (short for “Not only SQL”) databases, going beyond the design requirements (and limitations) of traditional relational database management systems. Several of these systems (including, but not limited to, Amazon’s DynamoDB[1], Apache HBase[2], Apache Cassandra[3], etc.) are now struggling to overcome their own limitations in query expressiveness and performance for complex workloads (join/rank-join queries, analytics, etc.). In order to produce an optimal query execution plan, one needs to first understand the inner workings of these systems to qualify and quantify the cost factors contributing to the query processing cost, in terms of processing time, main memory utilization, disk I/O bandwidth, network usage, etc.

The goal of this project is to produce an analytical cost model for a number of such systems. Students should first gain an in-depth understanding of the design and implementation of the systems at hand and identify core abstractions underlying them. Only then they should embark on developing and validating, through extensive experimentation, an analytical cost model taking into account all above mentioned cost factors.

[1] G. DeCandia, et al. Dynamo: amazon's highly available key-value store. In Proc. 21st ACM SIGOPS Symposium on Operating Systems Principles (SOSP), 2007.

[2] <http://hbase.apache.org/>

[3] <http://cassandra.apache.org/>

Special Requirements :

Solid background in statistics and probability theory. Good knowledge of principles of operating systems, distributed systems, and database management systems. Excellent knowledge of the Java programming language.

Development and validation of an empirical cost model for NoSQL database operations (ID: 7220)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

In the last few years we have witnessed a proliferation of new database management systems, collectively referred to as NoSQL (short for “Not only

SQL”) databases, going beyond the design requirements (and limitations) of traditional relational database management systems. Several of these systems (including, but not limited to, Amazon’s DynamoDB[1], Apache HBase[2], Apache Cassandra[3], etc.) are now struggling to overcome their own limitations in query expressiveness and performance for complex workloads (join/rank-join queries, analytics, etc.). In order to produce an optimal query execution plan, one needs to first understand the inner workings of these systems to qualify and quantify the cost factors contributing to the query processing cost, in terms of processing time, main memory utilization, disk I/O bandwidth, network usage, etc.

The goal of this project is to produce an empirical cost model for at least one of these systems. Students should first gain an in-depth understanding of the design and implementation of the systems at hand and identify core abstractions underlying them. They should then instrument their source code implementation in order to produce simple yet accurate profiles of the execution of various operations on these systems, ranging from simple point queries, scans, insertions, and deletions, to more complex operations such as evaluation of SPJ queries. The end result will include both extensive experimental evaluation results and an analytical approach translating these results into a cost model.

[1] G. DeCandia, et al. Dynamo: Amazon's highly available key-value store. In Proc. 21st ACM SIGOPS Symposium on Operating Systems Principles (SOSP), 2007.

[2] <http://hbase.apache.org/>

[3] <http://cassandra.apache.org/>

Special Requirements :

- * Excellent knowledge of the Java programming language
- * Good understanding of database system and distributed system internals

Development of Deployable Forensic Solutions for Programmable Logic Controllers (ID: 7562)

Owner/Supervisor: [Chris Johnson](#)

* Suitable for both IT/SD and CS+.

Description:

Cyber forensic capabilities are an essential component in incident response for all organisations and businesses. This is particularly the case within critical infrastructure sectors where there are mission and safety critical systems, such as industrial control systems (ICS) and supervisory control and data acquisition (SCADA) systems in sectors such as civil nuclear, aviation, and defence. This project expands on previous work that has been carried out in developing technical forensic solutions for programmable logic controllers

(PLCs), a device that is built to operate in safety-critical environments and that's found throughout ICS.

Special Requirements :

none

Does reputation matter ? Analysing how users' reputation affects collaboration in Q&A websites (ID: 7260)

Owner/Supervisor: [Iulia Popescu](#)

* Suitable for both IT/SD and CS+.

Description:

The aim of this project is to understand what are the different factors that enhance the user's reputation in Q&A communities. Initially, by looking at the content posted on Q&A websites, we intend to investigate whether the style of writing influences users' reputation. Then, we can explore how the level of engagement in the Q&A community impacts the reputation score.

Project plan:

- Learn how the user's engagement and linguistic factors (extracted from the content posted) influence the reputation score.
- Conduct the study in both science oriented Q&A platforms and non-science / everyday life Q&A platforms.
- Build a dataset with data acquired from Q&A websites (e.g. the communities operated by StackExchange <https://data.stackexchange.com/>)
- Extract textual data and transform it into vectorised attribute form (i.e. each text sample in the dataset should be represented as a feature vector).
- LIWC (<http://liwc.wpengine.com/>) could be used to analyse the text and build a "description" of the writing style of the user.
- Then, statistical analysis techniques can be applied to understand the factors that influence the user's reputation in Q&A communities.

This project will be co-supervised by Dr. Nikos Ntarmos.

Special Requirements :

None.

Drawing Editor for Digital Circuits (ID: 7325)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

General purpose drawing tools are generally not well suited to drawing digital circuit diagrams: they are hard to use, yet produce mediocre results.

The aim of this project is to design and implement, in Java, an interactive tool for drawing circuit diagrams, and for exporting them in a form that can be incorporated in documents. A diagram contains boxes (which represent either primitive components or sub-circuits); wires (which connect ports on the boxes); and labels and annotations (text which may be attached to a wire, or which may be freely placed anywhere).

The tricky aspect of editing circuit diagrams is to find a good placement for the boxes, and then neatly place the wires connecting them. People are better at this than algorithms, so the approach to be taken in this project is to provide good facilities for moving boxes around, and rearranging wires, but not to try to do these tasks automatically.

The program should allow the user to produce a good layout for a circuit, should save a design to a file and be able to reload the file, and should export the design to a standard vector graphics format (such as SVG).

The project requires good programming ability in Java, especially with GUIs, but doesn't require specialised knowledge of computer hardware.

Product: an interactive graphical editor for digital circuits, with the ability to edit layouts, write and read the data structures into files, and translate the schematic into vector graphics.

Special Requirements :

Requires good programming ability, and ability to work with graphics in a GUI.

Dynamic discovery of vulnerable nodes in social networks (ID: 7506)

Owner/Supervisor: [Inah Omoronyia](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

One major challenge in social network analysis and link discovery is the discovery of hidden structures. Examples of such structures are users that might be vulnerable to privacy violations as a result of their rather liberal behavior in disclosing sensitive information. Existing privacy preserving and encryption techniques might be useful in ensuring that such disclosed information does not end up with a malicious user outside a network. But such techniques cannot prevent sensitive information disclosed to a known user in a group from further being passed on to other users in the group, whom might themselves belong to other groups that are adversarial.

The goal of this work is to devise a preemptive approach to preserving the privacy of vulnerable users in information flow networks. The objective of this project is to investigate a technique for identifying vulnerable users in online social networks and chat rooms based on the entropy of the associated network.

Special Requirements :

None.

Emulating Glasgow's First Computer (ID: 7523)

Owner/Supervisor: [Jeremy Singer](#)

* Not suitable for CS+ students. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The first computer at Glasgow University was the English Electric DEUCE machine. Installed 60 years ago (in 1958) it was based on an original design by Alan Turing. We have some of the original documentation for the DEUCE, as well as a Java emulator program.

The aim of this project is to build a web-based front-end for the DEUCE. This will feature an attractive graphical console, based on the original desk console, see e.g.

[this photo](#)

The key aim is to be able to run DEUCE programs interactively via a web page. Further aims include the ability to 'program' and 'single step' the machine.

Skills required:

- - graphic design

- - web development
- - interest in Computer Architecture
- - interest in Computing History

Special Requirements :

None.

End to end learning for survival prediction from genomic alterations (ID: 7115)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Somatic alterations such as single nucleotide point mutation, insertions/deletions and structural variants are hallmarks of cancer. They are routinely examined by performing whole exome/genome sequencing on patient biopsies. Due to the complexity of the diseases, it is still challenging to predict patient outcomes from somatic alterations. In this project, we will develop a deep learning based end to end learning model to test how well we can predict survival in cancer patients.

Special Requirements :

None.

End to end learning for survival prediction from tissue images (ID: 7101)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Tissue images from cancer patients such as whole section images and tissue microarrays contain rich information about cellular features and interactions among different cell types.

In this project, we will develop a deep learning based end to end learning model to predict survival from patient tissue image data.

Special Requirements :

None.

Enforcing the Grammar of Graphics with Types (ID: 7435)

Owner/Supervisor: [Michel Steuwer](#)

* Suitable for both IT/SD and CS+.

Description:

The Grammar of Graphics is a way to structure the creation of plots. This theory has been developed by Leland Wilkinson and there exist multiple popular plotting libraries following this approach (e.g. ggplot2 in R).

While the theory nicely explains how plots can be understood as small composable components the existing implementations are all made in dynamically typed languages which makes the APIs harder to use, as no static type system can help the programmer to correctly use these APIs. A statically typed version would empower the IDEs to guide the programmer to correctly use the API (e.g. by code completion) and the compiler would help to spot errors early.

In this project, you are going to implement a statically typed version of the ggplot API in a suitable programming language (e.g. F# or Scala).

Special Requirements :

None.

Estimating subclonality in pancreatic cancer from whole exome sequencing data (ID: 7106)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Large-scale genomic sequencing studies such as the International Cancer Genome Consortium (ICGC) and the Cancer Genome Atlas (TCGA) are mapping the genomes of multiple cancer types. The data from these studies provides strong evidence for the high level of heterogeneity within tumours.

This project will focus on studying intratumor heterogeneity in pancreatic cancer, an aggressive type of cancer with limited therapeutical success. We will

- 1) Collect the publicly available whole-exome sequencing and SNP array data from TCGA
- 2) Estimate the copy number profiles from SNP arrays
- 3) Identify subpopulations of cancer cells from whole-exome sequencing data
- 4) Correlating features of results in steps 2 & 3 with patient survival data to exploring clinical utility.

Special Requirements :

None.

Evaluation of Firmware Security Capabilities for Industrial Control System Components (ID: 7564)

Owner/Supervisor: [Chris Johnson](#)

* Suitable for both IT/SD and CS+.

Description:

Components found within safety-critical systems don't typically run standard operating systems such as Windows. Instead, they contain firmware, which acts as the layer between the logic that is being executed, and the hardware components. For this reason, security solutions are required to mitigate the possibility of an attack on firmware. This project involves expanding firmware security solutions for programmable logic controllers (PLCs), and exploring real-world safety-critical environments that these solutions can be deployed in.

Special Requirements :

none

Evaluation of the ARM Cortex MPS2 FPGA board #1 (ID: 7367)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Choose this if you like: ARM processors, FPGAs, embedded systems

The aim of this project is to evaluate the ARM MPS2 FPGA board, which lets you create a System on Chip with one or more ARM Cortex M-series processors with FPGA fabric on which dedicated hardware can be created.

The purpose of the project is to evaluate how easy the provided toolchain and workflow is for use by software developers with limited hardware knowledge.

Special Requirements :

We will provide the FPGA board.

Evaluation of the ARM Cortex MPS2 FPGA board #2 (ID: 7368)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Choose this if you like: ARM processors, FPGAs, embedded systems

The aim of this project is to evaluate the ARM MPS2 FPGA board, which lets you create a System on Chip with one or more ARM Cortex M-series processors with FPGA fabric on which dedicated hardware can be created.

The purpose of the project is to evaluate how easy the provided toolchain and workflow is for use by software developers with limited hardware knowledge.

Special Requirements :

We will provide the FPGA board.

Evaluation of the PYNQ FPGA board #1 (ID: 7319)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Choose this if you like: FPGAs, embedded systems, embedded Linux

The aim of this project is to evaluate the PYNQ FPGA board, which contains a chip combining and ARM processor running Linux with FPGA fabric on which dedicated hardware can be created. The purpose of the project is to evaluate how easy the provided toolchain and workflow is for use by software developers with limited hardware knowledge.

Special Requirements :

We will provide the FPGA board.

Evaluation of the PYNQ FPGA board #2 (ID: 7345)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Choose this if you like: FPGAs, embedded systems, embedded Linux

The aim of this project is to evaluate the PYNQ FPGA board, which contains a chip combining and ARM processor running Linux with FPGA fabric on which dedicated hardware can be created. The purpose of the project is to evaluate how easy the provided toolchain and workflow is for use by software developers with limited hardware knowledge.

