Proceedings of the University of Oxford Department of Computer Science Student Conference 2011

Co-Chairs: Sophie Kershaw, Katie Leonard

18 November, 2011

Contents

Foreword

It gives us great pleasure to introduce the Proceedings of the 2011 Student Conference here at the University of Oxford's Department of Computer Science. Long standing members of the Department will hopefully agree that the Conference has gone from strength to strength over the years, growing from a biennial event to its current annual form. This year's event has included talks, a poster session and discussion throughout the day, and has drawn submissions from almost every research group in the department. The Conference aims to provide a forum for discussion of recent and ongoing work, a chance for current students to showcase their research, and an opportunity for the members of department – staff and students alike – to come together and acknowledge our shared aims of pushing research forward in Computer Science.

Following a number of years where the conference has been externally hosted, 2011 has seen the event return to the Department itself, with talks held in Lecture Theatre A and the poster session, wine reception and refreshment breaks held in the Atrium. Hopefully this helped to raise the profile of the Student Conference within the Department and encouraged your participation. The evening meal at Rewley House proved a pleasant end to the day and was attended by thirty current students, academics and their guests.

The quality of submissions this year was high and reducing these to a programme of accepted talks proved very difficult. Thankfully our extensive team of reviewers did a thorough job of appraising the contributions; our aim was to provide authors with constructive criticism and encouragement and we hope that we achieved this objective. The Selection Committee comprised postdoctoral researchers, current DPhil students and academics: furthermore, every paper was reviewed by at least one postdoc to ensure that an experienced eye had been cast over the work. Eventually eleven talks were accepted for presentation, with each speaker allocated twenty minutes for a talk followed by five minutes of questions from the audience. The remaining contributors were invited to present their work in the Poster Show in the final session of the day; moreover, some members of the department specifically chose to present posters rather than talks. We would like to thank all those who took the time to submit abstracts to the conference –

we enjoyed hearing about your work and hope your interest in the Student Conference will continue in years to come.

This year the conference was advertised to the undergraduates of both the Mathematics and Computer Science faculties, although it is a shame that attendance from this contingent was not overly apparent on the day. It would be really good to encourage our more junior members to attend in future years and we hope that this will be the case next year. In particular, our annual conference should prove useful for prospective research students to find out about the diverse research topics covered by the department.

The conference itself reflects the combined efforts of various individuals, without whose help and support the event would not have come together as it did. It is of course impossible to list all those we would like to thank within this Foreword – however, the following pages detail the names of those who contributed, whether through reviewing abstracts as part of our Selection Committee, providing administrative assistance during the organisation of the event, or providing technical or logistical support on the day itself.

Nonetheless, particular names call for special mention here. Our keynote speaker this year was Prof. Jim Davies of the Software Engineering group, whose sterling efforts in judging the talks throughout the day were greatly appreciated. Jim also opened the programme of talks with an engaging keynote speech and awarded the prizes for the best talks and best poster at the end of the wine reception. Thank you also to our new Director of Graduate Studies, Prof. Luke Ong, for his continued support for the Conference and for his role as part of the Steering Committee. Many thanks also to the Department of Computer Science for providing the generous funding that enabled the Student Conference to take place.

Last but by no means least, the Organising Committee owes a particular debt of gratitude to Julie Sheppard, whose endless patience, advice and support has been an immense help throughout the organisation of this year's Student Conference.

We wish next year's Organising Committee every success in carrying the event forward in 2012. Finally, we would like to take this opportunity to wish all members of the Department well for the coming year.

