400-Level Project Proposals 2018

A list of possible 400-level projects is given below. You do not have to restrict yourself to this list, and can make up your own project topic. However, to do so you will have to arrange for a supervisor. You can find more information about the research interests of the staff members on the Computer Science Department Web pages.

You should email your project selections to the project co-ordinator, 400projectadmin@cs.otago.ac.nz, by **Friday 2nd March**. Before making these selections you should talk to the supervisors of the projects you are interested in. In most cases the best way is to email them to make a time to do so – supervisors' email addresses are given with each project.

In order to allocate the projects as fairly as possible we ask that you give first, second, and third choice projects, and that you choose projects from three different supervisors.

There are two different project papers - COSC480 and COSC490. The relevant Web pages (for COSC480 and COSC490) have more details, but the short version is that if you are enrolled in the Computer Science Honours degree you take COSC490, students in the first year of a Masters take either paper, and most others take COSC480. Most projects are suitable for either COSC480 or COSC490, but a few might be only suitable for COSC480, which will be indicated after the project title.

1. **Deep Learning for Identifying Clouds**

   Andrew Trotman (andrew@cs.otago.ac.nz)
   David Eyers
   Brendan McCane

   "Whether the weather be fine, or whether the weather be not, whether the weather be cold, or whether the weather be hot, we'll weather the weather, whatever the weather, whether we like it or not?" (anonymous). One thing that will help us weather the weather is knowing what to expect. Cheap personal weather stations are readily available and can tell us the current temperature, humidity, and barometric pressure. They cannot predict the weather, and they do not tell us anything about the clouds. In this project we'll use a neural network (possibly TensorFlow) to learn to classify different cloud types then install a digital camera (probably a smart phone) to continuously identify what's happening in the sky. This, along with readings from the weather station will be used to predict the short-term (i.e. today's) weather.

2. **How to Stem**

   Andrew Trotman (andrew@cs.otago.ac.nz)

   Stemming is the process of transforming different word variants into a common base (e.g. swimming, swam, swims → swim). Search engines use stemming to increase the quality of their results. There are many different stemming algorithms including Porter, Krovetz, Paice-Husk, and so on. In this project, we will investigate a question about stemming: How can we predict which stemming algorithm will be most effective on a document collection we have never seen before? That is, what is it about the document collection (and queries, and ranking function) that make the different stemming algorithms effective in different places?

3. **Query Performance Prediction**

   Andrew Trotman (andrew@cs.otago.ac.nz)

   There are a number of ways to predict how good a search engine's results are likely to be before the search engine does the work of evaluating the query. There is some uncertainty as to whether or not these Query Performance Predictors work or not. In this project we will take a new approach to answering this question – we will use them as ranking functions. Specifically, we will use them to re-organize the top search results after the query. If QPP works then this should increase the quality of the results. If it does not then it will not.

4. **Searching Source Code**

   Andrew Trotman (andrew@cs.otago.ac.nz)

   Google has all but abandoned source code search – which is frustrating to many programmers. In this project we will examine state of the art in source code search (via literature review), add source code search to an open source search engine, and measure its performance on a set of code (which we'll extract from github). It is not clear what features best make results useful, but undoubtedly github forks, dates, authors, and so on will be important.

5. **Extension of a teaching OS (xv6) on Raspberry Pi Xv6 is Unix Version 6** (COSC480 only)

   Zhiyi Huang (zhuang@cs.otago.ac.nz)

   I ported it to Raspberry Pi (`https://github.com/zhiyihuang/xv6_rpi2_port`). Now it can run on Raspberry Pi 2/3. However, it does not support the virtual memory, e.g. cache, well enough. Other extension could be USB keyboard, network, etc. Since xv6 is a full-fledged OS, you will be able to master knowledge on OS, ARM architecture, Raspberry Pi, etc, regardless how much extension you could achieve. The final project description will be by negotiation.

6. **Sensor fusion with accelerometer, gyro, and magnetometer**

   Zhiyi Huang (zhuang@cs.otago.ac.nz)

   Sensor fusion is a very important technique in motion tracking using accelerometer, gyro, and magnetometer. We have used Madgwick filter for sensor fusion in tracking arm movement. However, to track the motion more precisely, the sensors need to be manually calibrated. However, due to the noise, for sensors like magnetometer, its readings could be very different from time to time. Could we create a dynamic calibration method to solve the problem? If you are very keen to solve this problem, come to talk to me in details.