Special Requirements :

We will provide the FPGA board.

Evaluation of the PYNQ FPGA board #3 (ID: 7346)

Owner/Supervisor: [Wim Vanderbauwhede](#)

* Suitable for both IT/SD and CS+.

Description:

Choose this if you like: FPGAs, embedded systems, embedded Linux

The aim of this project is to evaluate the PYNQ FPGA board, which contains a chip combining and ARM processor running Linux with FPGA fabric on which dedicated hardware can be created. The purpose of the project is to

evaluate how easy the provided toolchain and workflow is for use by software developers with limited hardware knowledge.

Special Requirements :

We will provide the FPGA board.

Extending student-project allocation software (ID: 7344)

Owner/Supervisor: [William Pettersson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A common problem in many universities is the assignment of students to courses with limited places available. Students would be able to rank their desired courses from most to least preferable. They would then be assigned to courses in a manner which would try to give students their most preferred courses whilst still treating all students equally. Traditionally such allocations would have been completed by hand, but in recent times algorithms to solve the problem have been developed. Some of these are even in use by various schools in the University of Glasgow.

However, often very simple changes to the student preferences can have complex effects on the outcomes. For instance, students may be allowed to mark a non-preference for courses, making those courses not acceptable to the given students. And some students may want to rank certain courses identically. Separately, each of these two conditions can be accepted without significant problems, but if both conditions are applied together then the problem can become significantly hard to solve.

The purpose of this project is to develop new models and software which can be extended to allow for more varied constraints on behalf of a school or department. This will involve discussions with clients to gather requirements, as well as the development of new tools to facilitate these requirements as they relate to matching students to courses.

This project would suit a student with a strong background in algorithms.

Special Requirements :

None.

Facial palsy assessment App - with NHS (ID: 7612)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+.

Description:

Facial palsy is a very common cranial nerve paralysis. Most common cause is idiopathic Bell's palsy which recovers in 95% of cases. However there are

other causes such as stroke, trauma, and malignancies which may not recover fully. It is important to know if the paralysis is complete or partial from the beginning so that we can plan the management strategies and monitor progress.

The current method of monitoring is a subjective scoring system called House-Brackmann system (Brian F. O'Reilly, John J. Soraghan, Stewart McGrenary, and Shu He., Objective Method of Assessing and Presenting the House-Brackmann and Regional Grades of Facial Palsy by Production of a Facogram, Otology & Neurotology 31:486-491 2010, Otology & Neurotology, Inc.) . However, it is very much dependent on observer and there is a lot of variability. It will be useful to have an objective assessment system to monitor progress. There were efforts to make such a system using digital imaging and monitoring change in pixels with facial movement, but were very cumbersome & involved multiple imaging at the medical illustration dept. and difficult calculations. So this was not universally accepted.

With the modern smartphones and use of AI, we can develop a very user friendly but validated scoring system. Objective of this project is to develop a smartphone app for objectively assessing Facial Palsy. This will help clinicians and patients to objectively assess the level of palsy at the onset and monitor the progress.

This project will be co-supervised by DR Arun Iyer, Consultant ENT surgeon / Otolologist, Royal college of physicians & surgeons of Glasgow

Special Requirements :

None.

Finding patterns in molecular names (ID: 6985)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Any molecule will have several names, all conforming to different naming conventions. For example, paracetamol also goes by the names 103-90-2 [RN], 203-157-5 [EINECS], 4-(Acetylamino)phenol, 4-13-00-01091 (Beilstein Handbook Reference) [Beilstein], 4-Acetaminophenol, 4'-Hydroxyacetanilide, 4-Hydroxyanilid kyseliny octove [Czech], and Acetamide, N-(4-hydroxyphenyl)- [ACD/Index Name]. Interesting features occur in each of these names. Faced with a group of molecules, it would be interesting to use all of their synonyms to find any textual features that were particularly enriched in a group. For example, perhaps paracetamol is grouped with other molecules that include the term 'phenol'. In this project, you will make use of a standards database coupled to a web resource such as

chemspider to mine the synonyms of groups of molecules for repeated terms that appear more often than would be expected by chance.

Special Requirements :

None.

Gamifying concurrency teaching (ID: 7156)

Owner/Supervisor: [John Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Students often struggle to understand correct patterns for dealing with multithreaded behaviour in conventional languages such as Java. This project will look at building a game around solving concurrency problems to help give students better insight into concurrent programming.

Specifically, it will involve building a game around the puzzles given in the open-source text book "The Little Book of Semaphores" <http://www.greenteapress.com/semaphores/>

The implementation will need to have a compelling and immediate graphical metaphor, the details of which will be up to the student, along with a testing mechanism that can reliably detect incorrect patterns (either stochastically or by more sophisticated search methods) and visually demonstrate their weakness.

Implementation can be in a language of the student's choice.

Special Requirements :

None.

GCViz - Visualization of Garbage Collection (ID: 7525)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Garbage collection (GC) is automatic memory deallocation, which takes place in managed runtimes like the Java virtual machine. There are many varieties of different GC algorithms, which organize the runtime heap in different ways. As a teaching tool, it would be extremely useful to have a simple visualizer that enables students to set up toy heaps, and run various GC algorithms to see their different behaviours.

This project involves building a Model-View-Controller architecture to implement a GC visualization tool. The experimental heap and sample garbage

collection algorithms will be provided by the MMTk library - see <http://jikesrvm.org/MMTk> - This will constitute the model. The project has to provide a visual representation of the model (i.e. the view) which shows how the heap is structured in a simple way. The project also has to provide a user interface (i.e. the controller) which allows users to add new items into the heap, to perform GC with various algorithms, to resize the heap, etc.

The underlying MMTk library is open source code, and it would be expected that the GCViz visualization tool would also be open source, and possibly contributed back to the main MMTk project if it is sufficiently useful and robust.

Useful background info:

- * The garbage collection page
at <http://www.cs.kent.ac.uk/people/staff/rej/gc.html>

- * Garbage collection article on wikipedia
http://en.wikipedia.org/wiki/Garbage_collection

- * Textbook
<http://www.amazon.co.uk/Garbage-Collection-Algorithms-Automatic-Management/dp/0471941484>

Special Requirements :

None.

Geo-distributed analytics (ID: 7234)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

Geo-distributed analytics have lately emerged as a major field of research and development, backed by the interests of and issues faced by several very large institutions/corporations. The source of related problems is the large volume of data generated at the edges of the network (e.g., in datacentres around the globe). Up to a couple of years ago, the state-of-the-art consisted of transferring relevant data from all datacentres to a central location for analysis and processing. This can be highly impractical now, and almost impossible in the years to come as the projection is that in a few years less than a handful of such computations would be enough to saturate the backbone of the internet.

The goal of this project is (a) to simulate a geo-distributed environment on top of a set of local computers, (b) to examine the applicability of currently available Big Data processing software (HDFS, Spark, etc.) for use in such

scenarios, and (c) to identify the pain points of such deployments and propose solutions.

Special Requirements :

- * Ansible
- * Linux tc/tcng/iptables
- * Java and/or Scala
- * Knowledge of principles of distributed/database/operating systems

Gesture based Interactive search system (ID: 7331)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A number of IoT enabled devices are being exploited. For example, a public display screen where useful information is provided in an entrance of a building (e.g., School of computing science foyer). Unfortunately such information is static and curated by an administrator. The challenge is to personalise such messages so that people with different needs can be catered effectively. For example, in a medical surgery deliver information based on surgeon's context, providing surgery based information on a public screen where it is possible to select, scroll using gestures. IoT linked devices provide new opportunities to engineer such interactive information solutions.

In this project we will develop and implement an Interactive search environment, using Microsoft Kinect, Leap Motion, depth cameras, public display screens. Kinect provides opportunities to track and gather user gestures. Leap supports flexible gestures. Camera can be used to detect user hands and depth. We will deploy the developed set-up in a public space, for example level 2 SAWB, and evaluate the effectiveness of the solution with occupants of that level.

Special Requirements :

None.

Gesture Sensor Attention (ID: 7358)

Owner/Supervisor: [Euan Freeman](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Mid-air gestures are hand or finger movements used to control a user interface. A problem with mid-air gestures is that users can't see the space where gestures are recognised, so the interaction area is ambiguous. Feedback can help users by telling them when their hand can be seen and is in a good position. This type of feedback is called "system attention" feedback, because

the user interface tells the user that it can see them and it is paying attention to their actions.

This project will design and implement a method for showing system attention, using simple light sources (LEDS). For example, the lights might turn on or become brighter when the system recognises the user's hand approaching the input sensor. We will vary the distance when feedback is given: e.g., at 15cm, 10cm, 5cm from the middle of the gesture sensor. The user evaluation will investigate how the feedback influences user behaviour: e.g., does the user stop moving their hand as soon as feedback is given, or do they continue reaching forward? This is an interesting question because if users stop as soon as feedback is given, they might not end up in the 'best' place to perform gestures; if this is the case, it will be better to delay feedback to force them to find a better hand position.

A Leap Motion sensor will be provided for implementation; this is a hand and finger tracker with a good SDK for most programming languages. Arduino-controlled LEDS will also be provided; these plug into a USB port and can be controlled via simple commands. The key part of this project is the user evaluation, because this will contribute to an ongoing research project.

Special Requirements :

A Leap Motion hand tracker and Arduino-controlled LEDS will be provided. These work with most programming languages, but something like Java, Python or C# is preferred.

Go Problem Solver with Alpha-Beta Tree Search (ID: 7327)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

Go is a board game for two players. It has simple rules but complex tactics and strategy.

Several aspects of the game can be studied through problems. In a typical problem, you are given a board position and a specific goal (e.g. black to play and kill white's group), and the aim is to find a sequence of moves that achieves the objective. There are several phases of the game (including life and death fights, tesuji problems, and endgame). There are large collections of problems that focus on each of these game phases.

A good method for solving problems is in tree searching, using algorithms such as minimax and alpha beta. This involves generating sequences of moves, and choosing the line that gives you the best result.

A useful tool for improving tactical skill is an interactive problem solver: a computer program that displays a problem, and allows the user to play through a solution. To do this, the program needs to be able to solve the problem itself, so that it can respond correctly to any moves that the user makes.

The aim of this project is to design and implement such a program. The software needs a graphical user interface, to enable the user to select a problem, to see the current board position, and to make a move.

The program needs to be able to determine the set of legal moves from any board position. It also needs a tree searching engine that enables it to solve problems; the minimax algorithm and the alpha-beta algorithm are suitable. The software needs specialised modules (a move generator and a position evaluator) to enable it to handle a particular class of problem.

The software will be evaluated by testing it with a set of problems and measuring its success rate.

Product: an interactive program, with GUI, that presents a go problem on a board, and interacts with the user as they explore the solution space of the problem. The program uses tree search algorithms to find good moves.

Special Requirements :

None.

Graphical editor for bigraphs (ID: 7198)

Owner/Supervisor: [Michele Sevegnani](#)

* Suitable for both IT/SD and CS+.

Description:

Bigraphs [1] are a universal formalism for modelling interacting systems that evolve in time and space. The formalism is based on a graphical model that emphasises both locality and connectivity.

The goal of this project is to implement a graphical editor for bigraphs.

[1] Bigraphs with sharing. Michele Sevegnani and Muffy Calder in Theoretical Computer Science, Volume 577, pp 43-73, Elsevier, April 2015

Special Requirements :

None.

Graphical Passwords (ID: 7398)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Graphical passwords use images rather than text when a user logs in. They choose 3 images when they register from the system. When they log in they are shown many images and have to recognise the three that they originally chose.

Registration process: The user is shown 60 images on three pages and must choose 3 images. They can come from any of the pages and don't have to be one from each page. The 60 images are chosen from a collection of 1000 images that I have. The user's choice is noted and 17 other images are also chosen and associated with the user. These are called distractor images. These 17 images can either (i) be chosen from the original 60 images, and are thus images that the user has rejected, or (ii) chosen from the rest of the collection and are images the user has not seen. The system should record the time taken to choose each image.

Login process: The user is shown 20 images, the 3 that they chose and the 17 distractor images. They should be arranged in a different random order each time the user logs in. The user must choose the correct images, in the correct order, to succeed. The system should record the order that the images are chosen and the outcome, whether the login was successful. It should also record the time taken to select each image.