Sophie Kershaw and Katie Leonard
Co-Chairs of the 2011 Student Conference

Organisation

Organising Committee

Sophie Kershaw*, Katie Leonard*, Joe Loughry

* denotes co-chair

Steering Committee

Julie Sheppard, Luke Ong

Selection Committee

Jan Botha, Rafel Bordas, Kevin Burrage, Ciara Dangerfield, Thomas Dunton, Shamal Faily, Eoin Hyde, Daniel James, Ernesto Jimenez-Ruiz, Sophie Kershaw, Hsiang-Shang Ko, Pablo Lamata, Katie Leonard, Joe Loughry, John Lyle, Eamonn Maguire, Philip Maybank, Gary Mirams, Pras Pathmanathan, Maciej Pirog, Ajitha Rajan, Steven Ramsay, Aadya Shukla, Blanca Rodriguez, Michael Vanden Boom, Jim Whitehead II

Keynote Speaker

Jim Davies

Thanks

Many thanks to Daniel James and Stephen Pulman for chairing; to Valentina Carapella, Sara Dutta and Konstantinos Sakellariou for helping with organisation on the day; to Sara Dutta for taking payments for the conference dinner, and to Lihao Liang for putting up the conference dinner posters.

Special Thanks

Special thanks to Julie Sheppard, Wendy Adams and Maureen York, as well as to all the members of the Finance team and Support staff of the Department of Computer Science.

Preprocessing for Non-Ground Answer Set Programs

Michael Morak

Oxford University Department of Computer Science
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

Answer set programming (short: ASP, [1]) is a declarative problem solving approach that has in recent years become a widely acknowledged paradigm and has gained much traction in the logic programming community, mainly due to its simple, yet expressive modelling language.

It consists of a set of logical rules, each being either ground or non-ground, where non-ground rules include variables that are implicitly universally quantified.

Answer set programming solvers follow a two-step approach: After the problem is modelled as such a set of logical rules, the so-called "grounder" instantiates all these rules by replacing the various variables with applicable constants. This then yields a propositional logic program (consisting of propositional or "ground" rules), that is equivalent for the given domain, and this program is then finally fed into the actual solver. In general, a single non-ground rule can therefore produce a number of ground rules that is exponential in the number of variables in the non-ground rule (i.e., given a domain with d elements and a rule with n variables, the grounder will produce at most d^n rules).

Often the rules exhibit a particular structure and therefore the groundings could be kept small if this were taken into account. However, current generation grounders do not in general do this. Several preprocessing and optimization techniques have been developed in the past, see e.g. [2,3], that work well in practice, however the grounding size still remains exponential in the worst case.

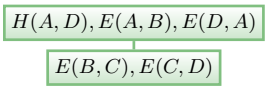
In this paper we present a novel preprocessing strategy that ensures that the grounding remains small by telling the grounder about the structure of the rules. This approach is based on [4], where a similar mechanism is used to efficiently solve the boolean conjunctive query evaluation problem.

For our preprocessing approach, each rule is represented as a hypergraph, where each variable in the rule is represented by a vertex and each predicate in the rule is represented by a hyperedge in the hypergraph. Using a hypertree decomposition of this hypergraph representation, the rule can then be split up into an equivalent set of smaller rules, whose grounding is only exponential in the size of the nodes in the hypertree decomposition (i.e. the number of variables in each node).

Example 1. Given the rule

$$r = H(A, D) \leftarrow E(A, B), E(B, C), E(C, D), E(D, A).$$

we compute a (simple) decomposition $HD(r)$, for instance the following:



This decomposition then yields the following rules (with T_1 being a new temporary predicate):

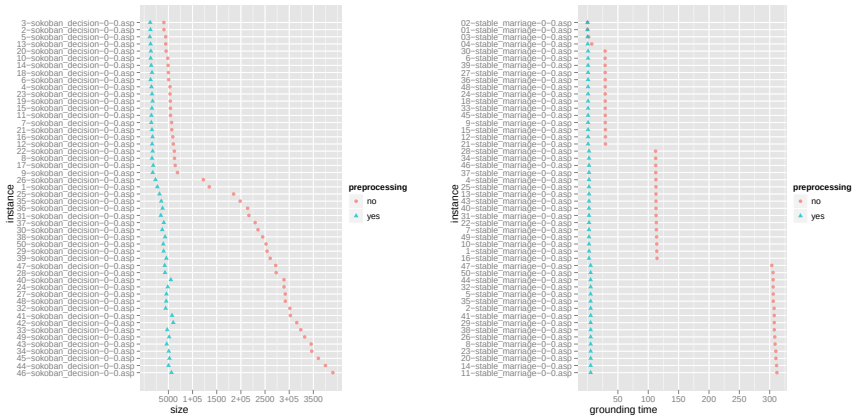
$$T_1(B, D) \leftarrow E(B, C), E(C, D)$$
$$H(A, D) \leftarrow E(A, B), E(D, A), T_1(B, D)$$

