My background is a BSc (Hons) in Computing Science from the University of Glasgow and an MPhil in Computational Biology from the University of Cambridge.

I am currently a Computing Science PhD student at the University of Glasgow. My PhD is on computational modelling and analysis techniques to study biological systems. More specifically I am interested in cellular signalling, the complex process by which a cell responds to its environment. Study of cellular signalling is extremely important in the fight against diseases such as cancer as mutation to the signalling process is often implicated in these diseases.

I was awarded the Jim Gatheral Scholarship in the first year of my PhD. The £3,200 award allowed me to visit Stanford Research Institute International (SRI) in California for 3 months. I visited Prof. Carolyn Talcott in the Pathway Logic group at SRI.

**Biological Background**

A cell has many different receptors on the cell wall that various biochemical signals attach to. A signalling network is a collection of biochemical reactions that govern how the cell responds to the incoming signals. Typical responses include genes being expressed in the nucleus, cellular differentiation and phenotype changes.

A signalling pathway is a series of biochemical reactions from a signal to some cellular output e.g. signal-receptor binding causing downstream protein activation in turn causing genes to be expressed. The thesis of my PhD is that there is much to be learned by treating a complex signalling network as a collection of signalling pathways and interaction between pathways.
Pathway Logic Project
A common approach to modelling a signalling network is to manually build the model using details in the literature, experimental data and personal knowledge. The Pathway Logic project is an approach to modelling that automatically creates a model from its database of biochemical reactions curated from the literature. The current signalling network model contains 675 biological entities and 542 biochemical reactions.

This project was immediately appealing to me as it provides an excellent platform to develop and test new computational algorithms on. Upon arriving at SRI I discovered that the team can generate a single pathway in the model to some output, however they cannot generate all pathways. This was the problem I worked on during my visit.

Research Outcomes

Current (Steady State) Techniques
A steady state technique of interest in this work is T Invariant analysis in which a sequence of reactions to achieve steady state is found. Hence, the concentration of each biological entity in a cell is the same before and after this sequence of reactions has occurred. The set of T Invariants in a model are computed as follows.

The stoichiometric matrix of a model is a matrix $C : P \times R \to Z$. Hence $C(p, r)$ is the chance in concentration of protein $p$ after the reaction $r$ occurs. The T Invariants in a model is found by enumerating all minimal nontrivial non-negative integer solutions to: $C \cdot y = 0$. 

Figure 1: A signalling pathway from a signal causing a protein to become activated, then causing a gene to become expressed (turned into a protein).
T-Invariant analysis has been applied to better understand both metabolic systems and signalling networks. A metabolic system is the set of chemical reactions that transform a nutrient set into a set of metabolites in order to maintain life. The study of metabolite output is suitable for steady state techniques as there is a constant uptake of nutrients and subsequent transformation into metabolites. In signalling networks the set of T-Invariants relate to the set of pathways in the network. However, during my visit to SRI I found that under certain conditions this approach produces incorrect results. These problems are due to a steady state approach being used to compute dynamic flows through a model. I found three model structures that are typical within signalling networks that make steady state analysis inappropriate.

One such structure is a trap. A trap is a set of biological entities in a model in which the concentration can never reach zero after becoming non-zero. An example of this is an enzymatic reaction, used often in signalling pathways, shown in Figure 2. In this case, steady state techniques cannot describe pathways.

![Figure 2: An enzymatic reaction turning a substrate into a product.](image)

There is a trap on \{Enzyme, Enzyme–Substrate\} because the concentration can never return to zero after being non-zero.

**New (Dynamic Flow) Technique**

To overcome the challenges of using steady state techniques to compute all pathways in a model, I looked at the area of model checking. A model checker is a piece of software that returns the truth value of a statement about a model written in a formal language. An interesting statement for our purposes is that a biological output cannot be produced, e.g. “Gene X can never be expressed”.

If the statement is false, then all counter examples can be produced. In this case the counter example would be a sequence of reactions that result in Gene X being expressed, hence a pathway. Model checking produces a counter example with the fewest number of reactions to each state that violates the statement.

The method I have developed is an adaption of current model checking techniques.
To illustrate the difference between my method and model checking counter examples, consider the example model to the right. There are two pathways between Input and Output however model checking techniques would only find one pathway, the shortest length one. My technique will produce the direct pathway and the indirect, though distinct, pathway through X.

**Result:** The new method produces all pathways in a model to an output even in cases where steady state techniques would fail.

**Impact of Jim Gatheral Scholarship**

The scholarship has allowed me to work with a great team of researchers in the world famous Silicon Valley. I learned a lot from this visit, especially to consider any computational tool that ‘does the job’ and the importance of interdisciplinary collaboration.

After the visit, we submitted a paper on the findings:


The collaboration with Prof. Talcott and the Pathway Logic team will continue on as we have identified several ideas made possible through the work in my visit. I already have ideas on improving the efficiency of the new algorithm to compute pathways such that we can then handle bigger models. I also wish to see how the set of pathways differs between normal and mutated (e.g. cancerous) cells – perhaps mutated cells contain extra pathways or lose important pathways.

A bigger question is to identify knockout targets, a protein within a cell whose removal would stop a biochemical output being produced. We can now achieve this by finding proteins that are common to all pathways to the output. I wish to extend this to finding knockout targets that are specific to the output, hence that minimise effects on other biochemical outputs. These relate to very specific drug targets to control disease.

The Jim Gatheral scholarship has also allowed me to have some non-work related fun. Apart from weekend trips up to San Francisco and Berkeley, I spent the Christmas period in San Diego and Palm Springs. It certainly was a change from the usual Scottish winter!

I wish to thank the Jim Gatheral scholarship for this opportunity – it has had a profound impact on my PhD studies.
A trip to Joshua Tree National Park during my visit to Palm Springs.

Sunny Stanford ... in November.