Experiments: The recorded data can be processed to see the choice of distractor images, either (i) or (ii) above has any effect on the success and time of login. The recorded data can also show reasons for failures, such as the correct images in the wrong order, how many images were wrong, etc.

User Study: Users can try out the system and we can find out which approach they preferred.

This project will involve a significant amount of programming, as well a user study.

Only choose this project if you are taking MSc Information Security or MSc IT Cyber Security.

Special Requirements :

None.

**Graphs through time: Adding time awareness to Neo4j
(ID: 7219)**

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

Neo4j[1] is an open-source NoSQL (graph) database written in Java. It supports full ACID[2] transactions, uses a custom disk-based storage engine to achieve durability, and can provide high availability through replication when distributed over multiple computing nodes. As it is purpose-built to handle graph semantics, it can scale to several billion vertices, edges, and properties. At the same time, its graph traversal framework allows for fast graph query processing.

Currently, Neo4j has no notion of time, other than annotating nodes and edges with a TIME property. This approach does allow Neo4j to support time-dependent query processing, but is surely not optimal. For example, queries engaging the complete graph at a given point in time, or even simpler queries with time constraints, need to scan through a large part of the datastore in order to produce their result set.

The goal of this project is threefold: first, to gain a deep understanding of the design decisions and implementation details of Neo4j; second, to explore indexing structure and techniques, expediting time-dependent queries; and third, to extend Neo4j's storage, caching, indexing, and query processing components with versioning and/or time-aware primitives.

[1] <http://www.neo4j.org>

[2] <http://en.wikipedia.org/wiki/ACID>

Special Requirements :

This project requires excellent knowledge of the Java programming language, and a deep understanding of database system concepts, data structures, and operating/database systems design and implementation principles.

Gravitational N-Body Problems (ID: 7395)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

Determining the orbits of 3 or more bodies with mutual gravitational attraction cannot be solved analytically and so numerical solutions have to be found.

This is also an interesting problem in chaos theory. The positions of the bodies can only be known for a limited amount of time in the future, no matter how accurate their positions and velocities are known at the start of the calculations.

It is also tricky to visualise what will happen to the bodies as their orbits develop through time. Initial orbits are interesting to watch, but it is important

to also see what happens when the program is run many times with slightly different starting conditions. We might not be able to predict exactly what will happen, but we can assign probabilities to a range of different things that can happen.

This project will start with the restricted three-body problem, with 2 massive and one tiny object orbiting together. This is important in looking at the Earth-Moon-Spaceship situation. If time permits it can then advance to more general N-body problems.

You should only choose this project if you are good at maths.

Special Requirements :

None.

Haptic feedback for driving simulators (ID: 7417)

Owner/Supervisor: [Graham Wilson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

As in-car infotainment systems become more complex, car manufacturers need to better ways of providing rich and informative feedback to the driver. Visual attention is needed for driving, and so vibrotactile feedback, provided from the steering wheel, may provide a solution. This project will involve either programming extensions to simulator Assetto Corsa in Python, to provide audio and vibrotactile feedback cues while driving, or creation of a simple Arduino-based hardware platform for connecting novel vibration motors to a steering wheel and PC. The project will require existing knowledge of Python or hobbyist electronics, and will involve both software/hardware development and experimental user testing, to measure the perception of feedback cues

Special Requirements :

None.

Humanising traditional music playback (ID: 7153)

Owner/Supervisor: [John Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

ABC notation format provides a convenient and compact notation for traditional/folk music. There are several tools for playing back ABC music, but the results are often very stilted and artificial.

This project will investigate generating more "humane" automatic playback of

traditional music in specific styles (e.g. Scottish solo fiddle music) by applying rule based transforms to the notation. The result will be an application which can take an ABC notation file, and generate a playable file (e.g. MIDI or direct audio) which has more realistic nuance and rhythm.

This project will require musical knowledge, and ideally interest in traditional folk music.

Special Requirements :

None.

Hybrid Music Recommendation (ID: 7075)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

How do you integrate multiple data sources in music recommendation? This project will evaluate state-of-the-art recommendation methods based on neural networks and matrix factorization.

The project will include:

- * Literature study and survey on music recommendation and deep neural network for recommendation
- * Formulation of a unified music recommendation solution/model based on state-of-the-art.
- * Data collection: Gather and analyse existing data sources (e.g. last.fm and the million song dataset)
- * Proof of concept interface for evaluation and real-time data collection for inferring context and user aspects
- * Hybrid semantic indexing using deep neural networks – integration of objective and subjective data sources, and dynamic ranking of tracks based on user profile and context.
- * Specific recommendation engine based on concepts such “diversity”, “familiarity”, “serendipity” etc to control exploration (give the user new unseen music which is believed to be relevant) vs. exploitation (return mostly known stuff)
- * User evaluation
- * Qualitative and quantitative/statistical visualisation and analysis of results

We will look to involve relevant industrial partners to carry out an evaluation on a large scale.

Special Requirements :

None.

Identifying estrogen receptor binding complex and its role in gene regulation from ENCODE data (ID: 7107)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Oestrogen receptor- α (ER) is the key transcription factor in the majority of breast cancers and its target genes dictate cell growth and endocrine response. ER does not act alone. It recruits a network of transcription factors to form different types protein complex, yet the compositions and roles for these complexes in dictating gene expression are limited.

In this project, we will address this question by study ChIP-seq data in breast cancers cell lines from the ENCODE project. In particular, we will

- 1) use clustering analysis to identify ER protein complexes
- 2) study variations in the compositions of ER complex in different breast cancers lines.
- 3) identify ER complexes role in gene expression by correlating complexes activities with expression data

Special Requirements :

None.

Identifying evolution signatures from whole genome sequencing data (ID: 7108)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Intratutormal heterogeneity analysis from whole genome sequencing data reveals the number of subclonal cancer populations and the population frequencies of somatic alterations in the genome.

In this project, we will develop a nonnegative matrix factorization (NMF) based approach to extract underlying evolution signatures from the population frequencies of somatic alterations across 2778 cancer genomes.

Special Requirements :

None.

Implementing a Crypto-Currency (ID: 7402)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

This project will implement a peer-based crypto-currency. The project will have to implement all the necessary components, including the tamper-proof distributed ledger, a mechanism for distributing transactions to all the participating nodes, a mining mechanism and so on. It should also implement a test harness so that a number of participating nodes can be created and run on a single computer. This will allow the basic mechanism to be tested. It should also implement an API that can be used by customer-facing wallet software. You are not expected to produce a fully distributed system capable of being released in the wild.

Only choose this project if you are taking MSc Information Security or MSc IT Cyber Security.

Special Requirements :

None.

Implementing a Crypto-Currency Wallet (ID: 7404)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A crypto-currency wallet is a user-focussed interface to a crypto-currency and deals with key management issues and synchronisation between hot and cold wallets. This project will run experiments to determine the usability, as well as testing the software to make sure that it is safe from attack. It should also implement a toy crypto-currency (just the user interface) for testing purposes. If another student is implementing the crypto-currency then it would be useful to make sure that the two pieces of software work together, but if not, this project will be able to work by itself.

Only choose this project if you are taking MSc Information Security or MSc IT Cyber Security.

Special Requirements :

None.

Improved Student Learning with Machine Teaching/Learning (ID: 7074)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

This project will investigate and build a web-based system for sequentially

optimizing student learning using Bayesian optimization.

Elements:

- A literature review of machine teaching based on [1].
- Design and implementation of a web-based system and interface for presenting questions or instructional examples to the student [based on 2].
- Application of Bayesian optimization for sequentially choosing the specific instruction/example/question to present in order for the subject to learn in an “optimal” way.

Literature:

[1] Machine Teaching: An Inverse Problem to Machine Learning and an Approach Toward Optimal Education, <http://pages.cs.wisc.edu/~jerryzhu/pub/MachineTeachingAAAI15.pdf>

[2] Optimizing Instructional Policies, <http://papers.nips.cc/paper/4887-optimizing-instructional-policies.pdf>

Special Requirements :

None.

Increasing interaction at public displays using mobile devices (ID: 7369)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Pairing mobile devices with public displays can increase engagement with display content, but the pairing process is often difficult or cumbersome. This project will explore different methods of enabling interaction between devices to improve interaction.

Special Requirements :

None.

Input Techniques for Mobile VR (ID: 7370)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Mobile VR headsets like the Gear VR offer the opportunity to use VR in more places than ever before. Gear VR requires no external sensors or any purpose-built setting to start a VR experience. This strength makes it easier to use VR

anywhere, but also means that input capabilities and tracking capabilities are severely limited.

This project will explore how input can be designed given the limited sensing opportunities available. This will make use of the Gear VR with controller and look for ways of exploiting all the sensors built into the device such as gaze and head gestures.

Special Requirements :

None.

In-silico fragmentation for natural product detection (ID: 7007)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Linking genomic features (presence of groups of genes) to the chemical products they produce is a key step in the discovery of new drugs (produced by e.g. bacteria). Algorithms exist to produce putative chemical structures from the genomic features, but currently it is not possible to map these structures to the chemical information observed when the bacteria are grown in the laboratory. In this project you will create a pipeline from a series of published tools that will allow the output of the genetic algorithms (predicted structures) to be virtually fragmented (using e.g. metfrag) and then compared to large datasets of molecular fragment data.

This project would be particularly suitable for someone with a background in chemistry / biology and/or someone interested in working within that area in the future.

Special Requirements :

None.

Interactive Graphical Proof Editor for Natural Deduction (ID: 7329)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

Natural deduction is a formal system for logical inference in propositional logic and predicate logic. It uses a set of inference rules that state precisely a set of required assumptions, and a conclusion that may be inferred if those assumptions are established. An advantage of natural deduction is that the inference rules are syntactic: they are defined precisely in terms of the form of the

logical expressions, so it is relatively straightforward to write an algorithm that determines whether an inference step is correct.

Natural deduction proofs are often written in a two-dimensional graphical layout, with a horizontal line across the page: the assumptions are written above the line and the conclusion is written below it. Since the assumptions may themselves be conclusions of other inferences, we end up with a tree structured layout containing many lines and inferences. The geometric layout helps to show the overall structure of the logical argument.

The aim of this project is to develop a program that allows the user to enter and edit a natural deduction proof and to check it. The software should show the graphical layout of the proof, and it should provide a graphical user interface that makes it easy to enter and edit the proof. The project involves both some GUI programming and some manipulation of the tree structures that arise in logical proofs.

You don't need to know about natural deduction or mathematical logic; this is straightforward material that you can learn at the beginning of the project. Suitable programming languages include Haskell and Java.

Product: a program that checks natural deduction proofs, and displays the structure of the proof graphically.

Special Requirements :

None.

Interactive Visualization of Statistical Models (ID: 7155)

Owner/Supervisor: [John Williamson](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Interactive representations of probabilistic systems. This involves building interactive visualisation software to probe the behaviour of statistical models. This, for example, might involve probing regression models to determine their sensitivity to outliers and showing the output as an animation. This involves instrumentation of probabilistic models, and interactive perturbation of their state.

Special Requirements :

None.

Interactive Water Feature for Public Squares (ID: 7371)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The physical elements of urban spaces that lead to better public spaces are well understood. The ideal placement of trees, benches, and water features has been extensively studied, and has informed the rules of city planning since William Whyte's seminal work on The Project for Public Spaces.

As technology integrates itself into daily life, we must also gain an understanding of the digital elements of public spaces. How does digital interactivity change the way physical elements shape a space? This project will explore how adding playful interactivity to a water feature changes a public space. This will make use of the PumpSpark Fountain Development Kit (<http://dl.acm.org/citation.cfm?id=2598599>) to create an interaction fountain.

Special Requirements :

Hardware for this project (and paddling pool) will be provided. Knowledge of micro-controllers not required, but would be helpful.

Internet Traffic Analysis and Characterisation (ID: 7053)

Owner/Supervisor: [Dimitrios Pezaros](#)

* Suitable for both IT/SD and CS+.

Description:

The design and provisioning of computer networks relies on the accurate understanding of the characteristics of network traffic, such as, e.g., the dominant applications, protocols, packet size distribution, burstiness of the traffic/packet arrival process, etc. However, the Internet continuously evolves in scope and complexity, much faster than our ability to understand and predict it.