This split-up rules contain only the variables that occur in the corresponding node of the hypertree decomposition. Therefore the grounding size clearly is no longer exponential in the number of the variables in the original rule, but only in the number of variables occuring in each node of the decomposition (i.e., $d^k \cdot n$ rules may be generated in the worst case, where k is the maximum number of variables in any node of the hypertree decomposition and n is the size of the original rule). □

In cases where the size of the nodes k remains bounded, the grounding remains polynomial when using current generation grounders. First experiments using a prototype implementation and the benchmarks from the well-known Third ASP Competition 2011[1] show a significant decrease both in grounding time and grounding size for certain problems.

As an example, for the benchmark instances of the Sokoban Decision problem the size of the grounded program decreases by an average of 80% using our approach, as shown in Figure 1(a). The Stable Marriage problem a benchmark aimed at current-generation grounders. For this problem, our approach can significantly improve the time needed for the grounder to calculate the grounding, as shown in Figure 1(b).

Future work will be done in the area of integrating this approach directly into current-generation grounders, as well as well as special-tailored algorithms for the actual solvers, which are currently slowed down by the preprocessing. This happens because of the number of introduced temporary predicates can be quite high, which makes it harder for the solver heuristics to determine the next predicate. However, by making the solver aware of the temporary predicates and their interdependencies, this problem can be remedied.



(a) Sokoban Decision problem

(b) Stable Marriage problem

Fig. 1. Grounding size of Sokoban Decision and grounding time of Stable Marriage.

References

1. C. Baral. *Knowledge Representation, Reasoning and Declarative Problem Solving*. Cambridge University Press, 2003.
2. W. Faber, N. Leone, C. Mateis, and G. Pfeifer. Using database optimization techniques for nonmonotonic reasoning. In *Proc. 7th International Workshop on Deductive Databases and Logic Programming, DDLP 1999*, pages 135–139, 1999.
3. M. Gebser, B. Kaufmann, A. Neumann, and T. Schaub. Advanced preprocessing for answer set solving. In M. Ghallab, C. D. Spyropoulos, N. Fakotakis, and N. M. Avouris, editors, *ECAI*, volume 178 of *Frontiers in Artificial Intelligence and Applications*, pages 15–19. IOS Press, 2008.

[1] `https://www.mat.unical.it/aspcomp2011`

Where is it?

Generating Locative Expressions by a continuous variant of the GRE algorithm

Jamie Frost

Oxford University Department of Computer Science
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

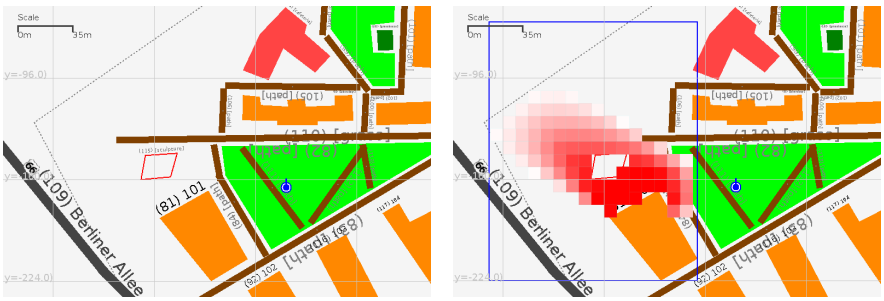


Fig. 1. The output of Algorithm 1 for "Where is the sculpture?" (pictured towards the left of the environment, with the observer position and direction indicated by the blue circle). The resulting description was "The sculpture is right of 101, just left of you and near the laboratory in front of me" when the terminating threshold was set low (i.e. resulting in longer descriptions). The picture on the right indicates the associated field F, with the intensity of red proportional to the probability.