7. **Gestures in the Sandbox**

   Steven Mills (steven@cs.otago.ac.nz)

   We have recently built and augmented reality sandbox (see `https://www.youtube.com/watch?v=CE1B7tdGCw0` for an example). People interact by moving the sand, and there is simple recognition of hands to create a rain effect. We would, however, like to add more complex gesture recognition, to allow users to point or otherwise identify areas of interest. This could be used to add additional functionality for teaching or outreach activities. This project would look at extending an existing AR Sandbox implementation to add gesture recognition and using it to implement some demonstration applications.

8. **Deep Feature Detection**

   Steven Mills (steven@cs.otago.ac.nz)
   Lech Szymanski

   Deep networks have been very successful lately for a number of tasks, and recently stacked hourglass networks have been shown to be useful for finding specific features in images, such as leaves on plants. We are interested in applying these methods to a number of applications, including architectural features for pedestrian navigation and individual letters for handwritten document analysis. This project would involve implementing stacked hourglass networks and testing them in one or more of these application areas.

9. **Line-Based 3D Reconstruction**

   Steven Mills (steven@cs.otago.ac.nz)

   There are established techniques for creating 3D models from collections of images. These are usually based on point features in the images, which leads to a representation in terms of 3D point clouds. Scenes of and inside buildings, however, are often rich in line features. This project will look at existing techniques for using line features in 3D reconstruction, implement some of them, and work towards making 3D line sketches from images of buildings.

10. **Vision with a Moving Light**

    Steven Mills (steven@cs.otago.ac.nz)

    Many computer vision techniques assume that the lighting in a scene is constant over time. One case where this assumption is broken is when the light moves with the camera. This happens when people take images with a flash; in endoscopic surgery; and when mobile robots operate in dark environments. This project will look at optical flow, one of the basic ways in which motion is estimated, and how established algorithms can be modified to relax the constant lighting assumption.

11. **Voxel-Based 3D Models**

    Steven Mills (steven@cs.otago.ac.nz)

    Many current methods for 3D reconstruction are based on point clouds, or triangulated meshes. While point clouds arise naturally from structure from motion, and meshes are convenient for rendering, voxels (regular cubes of space, like in Minecraft) are a promising intermediate representation for many applications. This project will explore the creation of voxel-based models by extending existing reconstruction methods and their application to a field such as scene understanding or mobile robots.

12. **Shape Analysis of Stone Tools**

    Steven Mills (steven@cs.otago.ac.nz)

    Pre-European Māori artisans produced some of the world's most sophisticated stone tools, and with researches in Archaeology we are trying to find out more about the processes they used. We do this by making 3D models of incomplete tools, and the flakes and fragments that are left behind during their manufacture. This project will look at ways to recover more information about the shape of these artefacts. Possibilities include detecting important features of the flakes, or looking at what the curvature of the surfaces can tell us about them.

13. **Quantifying conceptual density in text**

    Anthony Robins (anthony@cs.otago.ac.nz)
    Alistair Knott

    Can we identify the density of connections between individual concepts in a body of text? In other words, can we identify discrete concepts, and quantify the relationships between them? This project will involve investigating existing natural language processing tools, and perhaps developing our own.

14. **Attractor spaces in Hopfield nets**

    Anthony Robins (anthony@cs.otago.ac.nz)

    Hopfield networks (a kind of neural network) have dynamic behaviour that can be characterised in terms of gradient descent in a multidimensional attractor space. We know a fair bit about the structure of such spaces. This project will involve implementing and exploring a Hopfield type network to further develop our understanding of its dynamic behaviour.

15. **Serial learning in neural networks**

    Anthony Robins (anthony@cs.otago.ac.nz)

    Neural networks are very powerful learning systems. But most of them work best when all training data is available simultaneously (concurrent learning). In the real world humans and other animals learn different things on different occasions (serial learning). This project will review the current state of serial learning in neural networks, and further explore my "pseudorehearsal" solution.

16. **Extending DrJava (Anthony Robins)** (COSC480 only)

    Anthony Robins (anthony@cs.otago.ac.nz)

    DrJava is the IDE used in COMP160. It is open source. The purpose of this project is to extend the IDE to provide support for porting a Java program to an Android device. The goal is to make the process as simple as possible, so that COMP160 students can port simple graphical programs to Android. If time allows other platforms (iOS) and other extensions to DrJava could be considered.

17. **Partially Observable Markov Decision Process for scheduling wireless communication**

    Haibo Zhang (haibo@cs.otago.ac.nz)

    Partially Observable Markov Decision Process (POMDP) is a generalization of Markov decision process by permitting uncertainty on Markov states. It is general enough to model a variety of real-world sequential decision processes including robot navigation and automated planning. This project aims to apply it to solve a wireless communication scheduling problem. What are expected to be done include:

    - Re-implement the incremental pruning approach for POMDP based on an existing source package that is complex and unfriendly to use.
    - Use the new implementation to solve a wireless scheduling problem.