Over the past few years, a wealth of research has focused on the characterisation and classification of Internet traffic using statistical and machine learning algorithms, see, e.g., [1][2][3].

We have acquired large traffic traces from representative Internet backbone links. The purpose of this project will be to study these datasets, process them, and apply appropriate analysis and classification algorithms to identify trends in Internet traffic and study how it evolves. A major objective of the project will be to evaluate the validity of existing algorithms, try new ones out, and compare and contrast the applicability/efficiency of diverse data analysis approaches.

This project is intended for students with interests in (big) data analysis, classification, statistics and machine learning, as well as in systems and networking.

- [1]. Internet traffic classification overview: <http://www.caida.org/research/traffic-analysis/classification-overview/>
- [2]. Internet Traffic Classification Using Bayesian Analysis Techniques: <http://www.cl.cam.ac.uk/~awm22/publications/moore2005internet.pdf>
- [3]. Internet Traffic Classification Demystified: On the Sources of the Discriminative Power: http://conferences.sigcomm.org/co-next/2010/CoNEXT_papers/09-Lim.pdf

Special Requirements :

Product: data analysis and classification software

Suggested Platforms: Unix/Linux

Suggested Languages and Tools: R/Matlab, bash/C

Interval finder: Segmenting Strava Activities (ID: 7005)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Strava is a web application popular with runners and cyclists. Often runners will perform interval programmes (3 mins fast, 1 min recovery times 10 etc) but Strava doesn't provide a facility to be able to automatically segment the activity and allow athletes to track their performance changes over long time periods (e.g. the pace for the three minute intervals over a period of months). In this project you will use data extracted from Strava to develop an automatic segmentation method that can extract the intervals from an exercise. You will research different statistical and machine learning algorithms for this job, and compare their performance against labeled data (i.e. manually labelled data).

Special Requirements :

None.

Inverse bio-engineering --- a computational approach (ID: 7069)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+.

Description:

Can you from a sample of a biological material (e.g. a microscopy image) determine the process parameters (e.g. chemical components, light and

temperature) which produced it?

This project will investigate software pipelines and machine learning methods to aid biomedical engineers to invert their normal design process to speed up the exploration and design of novel biomaterials.

This project is in collaboration with the Biomedical Engineering department and it will be tailored to the successful candidate.

Special Requirements :

None.

Investigating partially observable controllers (ID: 6963)

Owner/Supervisor: [Gethin Norman](#)

* Suitable for both IT/SD and CS+.

Description:

Recently we extended the model checking tool PRISM (<http://www.prismmodelchecker.org>) to allow for the analysis of partially observable controllers or strategies.

Partially observability is import since for example:

- to verify that a security protocol is functioning correctly, it may be essential to model the fact that some data held by a participant is not externally visible;
- when synthesising a controller for a robot, the controller may not be implementable in practice if it bases its decisions on information that cannot be physically observed.

The aim of this project will be to develop case studies for this tool, for example to analyse security attacks or controllers for robots.

In addition, a component of the project will be to provide a translator of PRISM's low level textual description of the synthesised controller or strategy to a graphical representation.

For further details on this work see the paper:

<http://www.prismmodelchecker.org/bibitem.php?key=NPZ15>

Special Requirements :

Having a strong mathematical background is required.

Investigating retina-based vision systems (ID: 7183)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+.

Description:

This project aims to investigate retina-based vision systems in terms of:

1. how the retina samples the input image, for example, how a cat's retina samples the image surface is very different to how a human's retina samples or a frog's retina.
2. how to processes the equivalent retinal cell layers in software.

We currently have a GPU implementation of a software retina running on a GTX1080TI coupled with a Deep Learning Convolutional Neural Network (DCNN) for the purposes of evaluation the performance of the system for robot vision. In addition we have a simple "gaze control" module that decides where to direct the retina in the scene.

The first objective of this project is to explore different sampling patterns in the retina - we have a number implemented in Matlab and would like to port these to run on the GPU (this might not require GPU programming).

A second objective is to extend the model of the processing layers in the retina. Currently we model intensity and colour processing and a very basic temporal process. We would like to improve all of these models, but particularly the time-dependent processing based on the known operation of the eye.

The final objective, time permitting, is to demonstrate the system performing visual learning learning using DCNNs.

We have already developed:

- * GPU codes contained in Python wrappers to allow the software retina to compress input images to a size that the DCNN can handle directly,
- * Codes to find features in the cortical image output from the retina that can be used to direct where to look at next when either collecting training images, or attempting to recognise unknown objects.
- * Basic DCNN model architectures that interface with the retina and have demonstrated learning and recognition.
- * Colour retina cells that sample the red-green & blue-yellow components of objects and their contours or surface texture, to provide colour discrimination.
- * A basic time-dependent retina processing model.

All of the above are in Python with optional GPU acceleration.

Special Requirements :

No ethics required. CVMA4 desirable.

Investigating the balance between privacy and utility in mobile applications (ID: 7509)

Owner/Supervisor: [Inah Omoronyia](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Consumers and enterprises are increasingly relying on mobile systems, such as smart phones, to satisfy their social and business needs. This new generation of systems enables users to form localised, short- and long-lived communities to achieve common objectives. To this end, these systems may need to manipulate user's sensitive information, such as location, time, proximity to nearby services, and connectivity to other users. The disclosure of these attributes in an unregulated way can threaten user privacy. As such, the aim of a software engineer is to design systems that enable users achieve a balance between utility (the benefit of using the system) and privacy.

Special Requirements :

None.

Java Objects Security (ID: 7510)

Owner/Supervisor: [Inah Omoronyia](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The security of remote object invocations encompasses a wide range of potential considerations, ranging from the ability to utilize the security capabilities of middleware solutions (e.g., RMI, .NET, J2EE, DCE, REST etc.), to leveraging on programming language capabilities. These different approaches can also be used to realize complex and intricate security solutions such as discretionary access control, role-based access control, mandatory access control, etc. But the challenge with these potential considerations is that they are not amenable to highly dynamic and context sensitive environment. This is mainly due to the static nature of underlying security policy.

The objective of this project is to first explore and reflect on the security capabilities of java programming language, and that provided by dominant middleware platforms. You will then implement a simple lightweight toolkit for securing remote invocation of java objects.

Special Requirements :

None.

Keyword extraction from noisy speech-to-text (ID: 7070)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Speech recognition is quite reliable for high-quality recordings of well-articulated speech, especially when the domain/topic is known. Unfortunately, real-world multi-party conversations are quite the opposite. Here the task is to robustly extract reliable keywords and key phrases that best summarize the conversations from a noisy speech recognition output.

This project is potentially in collaboration with an industrial partner and will be tailored to the successful candidate.

Special Requirements :

None.

Least squares adjustment for the analysis of survey traverse data (ID: 7141)

Owner/Supervisor: [Alice Miller](#)

* Suitable for both IT/SD and CS+.

Description:

This project is to create software for Dr Elizabeth Petrie, a lecturer in the School of Geographical and Earth Science (GES) and will suit a student who has some experience with linear algebra (e.g. solving sets of simultaneous equations using matrices).

The software is to help students who are learning about using least squares calculations to analyse measurements taken using surveying equipment. Currently the students use existing software (n4ce - a survey processing package) to do these calculations. This software is very "black box", and does not show how functions work, or any intermediate steps. The purpose of this project is to develop a program to do the calculations that is constructed in such a way as to teach students about programming as well as the least squares adjustment method.

The program will be written in Python in such a way as to demonstrate to the students (a) how the least squares adjustment calculations work and (b) how good Python code is laid out and commented. This is a follow-on project. This version of the project will:

- be suitable for data in 3 dimensions
- have a fully functional GUI

- contain a set of worked examples for use by the students

[1] Charles D. Ghilani. Adjustment computations : spatial data http://encore.lib.gla.ac.uk/iii/encore/record/C__Rb3246281.jsessionid=58230257B0FDA9B45DA4F7ACA93F8970?lang=eng

Special Requirements :

The student will be required to visit GES to meet with colleagues in GES and use the n4ce package. Must have some maths background (especially linear algebra).

Local Search for Maximum Clique (ID: 7215)

Owner/Supervisor: [Patrick Prosser](#)

* Suitable for both IT/SD and CS+.

Description:

Given a graph $G = (V, E)$ a clique is a set of vertices in V that are all pair-wise adjacent. A maximum clique (MC) is then a largest clique in G . There are exact algorithms for MC. In this project we will develop & implement local search algorithms for MC and compare them to the exact algorithms

Special Requirements :

None.

Machine Learning Approaches to Sound Effect Synthesis (ID: 7065)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The standard in Foley and sound effects generation has been to record, sample and stored sound effects in simple, static libraries as uncompressed wave files with a minimum of metadata. Due to the versatile and complex nature of sound, a sound class (e.g. “fire” or “dragon”) may have up to 100 samples in the library associated with it. Each of these samples will differ very slightly on different dimensions to give the sound designer a wide variety of sounds to choose from.

Despite the seemingly large selection of samples available to sound designers today (e.g. [1]) it is however often the case that a designer does not find the “perfect” sound for an application. This is often due or to lack of variability within a particular class (on a particular aspect) or the designer being unable to navigate and express/define his needs via the system. Regardless this often leads to the designer either settling for a suboptimal sound sample which happen to be in the available library - or has to record his sound effect. This

leads low quality solutions; production time being wasted and sample libraries quickly becoming very large and unorganised.

This project will investigate deep learning for improving the current practices in sound effect generation and management. Specifically, we will use existing sound effect libraries (e.g. [1]) to train a state-of the neural network e.g. based on Keras/Tensorflow [2]. Once a certain type of synthesis model has been adapted to the characteristics of the phenomenon on which it has been trained, it can generate new sounds from the same space under similar conditions. The aim is to generate sound effects which are perceptually indistinguishable from recorded sound. A user study will be conducted to establish to which degree this has been accomplished.

[1]

https://offers.adobe.com/en/na/audition/offers/audition_dlc/AdobeAuditionDLCSFX.html

[2] <https://keras.io/>

Special Requirements :

None.

Machine learning for Analysing Biological Cells (ID: 7063)

Owner/Supervisor: [Bjorn Jensen](#)

* Suitable for both IT/SD and CS+.

Description:

This project will investigate and build a software pipeline for visualizing and analyzing images of biological cells using newly developed deep neural network-models which can segment, cluster and classify individual cells [3].

Practitioners without much experience with computer science or data analysis often apply computational methods in a suboptimal way [1,2]. This project aims at developing a web-based visualisation and machine learning pipeline which will provide non-CS experts with state of the art solutions for cell analysis based on recent results in machine learning and artificial intelligence.

The project is in collaboration with School of Engineering who will provide requirements and data.

Prerequisites:

Interest in artificial intelligence, machine learning, and image analysis.

Experience with Python and web-technologies (e.g. html, JavaScript) is an advantage.

Literature:

[1] cellprofiler.org

[2] Data-analysis strategies for image-based cell profiling, Caicedo et al, 2017, <https://www.nature.com/nmeth/journal/v14/n9/pdf/nmeth.4397.pdf?foxtrotcallback=true>

[3] Mask R-CNN, He et al, 2017, <https://arxiv.org/abs/1703.06870>

[4] Deeplearning, <http://www.deeplearningbook.org/>

Special Requirements :

None.

Matching social media accounts across different platforms (ID: 7259)

Owner/Supervisor: [Iulia Popescu](#)

** Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

The aim of this project is to help social media users understand what types of information can be shared online without making it possible for third parties (i.e. advertisers) to match accounts of the same person on different platforms (i.e. Facebook, Twitter). In order to achieve this, feedback mechanisms that inform and help users to avoid online identity recognition need to be developed. This project also offers the opportunity to design and implement novel techniques to find and match user accounts across different social media platforms.

Project plan:

- Familiarisation with Facebook and Twitter APIs.
- Given an user account, investigate what public information can be retrieved from the social media platforms.
- Develop a tool/script that can retrieve the user's public profile information from social media.
- Develop an algorithm that given a user account on social media platform A (e.g. Twitter), find the corresponding/matching account on social media platform B (e.g. Facebook), if the account exists.
- Build an appropriate feedback mechanism that will help users to prevent being identified across different social media platforms (e.g. this can be an browser addon/extension that advises the users on what information they can share and how this might be used by third parties, or a webapp that given a social media username as input, it can tell the user how quickly he/she can be identified on other platforms).