In the variety of linguistic systems that our 'pedestrian robot' is capable of, one of these is the task of generating coherent descriptions of objects in an environment that can disambiguate their location. Much research has been dedicated to a similar question of descriptions which can *uniquely identify* objects, originating with an algorithm presented in [1] (which we refer to as the *GRE algorithm*), and being extended to support relations by [2] and with further augmentation in [3] for example. In the case where all objects are visible or known in location to the user, yielding a description that ascertains their identity trivially also answers the question of where it is, but this in general is not the case. We therefore present an analogous algorithm which instead of disambiguating over distracting objects, disambiguates over points in space. If we have some model $\psi(\mathbf{x})$ which gives us a probabilistic spatial model for each preposition (e.g. for $Between(car_1, tree_1, tree_2)$ or "the car is between the two trees", $\psi(\mathbf{x})$ gives us the confidence in the range $[0, 1]$ that a car situated at point $\mathbf{x}$ satisfies the property of being between the two specified trees), then we require some metric $|F|$ which gives a *soft* measure of the ambiguity of a candidate description over space, i.e. for n relations: $|F| = \int_{-\infty}^{+\infty} \prod_{i=1}^{n} \psi_i(\mathbf{x}) d\mathbf{x}$. In some cases, we can obtain this result analytically without ever needing to evaluate ψ. For example, were we to allow prepositions based only on distance, e.g. *near* and *by*, and represented these using multivariate Gaussian functions whose means were set at each landmark object and covariance matrices set appropriately, then we can obtain an exact value for $|F|$ using calculus. However, since in general our spatial models are non-parametric (e.g. they are based on complex factors such as visibility), we

must instead evaluate ψ for a variety of values in a lattice of points across the environment. By deducing a minimal rectangular region in which we can guarantee all points outside the region have a validity of 0, we need only compute that lattice for this region, and given a quota of points we wish to calculate, can choose an appropriate *granularity* of data, i.e. the spacing of points in the lattice. We insert these into a *quad tree* structure for efficiency, and merge the quad trees for each observation in the candidate description using the appropriate algorithm. $|F|$ can then be computed by the sum over the leaves of the resulting tree.

The algorithm itself is presented below. While there is insufficient space to give full details (or any implementation/efficiency considerations), the algorithm greedily adds new relations to a set, where at each iteration, it chooses the relation which minimises $|F|$. The algorithm continues until some threshold is met (i.e. some tolerance of ambiguity, since descriptions need not identify a single point in space). Any landmark objects in the description must be uniquely described, so we use appropriate variable renaming and apply the GRE algorithm to each landmark involved. An example can be seen in Figure 1.

0. Initialise
$L \leftarrow \{\}$, $F \leftarrow \underline{\mathbf{1}}$, where $\underline{\mathbf{1}}$ is the function $(x, y) \rightarrow 1$.
$P_r \leftarrow$ the set of spatial relations true about the referent r, such that r appears as the reference object (first argument) of each relation.