18. **Heritage Document Analysis with Deep Networks**

    Lech Szymanski (lechszym@cs.otago.ac.nz)
    Steven Mills

    The Hocken Collections have a large collection of documents of historic importance to New Zealand. Many of these are hand-written and so not well suited to traditional character recognition techniques. However, some of them have been transcribed so we have some training data for machine learning techniques.The aim of this project is to apply convolutional networks to character or word recognition in these documents. This would allow the un-transcribed documents to be searched, making the document collections much more readily accessible to researchers.This project will give you a background in machine learning and practical experience with Tensorflow.

19. **Game Engine in Wireless Communication**

    Haibo Zhang (haibo@cs.otago.ac.nz)
    Aleksei Fedorov

    Existing open source wireless channel simulators do not take into account the realistic 3D environment for simulating signal propagation. Most of them generate the wireless propagation channels based on statistical characteristics of a specific environment. It is obvious that the same average channel won't work well in other environments. We propose to take into account the accurate 3D representation of an environment including landscape and buildings during channel simulation. This project aims to create a comprehensive simulation platform for wireless signal propagation based on realistic environment and visualize it in a game engine. We already have well-developed theory for channel simulation, and a 3D representation of central Dunedin with centimetre-level accuracy offered by School of Surveying. If you're familiar with game engines such as Unity Engine and Unreal Engine, this opportunity is for you! What are expected to be done include:

    - Import Dunedin CBD model into a game engine and add missed parts if needed.
    - Implement the ray tracing approach for simulation of signal reflection.
    - Investigate what is better in terms of efficiency: mesh objects or built-in colliders.
    - Simulate a number of users and a number of cellular base stations.

20. **Simulator for Li-Fi: data communications through LED light bulbs**

    Yawen Chen (yawen@cs.otago.ac.nz)
    Haibo Zhang

    LED bulbs can provide more than just illumination! Recent development of "Li-Fi" is an innovative way of wireless optical communications using the signals sent through the light bulb instead of radio waves as in Wi-Fi.The lightbulb is flicked on and off very quickly, up to billions of times per second. That flicker is so fast that the human eye cannot notice it. There are many advantages and potential applications to use Li-Fi in the future. One of the problems is how to design the layout of LED bulbs according to different application scenarios and requirements (e.g.illumination level/transmission rate/energy saving). This project aims to explore the characteristics of Li-Fi and design a simulator for Li-Fi, which is useful for the usage and development of Li-Fi. (`https://www.lifi.eng.ed.ac.uk`)

21. **AI in wireless communication**

    Haibo Zhang (haibo@cs.otago.ac.nz)
    Aleksei Fedorov

    Recently, we developed a novel localization algorithm that uses 4G LTE signals to localize users. Yes, your location can be well predicted if your cell phone is with you. The algorithm is based on the mathematical model of signal propagation. However, the real experiments on Massive MIMO prototype show that there are many stuff happening during propagation, which cannot be described by the mathematical model. The project aims to implement AI approaches in wireless communication to enhance the performance of localization algorithm and extend our understanding of how signals are propagated in the real world. If you're well-grounded in artificial intelligence and machine learning, and interested in wireless, this opportunity is for you! What are expected to be done include:

    - Develop an AI scheme to localize users using measurements from real experiments.
    - Generalize its predictions to the unseen measurements such that it can predict locations of users.
    - Investigate which is better in terms of localization accuracy: AI or the developed algorithm.

22. **Tracking the Motion of the Human Body using 3D Inertial Measurement Units** (COSC480 only)

    Haibo Zhang (haibo@cs.otago.ac.nz)
    Zhiyi Huang
    Steven Mills

    Human motion tracking has a wide range of applications such as in sports, gaming, and medical rehabilitation. In this project we are going to develop a motion capture system using wireless sensors and 3D inertial measurement units. Please watch the following YouTube videos to get a feeling of such systems.

    - `https://www.youtube.com/watch?v=nRzS39UtTF4`
    - `https://www.youtube.com/watch?v=b9TQcv0Asys`
    - `https://www.youtube.com/watch?v=tyiuocPvUes`

    We have already designed the hardware, implemented the communication protocol, and carried out some preliminary work on animating the captured body motions using blender. In this project we will add more functions and produce a demo for real-time body motion tracking. Background and experience with computer graphics and Python programming would be helpful to work on this project.

23. **Learning to predict the consequences of actions**

    Brendan McCane (mccane@cs.otago.ac.nz)

    This project has to do with robot self-learning. Just like an infant, it would be useful if robots could learn about their own capabilities and refine them to achieve certain goals. We have a robot arm and some cameras. The goal of the project will be to develop a system that can predict the next image that will appear in the camera based on the previous image and the motor actions of the robot arm.