This project will be co-supervised by Dr Nikos Ntarmos.

Special Requirements :

None.

Max Clique as a CSP (ID: 7214)

Owner/Supervisor: [Patrick Prosser](#)

* Suitable for both IT/SD and CS+.

Description:

Given a graph $G=(V,E)$ a clique is a subset of V , such that all vertices in that set are pair-wise adjacent. One of the algorithmic steps is to greedily colour vertices, placing vertices in colour classes, such that each colour class is an independent set. In the Constraint Satisfaction Problem (CSP) we have a set of variables, each with a domain of values, and a set of constraints. We must then assign values to variables, from their domains, such that all constraints are satisfied. In the clique problem, colour classes can be viewed as variables and vertices within colour classes as values. The constraints are then adjacency. In this project we will develop a new Max Clique algorithm inspired by a CSP perspective

Special Requirements :

None.

Maximum Weight Clique (MWC) (ID: 7216)

Owner/Supervisor: [Patrick Prosser](#)

* Suitable for both IT/SD and CS+.

Description:

Given a graph $G=(V,E)$, where vertices have weight, a max weight clique (MWC) is a subset of the vertices in V that are all pair-wise adjacent and have a combined weight that is a maximum. In this project we will invent, implement and empirically study new algorithms for MWC

Special Requirements :

None.

Measuring Attention for a Kinect Enabled Public Display (ID: 7376)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Public displays are increasingly interactive, giving users the opportunity to play, find information, and engage with each other in public spaces. Making displays interactive using Kinect means that passers-by can use their whole body to interact with a display. Previous work has shown that displaying the users' silhouette to indicate interactivity is very successful in capturing users' attention.

This project will involve creating a playful application using Kinect with a public display to evaluate the ideal parameters for capturing curiosity using a Kinect, specifically looking at how enabling interactivity at different depth ranges influences interaction at the display.

This HCI focused project will involve an evaluation of the display on the University of Glasgow campus.

Special Requirements :

None.

Merging molecular networks (ID: 6986)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Molecular networking and MS2LDA are both tools that put cluster molecules into groups based on their fragment spectra. Recent investigation has shown that they provide somewhat complementary views of the molecules being analysed and it is interesting to overlay the results of one analysis on the other. So far, we have investigated doing this via colouring edges in the molecular network according to shared membership of a 'cluster' in MS2LDA. In this project you will develop a more flexible visualisation platform that will enable the users to draw networks which can be manipulated in real time to help explore the similarities and differences between the groupings provided by these algorithms. For example, users will be able to switch on and off different clusters from either analysis, colour edges according to shared membership, add additional edges for new clusters, etc. Development could be in Python, Java, or using a web technology (e.g. D3).

Special Requirements :

None.

Metrics for understanding textual requirements quality (ID: 7511)

Owner/Supervisor: [Inah Omoronyia](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Substantial amount of the documents available for requirements analysis are written in text. One of the main reasons for this is the realisation that natural language is universal, flexible and widespread. These advantages normally come with a sacrifice on the quality and hence a common understanding of the stated requirements. A number of quality models for natural language

requirements are present in research literature. Some are based on language structure (grammar) analysis while others are based on the meaning (semantics and pragmatics) of concepts used in the expression of requirements

The aim of this project is to investigate metrics for understanding the similarities between requirement statements and provide a measure of ambiguity and inconsistency that can possibly exist in such statements. The student will be required to investigate existing quality models and how they can be applied to understanding the problem of ambiguity in textual requirement statements.

Special Requirements :

None.

Mining academic expertise from funded research (ID: 7624)

Owner/Supervisor: [Craig Macdonald](#)

* Suitable for both IT/SD and CS+.

Description:

We have been developing academic search engines that assist in identifying the relevant experts within Scottish Universities, based on their recent publication output. However, we have recently obtained access to data from research grants that can supplement publication data as evidence of expertise. In this project, the student will develop mining tools for the data, and research ways to integrate it with existing deployed academic search engines to obtain the most effective search results.

Technology: Java, Terrier, machine learning

Special Requirements :

None.

Mobile events system for University (ID: 7178)

Owner/Supervisor: [Roderick Murray-Smith](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Universities often organise daily talks on a range of topics. However, members of industry, students and other academics, or even members of the public may not be aware of talks or their relevance to their own work. We have built a system currently in use within the University

<https://samoa.dcs.gla.ac.uk/events/>

This project will involve creating a novel visualisations of talk series and topics to give insight into what is going on in a school.

Special Requirements :

Students should take IR, and would benefit from Text as Data

Mobile Memories app for refugees and asylum seekers (ID: 7581)

Owner/Supervisor: [Quintin I Cutts](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The project will involve working in collaboration with staff in the School of Education who are working with refugees and asylum seekers, particularly the UNESCO Refugee Integration through Languages and the Arts (UNESCO RILA) project. They are considering collecting and recording refugees' memories (e.g. songs, poems, proverbs, etc.) that are triggered by features around Glasgow. These memories would then be made available via an app, so that anyone who is interested to learn what resonance an object/space/museum exhibit/etc. has for someone coming from a different country, may access the recordings on their mobile device. The precise details will be worked out once the project is assigned, but we would hope to contact Glasgow Life to see if they would be willing to let us pilot the app in one of Glasgow's museums or art galleries. Here is one initial thought about how the system would work: having downloaded the app, someone could walk into a museum, see the exhibit of a woven blanket, see a logo and a QR code next to it, point their phone at the QR code and up would come a song/saying/story linked to the blanket. Staff at UNESCO-RILA would collect the stories, songs, etc. so the project's main focus is developing the web app to store and deliver them. The project is probably better suited to those with stronger programming skills.

Special Requirements :

None.

Mobile Phone Application 1 (ID: 7320)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

This project is to develop an interesting application for a mobile phone. The choice of application is up to you. Games and puzzles are a good source of ideas, but it could be anything. It could be for any type of phone: iPhone, Android, Windows. You will need good programming skills. You will need to have your own phone of a suitable type.

Special Requirements :

None.

Mobile Phone Application 2 (ID: 7328)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

This project is to develop an interesting application for a mobile phone. The choice of application is up to you. Games and puzzles are a good source of ideas, but it could be anything. It could be for any type of phone: iPhone, Android, Windows. You will need good programming skills. You will need to have your own phone of a suitable type.

Special Requirements :

None.

Modelling and analysis of search and pursuit-evasion games (ID: 6976)

Owner/Supervisor: [Gethin Norman](#)

* Suitable for both IT/SD and CS+.

Description:

Search and pursuit-evasion games have been developed for mobile robots. In the project such games would be formally modelled and the performance of different strategies would then be analysed (also finding what is the optimal behaviour). Here is a link to a survey paper on such games for further information:

<http://www-users.cs.umn.edu/~isler/pub/pe-survey.pdf>

There is scope to go in a number of different directions with the project. The basis would be using the game extension of the probabilistic model checker PRISM:

<http://www.prismmodelchecker.org/games/>

Special Requirements :

None.

Modelling and analysis of wireless sensor networks with bigraphs (ID: 7201)

Owner/Supervisor: [Michele Sevegnani](#)

* Suitable for both IT/SD and CS+.

Description:

A Wireless Sensor Network (WSN) consists of a homogeneous set of wireless sensor nodes, autonomously managing their physical connections. Nodes have a fixed, uniform transmission range and a bidirectional link exists between two nodes if their distance is lower than or equal to the transmission range.

Deploying applications directly on this physical layer is often overly complicated and not very flexible. Therefore, it has been increasingly popular to adopt managing and programming paradigms based on overlay networks in which nodes can be thought of as being connected by virtual links, each of which corresponds to a path in the underlying network.

The aim of this project is to develop models of WSNs including both the underlay physical network and the overlay network. Analysis will focus in particular on the interplay between events in the underlay network and invariants on the applications deployed on the overlay network. For example, a node failure can affect the status of a virtual link and therefore break some application requirements.

Models will be implemented using bigraphs, a graphical formalism introduced by British computer scientist Robin Milner (Turing Award in 1991) to model systems evolving in time and space. Analysis and simulations will be carried out with BigraphER [1].

[1] BigraphER: <http://www.dcs.gla.ac.uk/~michele/bigrapher.html>

Special Requirements :

None.

Modelling and verification of DNA programs (ID: 6977)

Owner/Supervisor: [Gethin Norman](#)

* Suitable for both IT/SD and CS+.

Description:

DNA molecules can be used to perform complex logical computation. DNA computation differs from conventional digital computation and is sometimes referred to as 'computing with soup'

(see <http://www.economist.com/node/21548488>).

The aim of this project is to model and analyse using the probabilistic model checker PRISM the DNA implementation of logic inference proposed in

Alfonso Rodríguez-Patón, Iñaki Sainz de Murieta, and Petr Sosík

Autonomous Resolution Based on DNA Strand Displacement

DNA Computing and Molecular Programming, volume 6937 of LNCS, 2011.

Special Requirements :

A strong mathematical background is required.

Modelling Continental Drift (ID: 7399)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

It is fascinating to watch a visualisation of continents drifting over the surface of the earth during millions of years. It is also difficult to see what is happening because of the shape of the globe.

This project will use a 2D flat square surface to model how continents drift. The left and right, and top and bottom edges will be joined so that there are no boundaries to impede the slow drift of the continents. It should then be easier to see these artificial continents as they drift over the flat earth.

The project will model the physical processes involved when continents collide. Energy is absorbed to slow down the relative movement, while momentum is conserved. This energy will create mountains which will then be eroded to form sedimentary rocks in the oceans. There will also be a process where continents randomly split and move apart to put energy back into the system.

Once the basic system has been built, interesting geological and environmental questions can be investigated. For example, mountain building puts carbon dioxide into the atmosphere while mountain erosion takes it out again.

Additionally, large continents over the poles lead to the build up of ice layers which lower the sea level, provided the temperature is not too warm. The flat earth can have a hot spot and a cold spot to represent the poles and the equatorial region.

Special Requirements :

None.

Multi-Level Parallel Applications with the C++ Parallel STL (ID: 7438)

Owner/Supervisor: [Michel Steuwer](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

High-performance applications are essential for many sciences. NASA has developed a benchmark suite for a representative set of their applications called the NAS Parallel Benchmarks. These include a set of benchmarks designed to exploit multiple levels of parallelism.

C++ has recently adopted a new version of the standard template library (STL)

which is defined in terms of composable "ranges" which replace the idea of iterators.

In this project, you are going to implement a subset of the NAS Parallel Benchmarks using modern C++ and the range-based STL. We are interested in understanding the benefits and potential overheads of implementing high-performance applications when using the abstractions provided by the STL.

Special Requirements :

Familiarity with C++ is a beneficial prerequisite. You don't need advanced experience with the STL for this project, this will be acquired in the project.

Multi-threaded graph caching (ID: 7218)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

The project will investigate, design, and implement multi-threaded support for the GraphCache, a system developed by the IDEAS group to expedite graph queries.

The student will need to implement a task scheduling system, using Java threads/synchronisation primitives and (optionally) thread pools. Thus excellent knowledge of the Java programming language is a must. The student should also have a good knowledge of operating systems principles, especially regarding synchronisation of parallel processes/threads, race condition and deadlock avoidance, remote procedure calls, and client-server design patterns.

Special Requirements :

Excellent knowledge of the Java programming language; good knowledge of operating systems principles.

Multi-User BibTeX Manager (ID: 6941)

Owner/Supervisor: [David F Manlove](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

BibTeX is a way of encoding references for use by the LaTeX mathematical typesetting tool. In this context, a reference might be a journal article, book chapter, or contribution to a conference proceedings volume, for example.

BibTeX is based on .bib files, which are plain text files encoding lists of references. Each reference contains standard fields, such as the author, title, volume number, pages etc. Here is an example BibTeX reference:

```
@article{Man02  
author = "D.F. Manlove",
```

title = "The structure of stable marriage with indifference",
journal = "Discrete Applied Mathematics",
volume = "122",
number = "1-3",
pages = "167-181",
year = "2002"}

The purpose of this project is to provide a web application that can allow multiple users to work on sets of shared BibTeX records that are stored on a central database. It involves extending an existing project that already provides features for adding, editing and deleting BibTeX records. Various additional features could be added depending on the student's interests and time constraints.