1. Check Success
if $|F| < threshold(r)$ **then**
  goto Step 4
else if $P_r = \emptyset$ **then**
  fail or goto Step 4 as a compromise
else
  goto Step 2
end if

2. Choose Property
for each property $p_i \in P_r$ **do**
  $N_i \leftarrow N \oplus p_i$ where
end for
Chosen predicate is p_j, where $|F_j|$ has the smallest size.
goto Step 3

3. Extend Description (w.r.t. the chosen p)
$P_r \leftarrow P_r - \{p\}$, $N \leftarrow N \oplus p$
goto Step 1

4. Resolve relative objects to produce final description.
Associate referent r with a unique variable v.
$C \leftarrow \langle C_v = r \rangle$, $L \leftarrow [r'\backslash v']L$
$R \leftarrow$ set of all relative objects used in description L.
for $r \in R$ **do**
  associate r' with a new, unique variable v'
  create a new stack, add $Describe(r', v')$ and apply GRE algorithm to produce description $N' = (L', C')$
  $L \leftarrow [r'\backslash v']L \cup L'$, $C \leftarrow C +\!\!+ C'$
end for
return (L, C)

References

1. Dale, R. *Generating Referring Expressions in a Domain of Objects and Processes.* PhD thesis, Centre for Cognitive Science, University of Edinburgh, 1989.
2. Dale, R., and Haddock, N. Generating referring expressions involving relations. In *Proceedings of the fifth conference on European chapter of the Association for Computational Linguistics* (Morristown, NJ, USA, 1991), Association for Computational Linguistics, pp. 161–166.
3. Horacek, H. An algorithm for generating referential descriptions with flexible interfaces. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics* (1997), Association for Computational Linguistics, pp. 206–213.

Abstract Satisfaction

Unifying Static Analysis and Satisfiability Solving

Leopold Haller, Vijay D'Silva, and Daniel Kroening

Computer Science Department, Oxford University

The problem of program verification is to decide whether a program P satisfies a specification Φ. It is a basic result in computer science that this problem is undecidable [3]. Therefore, approaches that attempt to solve it resort to computing approximate instead of precise answers. One way to do this is to search for a proof of correctness, for example by identifying a property that is shared by every possible run of the program and that is sufficient to establish Φ. Another way is to prove the presence of errors, for example, by providing a witness in the form of a program trace that violates Φ. If neither neither one nor the other can be found, no conclusions can be drawn about the correctness of P.

These two approaches correspond loosely to two dominant verification communities and their respective paradigms. The *abstract interpreter* or *static analysis* approach to verification is to compute an abstract characterisation of program behaviour that preserves only properties of interest. Examples of static analysis include constant propagation or pointer analysis. The theory of *abstract interpretation* [1], provides a formal framework for describing and reasoning about static analyses. The *decision procedure* approach is to approximately reduce a verification query to a satisfiability problem in some decidable logic. Presence of errors is then established by running an efficient satisfiability procedure, such as a modern propositional SAT solver that uses lemma learning [2].

To illustrate these two approaches, consider the program P given below and the property Φ that states that the integer variable x has value greater or equal to zero immediately after executing the assignment.

$$P \quad \hat{=} \quad \mathsf{if}(x < 0)\ x := x * x$$

One static analysis approach to solving this problem is to perform sign analysis by picking values for x from the set $\{-, +, \top\}$, which respectively denote cases where x is negative, positive, or where it's value is otherwise or unknown. At program entry, x is assigned the abstract value "$\top$". Upon entering the conditional statement, x is assigned "$-$". We then abstractly execute the multiplication. Since a negative value multiplied by a negative value is always positive, we obtain that the result must be "+". This information is sufficient to determine that Φ holds. In contrast, for the decision procedure approach one would construct a formula $\varphi \hat{=} (x < 0) \wedge (x * x < 0)$ and decide its satisfiability. If φ is a contradiction, the program is safe. Otherwise, an error exists.

Methods that follow these approaches tend to show typical strengths and weaknesses. Abstract interpreters are often very efficient, since they can evaluate a program over a computationally simple abstraction. The downside is that they introduce imprecision. Consider for example the above program and the property "x may not be 13" immediately after executing the assignment. Since 13 is prime, the property holds, but the sign abstraction loses the information necessary to establish this fact. In contrast, the decision procedure approach tends to be precise, but lacks scalability. In some cases, SAT solvers expensively explore the space of possible solutions even though safety can be determined immediately in an abstraction.