24. **Simulator for Multicasting in Optical Network on Chip (ONoC)**

    Yawen Chen (yawen@cs.otago.ac.nz)
    Haibo Zhang

    With the increasing number of CPU cores on a processor chip, conventional electronic interconnects for core-to-core communication are becoming a bottleneck. Recent rapid advances have made the optical-based interconnection, Optical Network-on-Chip (ONoC), an attractive solution to breakthrough electronic interconnect limitations. Multicast communication, in which packets from one source need to be sent simultaneously to multiple destinations, intensively exists in chip multiprocessors due to the need for cooperative computing and cache coherence. This project aims to simulate the behaviour of multicast traffic on Optical Network on Chip (ONoC). The simulator will be very useful for the hardware design and development.

25. **Machine learning for remote robot control**

    Lech Szymanski (lechszym@cs.otago.ac.nz)
    Brendan McCane

    The aim of this project is to implement a software agent capable of training itself to remote control a robot. The setup consists of a robot in an environment and a camera capturing the scene from above. The software agent will interpret images from the camera, and send commands to the robot. The student will build a learning model (most likely a neural network) into the agent. The aim is for the agent to teach itself how to pilot the robot to an arbitrary spot in the environment by sending commands and observing the effects through the camera.

26. **Computational methods for genetic interactions detection**

    Alex Gavryushkin (alex@cs.otago.ac.nz)

    Understanding how genes define functions of an organism, so-called genotype-phenotype association, is a central problem in modern biology and biomedicine. When the effect of a certain gene depends on the genetic background, we say that genetic interactions are present in the system. These genetic interactions often produce unexpected effects on the organism, and understanding them is crucial in, for example, cancer treatment. In this project, you will design and implement a software tool for detecting genetic interactions in genotype-phenotype association data. Following recent mathematical and computational development (`https://elifesciences.org/articles/28629`), geometric methods can be used to understand the fine structure of interactions in genomic systems. However, a readily available software package that can be used by the wider biomedical community is still missing. In this project, you will contribute to the design and development of such a package.

    No biological background is necessary for this project but a strong interest in biomedical applications and willingness to learn the basics is desirable. The project would suit someone with a solid background in computer science (or mathematics) willing to continue their study at Masters and/or PhD level in the future. Excellent programming skills are essential. Prospective students are strongly encouraged to approach Alex Gavryushkin to learn more about this work before committing to the project. An important technical detail is that Alex will work and supervise the project from Germany from 1 June onward, so the main stages of the project will need to be completed by then.

27. **Web-based framework for AI competitions** (COSC48 only)

    Lech Szymanski (lechszym@cs.otago.ac.nz)
    David Eyers

    The second assignment in COSC343 is about creating agents that learn how to survive as predators or prey in a simple simulated environment. In the future we would like to stage rounds of online competitions, where these agents can battle/interact with each other. This project is about engineering a web-based framework for running these competitions. Most of the work will be integrating existing technologies (such as git and programming contest software systems) to accept source code submissions, auto-build them, run agent duels and to provide feedback to the participants and overall scoring.

28. **Timing evolutionary histories**

    Alex Gavryushkin (alex@cs.otago.ac.nz)

    Phylogenetic trees are used throughout evolutionary biology to describe the ancestral history of organisms. Typically molecular sequence data is used to reconstruct the mutation history and visualise the result as a phylogenetic trees. Following recent advances in molecular biology, mathematics, statistics, and computing, it is now possible to time phylogenetic trees, that is, to estimate the time of mutation (divergence) events. Among many important applications of this method is the reconstruction of the HIV spread history. Despite these successes, the underlying mathematical structure of phylogenetic time-trees (evolutionary histories) is poorly understood and scalable statistically sound methods are yet to be developed. Our recent advances (`https://link.springer.com/article/10.1007/s00285-017-1167-9`) show that, despite common expectations, many of the methods developed in classical phylogenetics do not translate to when it comes to timing evolutionary histories. Specifically, new computational methods and algorithms need to be designed for this class of trees. In this project you will contribute to the ongoing work aimed at developing sampling algorithms over the space of time-trees.

29. **Medical Image Segmentation**

    Brendan McCane (mccane@cs.otago.ac.nz)
    Ming Zhang (Department of Anatomy)

    The aim of this project is to develop automatic or semi-automatic segmentation methods for segmenting images of sliced plastinated bodies or body parts, with a view to generating 3-dimensional visualisations of the anatomical structures. The data set is rather unique, and will allow anatomists and surgeons get a much greater understanding of the soft tissues in the body. In collaboration with a colleague in the Department of Anatomy.