This project would probably be best suited to a student who has enjoyed the ITECH course. For more information, see <http://en.wikipedia.org/wiki/BibTeX>.

Special Requirements :

None.

Network Asset Discovery Tool Development for Industrial Control Systems (ID: 7565)

Owner/Supervisor: [Chris Johnson](#)

* Suitable for both IT/SD and CS+.

Description:

This project involves research and development of a bespoke network discovery tool for use in critical infrastructure environments, in order to build an organisation's asset register.

Special Requirements :

none

Next Generation Issue Tracking (ID: 7390)

Owner/Supervisor: [Tim Storer](#)

* Suitable for both IT/SD and CS+.

Description:

Current state of the art software project issue tracking tools, such as GitHub, GitLab, JIRA and Trac store issues separately from the version control repository that contains the project's source code. This creates *friction*, because software developers must remember to keep both the issue tracker and the version control repository up to date as progress is made on completing tasks. The aim of this project is to develop

a web-based front end for managing issues that are stored within the version repository. This has a number of advantages for eliminating friction:

- * The package that an issue refers to can be identified implicitly, by storing the issue in the most relevant sub-directory of the code base.
- * The assignee for an issue can be identified based on who is making commits to a repository.
- * Implicit links can be identified between source code affected by an issue, if both the issue and the source code are altered in a single change set.
- * Comments on issues can be automatically recovered from the version control log.

The aim of this project is to develop a demonstrator application or plugin for existing software project management tools that allow issues to be managed **within** a version control repository rather than as a separate database.

Special Requirements :

None.

Pan-tilt camera, gaze control & retina-deep learning in a real-time vision pipeline (ID: 7182)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+.

Description:

This project aims to combine a GPU implementation of a software retina running on a GTX1080TI with a Deep Learning Convolutional Neural Network (DCNN) and gaze control of a pan-tilt camera.

The objective of this project is to code together the following components into a complete system that can direct the pan-tilt to capture training images, train a DCNN classifier and then direct his camera to recognise unknown objects from the training set.

We can then validate the performance of this recognition system which should be fast enough to be able to execute in real-time (> 15 frames/second) to allow continuous perception of a robot's workspace.

If time permits, we would like this camera to be controlled from within ROS, the Robot Operating System, to facilitate integration with our robot systems.

We have already developed:

- * GPU codes contained in Python wrappers to allow the software retina to compress input images to a size that the DCNN can handle directly.
- * Codes to find features in the cortical image output from the retina that can be used to direct where to look at next when either collecting training images, or attempting to recognise unknown objects.
- * Basic DCNN model architectures that interface with the retina and have demonstrated learning and recognition.
- * Colour retina cells that sample the red-green & blue-yellow components of objects and their contours or surface texture, to provide colour discrimination.

Special Requirements :

No ethics required. CVMA4 desirable.

Parsing protocol specifications (MSc IT+) (ID: 7026)

Owner/Supervisor: [Colin Perkins](#)

* **Not suitable for CS+ students.**

Description:

The Internet Engineering Task Force (IETF) has a long history of using a combination of an ASCII art diagram and some explanatory text to define packet formats. An early example of this is the IP specification

<https://tools.ietf.org/html/rfc791#section-3.1>

The precise format of the diagram and the following text has never been defined. This makes it difficult to write code to parse and check the specifications for consistency. As a result, errors can occur (e.g., <https://tools.ietf.org/html/rfc6958> section 3, where the description disagrees with the diagram).

The aims of this project are: 1) to write a tool that can find and extract ASCII art packet diagrams and their associated explanations in the standards document; and 2) attempt to check that the diagrams and text match.

Finding and extracting the diagrams is expected to be a conceptually straightforward text parsing problem.

Checking for consistency is unlikely to be possible in the general case, since it requires natural language processing. The idea is to find the simple cases that are possible, and write a tool to check those.

Code for this project should be written in Python.

Special Requirements :

None.

Pascal to Java/JVM compiler (ID: 7313)

Owner/Supervisor: [Simon J Gay](#)

* Suitable for both IT/SD and CS+.

Description:

The aim is to design and implement a Pascal compiler, either by translating to Java, or by translating directly to Java Virtual Machine code. If time permits, language extensions can be explored.

This is an opportunity to learn about programming language implementation, as well as gaining an in-depth knowledge of the Java Virtual Machine.

Special Requirements :

None.

Phylogenetic analysis of cancer evolution (ID: 7104)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Background:

Next generation sequencing data has provided a unique opportunity to study intra-tumor

heterogeneity in cancers. A recent study by Nik-Zainal et al. (2013) has shown an attempt

on pinpointing the phylogeny of clones in patient samples using deep sequencing data.

Traditionally, phylogenetic methods (Felsenstein, 2004) are designed to accommodate

allele evolution from sequences of different species. In cancer, many assumptions of

traditional phylogenetic methods no longer hold. For example, samples are not well-defined species and ancestral clones (inner nodes) can coexist with descendant clones

(leaf nodes) in the same tumour. In order to get a clear picture of how these conceptual

changes may affect the result from phylogenetic analysis, we want to

understand the

performance of classical phylogeny reconstruction approaches on NGS data from cancer

samples.

Details of the project:

In this project, we will comprehensively examine classical phylogeny reconstruction

approaches on NGS data from patient samples. We will address the following questions:

1. Unbiased comparison of the performance of the classical phylogeny reconstruction

methods.

2. Examine what assumptions are/aren't suitable in the context of reconstructing

cancer evolution history.

As a benchmark, we will use both simulated data and patient data including i) DNA

methylation data from colon cancer patients in Sottoria (2013). ii) Single cell sequencing

data from a kidney cancer patient in Xu (2012).

Methods: We will use the three mainstream classes of phylogeny reconstruction methods:

1. Parsimony methods.

2. Distance matrix methods.

3. Probabilistic (likelihood and Bayesian) methods.

References:

1. Nik-Zainal et al. (2012). The life history of 21 breast cancers. Cell 149, 994-1007.

2. Sottoria et al. (2013). Single-molecule genomic data delineate patient-specific

tumor profiles and cancer stem cell organization. Cancer Research 73, 41-49.

3. Xu et al. (2012). Single-cell exome sequencing reveals single-nucleotide mutation

characteristics of a kidney tumor. Cell 148, 886-895.

4. Felsenstein (2004). Inferring Phylogenies. Sinauer Associates,

Sunderland, Massachusetts.

Special Requirements :

None.

Probabilistic Model Checking Versus Planning (ID: 6979)

Owner/Supervisor: [Gethin Norman](#)

* Suitable for both IT/SD and CS+.

Description:

PRISM is an open source formal verification tool for analysis of probabilistic systems. PRISM has an extensive website

(<http://www.prismmodelchecker.org/>) which includes software for download, tutorial, manual, publications and many case studies.

One of the things PRISM can do is solve optimisation problems on Markov decision problems (MDPs). These are also widely used for planning problems, e.g. in the fields of robotics and AI. This project will implement connections between planning tools and languages (e.g. RDDL) and investigate how well PRISM performs in this field.

Special Requirements :

None.

Python Dependency Graph Generator (ID: 7418)

Owner/Supervisor: [Tim Storer](#)

* Suitable for both IT/SD and CS+.

Description:

This project will be supervised by Tom Wallis and Ben Bumpus.

Consider the scenario: you've developed a small project which quickly gained

traction and turned out to be rather successful, and what was once a hastily hacked-together codebase requires major refactoring and careful architectural engineering! However, the code as it exists is brittle and breaks in lots of unexpected places when even small edits are made. If only a tool existed which could help to refactor the codebase and work out a better architecture...

One way to do this could be to build a graph of dependencies between functions. If we know which parts of our code rely on other parts to operate properly, we could construct a graph representing how different parts of our code interact! Then, we might analyse various properties of the graph to learn about our codebase, and how it could be refactored. A question is left to be answered: how do we build this dependency graph, and what can we learn about our code by measuring the graph's properties?

This project attempts to help software engineers understand how different parts of their codebase depend on others, and potentially predict / visualise technical debt and coupling, by producing dependency graphs of the functions in a codebase. The project will then attempt measurements of the graph which might identify properties of the codebase it represents, such as coupling, potential modularisation, and the potential impact of editing code through observing density and tracing paths.

The objectives for this project are:

- Implement a dependency graph generator for Python code
- Produce a library of graph measurements and assess whether they can be useful to a developer in understanding their codebase
- Evaluate the effectiveness of these graph measurements

Special Requirements :

None.

Real-time machine learning for sentiment analysis in multi-party conversations (ID: 7071)

Owner/Supervisor: [Bjorn Jensen](#)

** Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

This project will investigate machine learning techniques for real-time sentiment analysis of audio streams, i.e predict and visualize the emotional state of the individual speakers.

This project is potentially in collaboration with an industrial partner and will be tailored to the successful candidate.

Special Requirements :

None.

Real-time privacy monitoring for smart phones. (ID: 7512)

Owner/Supervisor: [Inah Omoronya](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Smartphone applications are increasingly gathering and disclosing personal information about users. While disclosing this information is essential for their functionality, it may also raise privacy concerns. A key shortcoming is that these applications do not provide users with adequate control over and visibility into the usage and subsequent transfer of such personal information. Where such control and visibility is missing, it becomes difficult for users to make appropriate disclosure decision that preserves their privacy. The aim of this project is to address this setback by building an efficient, system-wide dynamic tracking and analysis system for smartphones.

Special Requirements :

None.

Recommending study buddies (ID: 7222)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

The goal of this project is to produce a system that will recommend "study buddies" to students so as to help them improve in areas where they have weaknesses. In this system, students will be represented via their "knowledge graph". The student will first be briefed on how these graph representations are produced. Then the student will (a) review relevant literature on recommender systems and graph clustering, (b) create a simple data generator to produce random but meaningful "knowledge graphs", (c) implement at least one recommender system putting the above together, and (d) evaluate the system's performance from both a qualitative and quantitative point of view.

The student will receive guidance from Dr Nikos Ntarmos and Dr Christos Anagnostopoulos.

Special Requirements :

Excellent programming skills. Good understanding of data structures and algorithms. Good understanding of machine/statistical learning principles and methods.

Regression Testing for the Lift Open-Source Project (ID: 7429)

Owner/Supervisor: [Michel Steuwer](#)

* Suitable for both IT/SD and CS+.

Description:

Performance regression testing is a form of testing that ensures that the performance of software stays the same (or improves) during a software development process.

The Lift project (<http://www.lift-project.org/>) is an open-source research project which currently uses continues testing to reduce the number of bugs in the program but which does not use performance regression testing.

In this project, you will setup a regression testing service used by the Lift project. As Lift has challenging hardware requirements this involves the configuration and setup of multiple machines, the writing of scripts, and the integration of the regression testing in the software development workflow.

Special Requirements :

None.

Robot behaviour prototyping using block programming (ID: 7472)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

* Suitable for both IT/SD and CS+.

Description:

Programming a robot is a time-consuming task and programmers spend most of their time debugging and making things work. The aim is to implement common and isolated robot functionalities in Blockly [1, 2] to mitigate the above problem. You will implement code recipes for common robotic tasks, such that complex behaviours can be prototyped. Specific objectives are:

- * Define and interface code recipes of ROS packages (www.ros.org) for a set of simple tasks for Baxter
- * Interface "robot_blocky"
- * Design a set of complex behaviours for pick-and-place tasks

Technologies: ROS, Python, OpenCV

[1] <https://developers.google.com/blockly/>

[2] https://github.com/erlerobot/robot_blockly

Special Requirements :

None.

Robotic Pick-n'-Place with a Smart Grasping System (ID: 7475)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

* Suitable for both IT/SD and CS+.

Description:

This project is about integrating a robotic system using a Smart Grasping System (SGS - <https://www.shadowrobot.com/smart-grasping-system/>) for pick and place tasks. The CVAS group at the School currently collaborates with Shadow to develop an integrated grasping and vision system. The aim is to implement computer vision modules to detect objects using an RGB-D sensor (e.g. Microsoft Kinect) and develop robot manipulation and grasping capabilities to solve this task.

For this project, you will use a robotic simulator featuring the SGS to test and evaluate your developed methods. You will also use Python and ROS.

Special Requirements :

None.