Recent research efforts have aimed at combining the strengths and weaknesses of the two approaches by *shallowly* integrating abstract interpreters and decision procedures and treat-

ing them as communicating black-boxes. We intend to show that this black-box separation is conceptually unnecessary and practically inefficient.

Our work is based on a very simple, but surprising insight: Modern SAT decision procedures *are* abstract interpreters in a precise formal sense. Furthermore, implementations of SAT solvers and abstract interpreters are similar enough to enable immediate technology transfer between the two areas: Decision procedures can be built to operate directly over rich abstract representations of the problem domain which allows them to use advanced domain-specific reasoning in a way that is entirely distinct from existing approaches. Abstract interpreters can adopt techniques such as case-based reasoning and lemma learning to precisely and efficiently refine an abstract analysis.

To give an approximate idea of the identity between SAT and abstract interpretation, consider the following example, where we equate deduction in a SAT solver line-by-line with the corresponding concepts in constant propagation.

Consider the propositional formula

$$\phi = x \wedge (\neg x \vee y)$$

SAT solvers operate over a *partial* function that maps propositions to true or false. The *unit rule* is applied to subformulae to extend an initially empty partial assignment ρ through deduction.
Applying the unit rule to x extends the partial assignment $\emptyset$ to $\rho' = \{x \mapsto \mathsf{true}\}$.
From ρ' and $\neg x \vee y$ we can deduce

$$\{x \mapsto \mathsf{true}, y \mapsto \mathsf{true}\}.$$

Consider the program

$$\mathsf{bool}\ x, y;\ \mathsf{if}(x)\ \mathsf{if}(\neg x \vee y)\ \mathsf{assert}(\mathsf{false});$$

We use the abstract domain of constants, which maps variables to true, false or $\top$. We use the *best abstract transformers* to model execution starting from the initial element $\pi = \{x \mapsto \top, y \mapsto \top\}$.
Abstractly entering the conditional $\mathsf{if}(x)$ from π yields $\pi' = \{x \mapsto \mathsf{true}, y \mapsto \top\}$.
Entering $\mathsf{if}(\neg x \vee y)$ from π' yields

$$\{x \mapsto \mathsf{true}, y \mapsto \mathsf{true}\}.$$

We have established a formal framework that gives an abstract interpretation based account of modern SAT procedures. We identify lattice theoretic generalisations of concepts and techniques in propositional SAT solving, and obtain a generalised satisfiability procedure that is parametrised by a concrete and abstract domain. Instantiations of this generalised procedure yield a new class of efficient satisfiability procedures *and* a new class of program and property dependent refinement procedures for static analysers.

Current results show that a program verifier based on our ideas can significantly outperform existing static analysers and decision procedures in terms of precision and efficiency, respectively. As future work, we intend to demonstrate the efficiency and generality of our framework by instantiating it over a variety of application domains.

References

1. Cousot, P., and Cousot, R. Systematic design of program analysis frameworks. In *POPL* (New York, NY, USA, 1979), ACM Press, pp. 269–282.
2. Moskewicz, M. W., Madigan, C. F., Zhao, Y., Zhang, L., and Malik, S. Chaff: Engineering an efficient SAT solver. In *DAC* (2001), ACM, pp. 530–535.
3. Rice, H. G. Classes of recursively enumerable sets and their decision problems. *Transactions of The American Mathematical Society 74* (1953), 358–358.