Robot-vs-human: A chess game (ID: 7464)

Owner/Supervisor: [Gerardo Aragon Camarasa](#)

* Suitable for both IT/SD and CS+.

Description:

Baxter is a dual-arm robot manufactured by Rethink Robotics. It features cameras, joint sensors and an easy-to-learn software API based on the Robot Operating System (ROS). Baxter is safe and can interact with people as its movements are elastic and not dangerous to humans.

The project will consist of investigate state-of-the-art deep learning approaches to enable Baxter to understand the state of the chessboard and identify if it is playing against a human. The project has the following objectives:

- * To develop computer vision modules that allow Baxter to understand the state of the chessboard — previous projects have built a considerable training set to be use with state-of-the-art deep learning frameworks. In this project, the aim is to test at least one deep learning architecture against the currently developed versions.

- * To interface an emotion Python package [1] to allow Baxter to react to different human emotion states while playing chess.

[1] <https://github.com/Selameab/emopy>

Special Requirements :

None.

ROS control of Mitsubishi RV-M2 robot for 3D sensing (ID: 7172)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+.

Description:

The RV-M2 <http://www.rixan.com/Portals/0/RV-M2.pdf> is a solid and functional industrial robot that has a now very dated PC software interface, housed in the Computer Vision for Autonomous Systems digitisation lab in Lilybank Gardens. We have investigated porting the ROS operating system to this robot and this appears to be feasible.

The objective of this project is to write a new interface to allow the RV-M2 to be controlled by the inverse kinematics package MoveIt from ROS (Robot Operating System) www.ROS.org

Details of the RV-M2 and

programming: <http://www.mitsubishirobots.com/Mitsubishi%20Movemaster%20Robot%20Manual/rv-m2.pdf>

The RV-M2 is currently connected to a modern PC running Ubuntu Linux via an RS232 interface and can be controlled over this link.

A ROS node will be implemented in this project allowing ROS services to communicate with the MV-R2

A 3D model of the MV-R2 is to be constructed as part of the project to allow it to be used in the Rviz and Gazebo visualisation packages and the principal deliverable for the project is to demonstrate the RV-M2 being operated under ROS.

If time permits, the demonstration should also include the robot arm scanning a 3D camera in a preset path to digitise an object and produce a 3D model using Kinect Fusion.

Special Requirements :

Access to the Mitsubishi RV-M2 robot in G142

Secure web server development (ID: 7338)

Owner/Supervisor: [Joemon Jose](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Given that malicious attacks or hacking on web servers are shown an increasing rate, the need for Secure web technologies are becoming a necessity. It is not clear what and how such secure servers could be developed.

The goal of this project is to develop a secure web server using technologies such as cryptographic methods for connections, passwords, or encrypted database components. It is also expected to use Open authentication framework for access to the server. It is also important to develop and study various threat models.

Specific objectives:

Communication between server and the client should be encrypted

all database entries must be encrypted

all cookies must be secured

the system should be resistant to session impersonation

study the efficiency and scalability issues of various encryption schemes

implementation of the open authentication framework

development and study of various threat models

development and experiment of cloud based server solutions

Special Requirements :

None.

Security Impact of Wireless Networking & IoT Devices in Industrial Control Systems (ID: 7566)

Owner/Supervisor: [Chris Johnson](#)

* Suitable for both IT/SD and CS+.

Description:

This project involves analysing the potential security implications of using wireless networking technology in industrial environments (vs traditional wired networking), with a particular attention to IIoT (Industrial Internet of Things) vulnerabilities and potential attack vectors.

Special Requirements :

none

Simulating Actuated Devices in Virtual Reality (Unity) (ID: 7449)

Owner/Supervisor: [Graham Wilson](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

The aim of this project is to extend an existing Unity Virtual Reality platform that can detect the location of a user, their hand positions and their gaze direction relative to objects placed in the virtual environment. The platform currently presents a phone or a horizontal pin-array, which ‘actuate’ (move) in a range of ways when approached/looked at: changing shape, position and texture/colour. The purpose is to prototype “expressive” devices that can move instead of just showing things on a screen. The project will expand the types of objects that can be shown in the environment (e.g., a watch or a levitating display of objects), and the number of ways in which the objects can actuate (e.g., bending, 2D/3D deformation). It will also involve a user study where the actuations are presented to users and their subjective opinions about them are gained. The project would benefit from familiarity with C#.

Unity programs are developed in C#, and the project will give you experience in game development, 3D graphics and virtual reality. The project will be based on the HTC Vive or Oculus Rift VR headsets, depending on requirements

Special Requirements :

None.

Sketching app for Do-It-Yourself projects (ID: 7583)

Owner/Supervisor: [Quintin I Cutts](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Home Do-It-Yourself folk often sketch out plans for things they are going to build. These are usually on scraps of paper and such like. The purpose of this project is to create a seemingly simple, but incredibly useful, app to support the fast development of such sketches. This is a start-simple, sky’s-the-limit kind of project, where you can do as little or as much as you choose. There are folk in the School and elsewhere who will act as clients for the project, to help set the requirements and to act as evaluators later on.

Special Requirements :

None.

Software Quality Metrics for Unit Tests (ID: 7391)

Owner/Supervisor: [Tim Storer](#)

* Suitable for both IT/SD and CS+.

Description:

The development of both quality metrics and design patterns for application software is well understood (think cyclometric complexity, the factory pattern, observer observable etc.). But

what are the characteristics of a well designed, maintainable unit test? These are likely to be very different from the characteristics of maintainable application code, because the constraints of writing an effective test contradict some of the practices of writing good application code. For example, tests should be concise, self-contained and readable, and focus on 'one thing', but this also encourages code cloning.

The aim of this project is to develop and validate metrics and/or design patterns for unit test code.

Special Requi

Spectral consensus scoring (ID: 6983)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Finding common substructures amongst a group of molecular fragment spectra is a useful task in many applications. The general problem is: given a set of spectra and, for each spectrum, 0 or more candidate molecular structures, are there substructures common to the group that might help to determine what the group represents (and help to identify individual members). In this project, you will develop and compare different methods for solving this problem when given a set of spectra and their candidate structures. Code will be developed in Python.

Special Requirements :

None.

Spectrum ID into MS2LDA (ID: 6984)

Owner/Supervisor: [Simon Rogers](#)

* Suitable for both IT/SD and CS+.

Description:

Our MS2LDA pipeline allows the decomposition of metabolomic mass spectrometry data into 'topics'. It would be useful to add functionality to this software to allow standard spectral ID tools to ID the spectra that exist in standard databases and allow in-silico fragmentation predictors to ID the others. This will involve creating the pipelines (most likely in python) that will link our platform to various external tools, allow users to submit query datasets and store the resulting annotations.

Special Requirements :

None.

SSPPFlow: TensorFlow implementation of state-space models with point process observations (ID: 7103)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

State-space models with point process observations (SSPP) are a class of probabilistic models that can be used to deal with a wide range biological applications from neural spike trains to RNA sequencing analysis in cancer.

In this project, we will develop a Python software package that implements SSPP using Google's TensorFlow engine. We will test the software with real world neural spike trains to RNA sequencing data.

Special Requirements :

None.

Startup Bootstrap (ID: 7386)

Owner/Supervisor: [Tim Storer](#)

* Not suitable for CS+ students. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

I had an MSc project team build a prototype website to facilitate very early stage startups formation. This will need refining and extending over the next year and could lead to something quite exciting as a forum for generating a pipeline of tech-startups. The aim of this project would be to continue to develop the features of the platform, introducing ideas such as recommendations and notifications.

Special Requirements :

None.

Steganography (ID: 7397)

Owner/Supervisor: [Ron Poet](#)

* Suitable for both IT/SD and CS+.

Description:

Steganography is a way of hiding or communicating secret information by pretending it does not exist. In the modern era this is normally achieved by embedding a digital payload in a larger digital cover document. The payload can also be encrypted for additional security. Images and videos make good cover documents and the aim of steganography is to embed the payload without changing the appearance of the cover file. Electronic countermeasures against steganography involve detecting cover documents that contain a payload and distorting the cover documents to destroy the payload without

changing their visual appearance.

This project involves implementing a common steganography algorithm: BPCS and investigating the effectiveness of various countermeasures. One goal is to investigate the maximum carrying capacity of various cover documents while ensuring that the secret message cannot be detected or distorted.

This project will involve a significant amount of programming, as well as a user study.

Only choose this project if you are taking MSc Information Security or MSc IT Cyber Security.

Special Requirements :

None.

Student simulator (ID: 7217)

Owner/Supervisor: [Nikos Ntarmos](#)

* Suitable for both IT/SD and CS+.

Description:

Currently MOOC and e-learning providers tailor their course offerings and recommendations based on the interests of their students. However, key deciding factors – such as the students' attainment, skills, weaknesses, and aspirations – are not taken into account. For such systems to be developed, a large amount of data needs to be gathered through monitoring of the interactions and progress of actual students. Alas, this information is not always readily available.

The goal of this project is to produce a simulator for the attainment of students in a virtual class. The student will need to review theoretical models of student attainment for such use cases, as well as the fundamentals of probabilistic data generators. Then the student will (a) examine anonymised records of student attainment (possibly including attendance records) and deduce a model, using machine/statistical learning techniques, (b) create a program that can be parameterised (based on multiple input dimensions, including student characteristics and course/assessment requirements, but also taking into account the aforementioned models), and which will generate student attainment data approximating what a real class would produce.

The student will work under the joint guidance of Dr Nikos Ntarmos and Dr Christos Anagnostopoulos.

Special Requirements :

Excellent programming skills, good knowledge of statistics and machine learning primitives

Supervised learning approaches to estimate cancer cell fractions (ID: 7112)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Cancer cell fraction (CCF) refers to the percentage of cancer cells bearing a given mutation. It has become a key measurement of intratumor heterogeneity in recent large-scale cancer genome projects, such as the PanCancer Analysis of Whole Genomes (PCAWG) project of ICGC. Currently, CCFs are mainly estimated by clustering analysis via mixture models. Mixture models suffer from high computational cost and problems with model selection.

In this project, we will develop supervised learning approaches to estimate cancer cell fractions. In particular, we will

- 1) use several state-of-art CCF estimator to form consensus CCF classes
- 2) develop supervised learning approach using consensus CCF as labels.

Special Requirements :

None.

Taint Visualisation Tool (ID: 7513)

Owner/Supervisor: [Inah Omoronya](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

A taint visualisation tool with the following properties:

- Associate taint marker with untrusted input as it enters the program
- Propagate markers when string values are copied or concatenated
- Report vulnerabilities when tainted strings are passed to sensitive sinks
- Will be a plus if the tool can also be integrated into Eclipse IDE

Special Requirements :

None.

Topical modelling for identifying copy number signatures from cancer genomes (ID: 7102)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

Features of the copy number alterations across cancer samples can be used as signatures for certain biological process that causing the alterations. It can provide key insights into the disease progression and novel treatment options.

In this project, we will develop a Bayesian hierarchical model, similar to the topical models in natural language processing, to better define those copy number signatures.

Special Requirements :

None.

Tracing evidence of intratumour heterogeneity from DNA methylation profiles (ID: 7094)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

DNA methylation is an epigenetic marker which acts as a switch of the genetic information it covers. It is also well-known that DNA methylation is largely inherited between cell division. In cancer, errors are frequently incurred as results or indicator as the disease progresses, leaving a trace of the life history of a tumour.

Emerging sequencing technologies has allowed genome-wide profiling of DNA methylation in tumour samples. In this project, we will investigate evidence of diversity among subpopulations of cancer cells within a single tumour. We will work with large collections of such data from publically available dataset such as The Cancer Genome Atlas and International Cancer Genome Consortium.

Product: 1. Detailed analysis of the results of the research. 2. Codes to reproduce the analysis.

Special Requirements :

None.

Tracing evidence of intratumour heterogeneity from DNA methylation profiles (ID: 7098)

Owner/Supervisor: [Ke Yuan](#)

* Suitable for both IT/SD and CS+.

Description:

DNA methylation is an epigenetic marker which acts as a switch of the genetic information it covers. It is also well-known that DNA methylation is largely inherited between cell division. In cancer, errors are frequently incurred as results or indicator as the disease progresses, leaving a trace of the life history of a tumour.

Emerging sequencing technologies has allowed genome-wide profiling of DNA methylation in tumour samples. In this project, we will investigate evidence of diversity among subpopulations of cancer cells within a single tumour. We will work with large collections of such data from publically available dataset such as The Cancer Genome Atlas and International Cancer Genome Consortium.

Product: 1. Detailed analysis of the results of the research. 2. Codes to reproduce the analysis.

Special Requirements :

None.

Tracing malicious behaviour on Twitter (ID: 7514)

Owner/Supervisor: [Inah Omoronyia](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

“Trolls” - individuals who engage in malicious behavior - are a common occurrence within online communities. Yet simply banning accounts associated with trolls is often ineffective as individuals may register new accounts under pseudonyms and resume their activity. This project will investigate how this can be addressed through a behavioral trace.

The project aim is as follows:

- 1) To analyse Twitter posts of an account to derive semantic signature unique to the account's owner.
- 2) To compare the signatures of two accounts, to determine whether they belong to the same user.
- 3) Validate the technique on a dataset of Twitter users.

Special Requirements :

None.

Training the Baxter Research Robot using a VR interface based on the HTC Vive (ID: 7195)

Owner/Supervisor: [J Paul Siebert](#)

* Suitable for both IT/SD and CS+.

Description:

We recently coupled the Oculus Rift VR headset, and now an HTC Vive headset, to a stereo-pair camera mounted on Baxter's head (this can pan-tilt),

such that when the headset wearer pan-tilts their head, Baxter's head and hence the stereo-pair camera moves accordingly. We have also coupled both brands of VR hand-controllers to MoveIT in ROS, the Robot Operating System, hence by moving the VR controllers, Baxter's end effectors also move broadly as the user would expect. Therefore a user wearing the VR headset would see through Baxter's eyes (i.e. cameras) and see Baxter's end effectors (instead of their hands) and control these using the Oculus Touch.

The objective of this project is threefold:

1. Firstly, to improve the operation of the HTC Vive control of Baxter by completing missing features or fixing communications issues that the original implementation has not yet resolved. In other words, improve the functionality and software quality of the system. Budget will be available to update hardware components, such as cameras and perhaps their control, as required by the project as it progresses.
2. Build a good demonstration of the system for Open-day visitors to control the robot using the HTC Vive VR facilities.
3. Investigate collecting image and joint angle data with a view to using the facility to train a Deep Convolutional Neural Network (DCNN) to control Baxter's arms to execute a task in response to visual input

Special Requirements :

Ethics may be required depending on the type of validation undertaken.
CVMA4 desirable.

Uncertainty Quantification in Digital Humanities (ID: 7082)

Owner/Supervisor: [Bjorn Jensen](#)

** Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.*

Description:

Researchers from the Digital humanities rely heavily on computational methods for cataloging, understating and interpreting physical artifacts such as painting, sculptural, books, and sound recordings. However, the research field has yet to fully adopt the digital and often hypothesis-driven approach to scientific discovery. A particular issue in adopting computation methods seems to be an inherent mistrust in computation methods which do not always produce results that are 100% aligned with the researchers' own view. This is typically due to numerical issues and limitation in the computation algorithm. These can however often be quantified and disseminated.

In this project, we will aim to build a software tool which can dissimilate this uncertainty and aid researchers from the digital humanities to reliably and confidently interpret results produced by computational methods.

This project is in collaboration with digital humanities and will be tailored to the successful candidate.

Special Requirements :

None.

Urban Interaction (ID: 7372)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

As people walk through cities, there are numerous opportunities for interaction with other people and the environment around them. This project will look at embedding touch interactivity into the urban environment for playful interaction. For example, this could include augmenting chairs, railings, and other physical features of public spaces with an interactive device.

Special Requirements :

None.

Viking Board Games (ID: 7530)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

[Tafl](#) is a Viking board game. It's a little like Chess or Draughts.

One team has a King and some defender pieces. The other team has lots of attackers. The aim of the game is for the King to escape from the board - or for the other team to catch the king. The board is square - perhaps 9 by 9. Pieces can only move in straight lines (like [Rooks](#) in chess). Variations of the game exist.

This project is about creating a nice GUI-based Tafl board viewer, on top of a Tafl engine. There are lots of ways to do this - perhaps in Java FX, or Python tkinter, or a neat JavaScript library like d3.

You could start by making it a two-player human vs human game. For an extension, you could add an AI component so one human player could take on the computer.

Special Requirements :

None.

Virtual reality Workspaces (ID: 7440)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Virtual and Augmented Reality (VR/AR) headsets have the potential to improve upon physical multi-display environments (such as your desk with monitors, or laptop) by rendering wide, ergonomic virtual display spaces that can adapt to the task at-hand. However, there are open questions around how we view, and interact with, these virtual workspaces, how they are presented/placed around us, and even how cursor-style interactions are best facilitated. This research project is open-ended with the intention the student pick one of these aspects to examine in detail, developing and evaluating the given technique using a VR headset.

Special Requirements :

None.

Visualising algorithms for the Travelling Salesman Problem (ID: 6940)

Owner/Supervisor: [David F Manlove](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

In the Travelling Salesman Problem (TSP) there is a set C of cities and a distance $d(i,j)$ between every pair of cities i, j . The task is to form a travelling salesman tour (also known as just a tour) of minimum length. That is, we seek a tour of all the cities that visits each city exactly once, returning to the starting city, such that the total distance travelled is minimised.

For example suppose there are 4 cities, labelled w, x, y and z , where the distance between every pair of cities is 1 unit, the distance between w and y is 2, and the distance between x and z is 2. Then the tour $w-x-z-y-w$ has length 6, whereas the tour $w-x-y-z-w$ has length 4, which is minimum.

TSP is “NP-hard”, which means that it is computationally difficult to find an optimal solution. That is, there is unlikely to be an efficient algorithm to solve the problem optimally. Instead, researchers have considered “approximation algorithms” and heuristics, which aim to find solutions that are close to

optimal within a reasonable length of time.

The aim of this project is to explore approximation algorithms and heuristics for TSP with the aid of a Graphical User Interface. The idea would be to display instances of TSP and then show approximate solutions obtained from different algorithms. This would work well for instances of “Euclidean TSP”, in which the cities are points in the plane and the distance between each pair of cities p, q is just the Euclidean distance (i.e., “as the crow flies”) between p and q .

Examples of different algorithms that could be implemented include the Nearest Neighbour heuristic, the 2-approximation algorithm based on computing a minimum weight spanning tree, and Christofides' algorithm.

This project would suit a student who has taken a course in data structures and algorithms either during their Masters programme or during their undergraduate degree.

Further information:

http://en.wikipedia.org/wiki/Travelling_salesman_problem

http://en.wikipedia.org/wiki/Christofides_algorithm

Special Requirements :

None.

VR and Digital Tourism (ID: 7373)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Unstewarded historical sites can be filled with interesting stories, histories, and memories that are impossible to experience for the average visitor. For example, the Clyde waterfront has a long history of innovation in shipbuilding and labour activism, but this isn't easily visible as you walk along the waterfront. Only a detailed knowledge of history can add depth the this important historical site.

Using VR, one can recreate historical settings in an immersive way that shows visitors interesting historical content. This project will explore creating a tourism application for the Clyde waterfront using a Gear VR as a historical telescope.

Special Requirements :

None.

VR Chat and Developing Cultural Norms and Standards (ID: 7375)

Owner/Supervisor: [Julie Williamson](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Virtual hang out and chat spaces are shaped by the technical capabilities and features implemented in the chat application. This project will explore how different chat features lead to different cultural norms and standards to develop a core set of features that support positive and productive virtual social spaces.

Special Requirements :

None.

VR spectators (ID: 7461)

Owner/Supervisor: [Stephen Brewster](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Whilst previous work has examined VR usage in high-throughput spectator situations (e.g. museums), physical and cost constraints dictate that not all spectators may join in on the VR experience at the same time. This leaves some as passive bystanders, often able to view the activity of the VR user, but unable to interact. This project would look at enabling bi-directional communication/interaction when in VR, such that the VR user can see/hear those spectators, and spectators can both see the VR user's view, and both draw their attention to areas, as well as influence events in VR.

Special Requirements :

None.

Web-based tool for map projection (ID: 7389)

Owner/Supervisor: [Alice Miller](#)

* Suitable for both IT/SD and CS+.

Description:

This project will suit a student who has some background in applied maths and/or geographical science (or similar). It involves programming an app (probably web-based) to illustrate different map projections. The project will be undertaken in conjunction with Dr Elizabeth Petrie, School of geographical earth sciences.

Some background on map projection be found below:

<http://geoawesomeness.com/best-map-projection/>

https://www.revolvvy.com/main/index.php?s=Tissot%27s%20indicatrix&item_type=topic

Maths etc on Map projections:

Snyder, John P. (1987). Map projections—A working manual.

A Java applet:

<http://www.se16.info/js/mapproj.htm>

Special Requirements :

None.

Who's There? Occupancy Detection with Smart Sensors (ID: 7528)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+. *Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Most public meeting rooms in the Sir Alwyn Williams Building are equipped with Raspberry Pi based sensor boxes. Live feeds of these sensors are available [online](#).

This project involves inferring how many individuals are present in a meeting room, based on the sensor readings. This requires some machine learning, combined with some scripting to read sensor data. The ideal output from this project is a web front-end that shows live room occupancy predictions for each sensor-equipped room in our building.

Required background:

- - Electronics experience
- - Some understanding of machine learning
- - web development

Special Requirements :

None.

Why is This Sensitive?: Visualising Important Sensitivity Classification Features (ID: 7355)

Owner/Supervisor: [Graham McDonald](#)

* Suitable for both IT/SD and CS+.

Description:

Government departments, such as the Ministry of Defence, produce a large amount of written communications that are essential for the international operations of government. Freedom of Information laws legislate that these documents should be released to the public. However, many of these documents contain sensitive information and, therefore, the documents need to be reviewed to identify and withhold any sensitivities. The aim of this project is to develop a web-based sensitivity review application that integrates automatic sensitivity classification. The system should be able to display documents to a reviewer along with a classifier's predictions as to whether the documents contain any sensitive information. Importantly, the system should provide the reviewer with useful information to explain why the classifier has predicted that a document is (not) sensitive, such as colour coded highlighting of important term features within a document. The system should produce useful reports to visualise the similarities and differences between what the classifier believes to be the most important features of sensitivity and features that are present in the collection being reviewed. Moreover, the system should also allow the reviewer to provide positive and negative feedback to the classifier about document features that the reviewer believes to be indicative of (non-)sensitivity, and inform the reviewer of how this new information will be integrated to update the current classification model. The student will be supplied a collection of government documents and will evaluate the system using this collection and user studies.

Special Requirements :

None.

Word Clouds for Haskell Source Code (ID: 7527)

Owner/Supervisor: [Jeremy Singer](#)

* Suitable for both IT/SD and CS+.*Project may require participation from people other than the student and the supervisor as part of the evaluation.

Description:

Word clouds are beautiful visual representations of a corpus of text. Generally, the larger an individual word, the more frequently that word occurs in the text. There are lots of word cloud generator websites (e.g. wordle.net) and word cloud libraries (e.g. https://github.com/amueller/word_cloud).

Standard word cloud generators use natural language processing algorithms like stop word elimination and stemming to make the content more meaningful. Word cloud generators for programming language source code need more intelligence:

* they should recognise language keywords, and possibly remove them.

- * they should recognise multi-word variable names, and possibly split them up.
- * they should enable high-level navigation into the source code from a clickable word cloud.

This project involves developing a word cloud generator for Haskell source code programs. As an extension, it could become a word cloud generator for multiple programming languages.

Students should be interested in:

- * programming languages
- * text processing
- * data visualization

Special Requirements :

None.

Writing a compiler for a small language (ID: 7330)

Owner/Supervisor: [John T O'Donnell](#)

* Suitable for both IT/SD and CS+.

Description:

The aim of this project is to write a complete compiler that takes input written in a small programming language and translates it to assembly language. The basic part of the project uses a simplified syntax that is easy to parse into an abstract syntax tree; if time permits the second phase of the project uses a parser to handle conventional language syntax.

The source programming language should include assignment statements and basic control structures. The target assembly language can be for any suitable architecture.

The product would be a program that compiles the source language into assembly language and then runs it.

Special Requirements :

None.