Interpreting optical mapping recordings in the ischaemic heart

a combined experimental and computational investigation

Sara Dutta

Computational Biology Group, Department of Computer Science

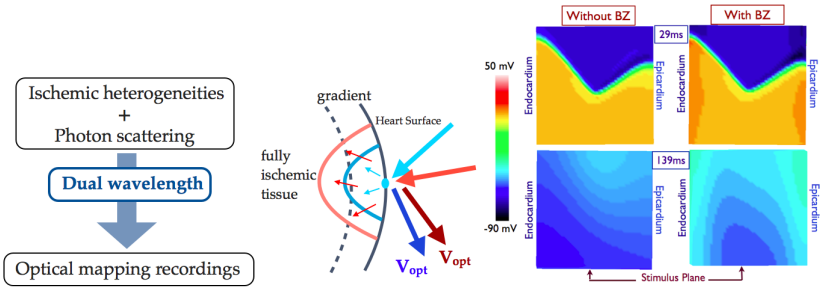
Sudden cardiac death (SCD) accounts for upwards of 60 000 deaths in the United Kingdom, each year [4]. One of the major causes of cardiac arrest is coronary artery occlusion, reducing the supply of blood to the heart, and resulting in a phenomenon known as acute myocardial ischaemia. Since the first cardiac cell model was built over 50 years ago, the use of mathematical models and computer simulations has provided significant insight into multiscale cardiac electrophysiological phenomena and has become a state of the art technique used in the field. We can now simulate whole heart electrical activity on our desktop computers.

Under normal conditions, cardiac cell excitation, also known as the action potential (AP), has a regular and stable shape. During the first 10-15 min of acute ischaemia, the AP displays morphological changes. Ischaemic effects vary not only through time but also spatially [2]. Scientists have shown there exists a layer of cells, referred to as the border zone, at the surface of the ischaemic region that display progressive changes in electrophysiological properties between healthy and fully ischaemic tissue [3]. The highly heterogeneous cardiac tissue that exists during ischaemia affects the success of clinical treatments and increases the likelihood of arrhythmogenesis [6].

Biologists attempt to shed light on underlying mechanisms with optical voltage mapping, a widely used technique for non-contact visualisation of surface electrical activity. Optical mapping utilises voltage-sensitive fluorescent dyes to visualise the electrical activity of the heart. Upon excitation at a specific illumination wavelength, dye molecules transduce differences in membrane potential (V_m) into changes in emitted fluorescence. However, penetration of the illuminating light into the tissue (with depth dependent on illumination wavelength) and scattering of the emitted fluorescent photons, means that the detected signal represents a weighted-average of V_m levels from within a volume of tissue beneath the surface recording site. Such effects have been shown in modelling studies to distort optical recordings, in particular causing a prolongation of the AP upstroke [1].

The distortion effect may be more pronounced in the presence of a border zone during ischaemia, when a gradient of electrophysiological properties exists at the surface of the heart. Therefore, the optical signal gathers information from a volume of tissue which comprises progressive changes in electrophysiological properties and therefore displays different APs. As can be seen in Figure 1a our work focuses on understanding how the two lights (blue and red) can be used to characterise the border zone at the surface of the heart, represented by the dashed line.

Computer models have provided a wealth of information and insight that is not possible to obtain through experiments alone. Computer models allow scientists to change parameters relatively easily, therefore, speeding up scientific investigations. Experiments are used to validate these complex models in order to be able to use them as a predictive tool to guide further experimentation. Computational models of optical mapping recordings have provided a detailed understanding of the precise mechanisms by which photon scattering contributes to the distortion of the optical signal.



(a) Diagram of goal and hypothesis (b) Simulation results

Fig. 1

A combined experimental and computer simulation investigation into how photon scattering, in the presence of ischaemia-induced spatial heterogeneities, distorts optical mapping recordings is performed. Dual excitation wavelength optical mapping experiments are conducted in rabbit hearts, in order to excite a bigger or smaller volume of the tissue at the surface of the heart. Therefore, the resulting optical signal is more or less affected by the border zone. In order to interpret experimental results a computer simulation study using Chaste [5] is performed using a 3D model of ischaemic rabbit cardiac tissue combined with a model of photon diffusion to simulate optical mapping recordings. Simulations were run on two types of tissue: one with a border zone and one without, results of the transmembrane potential across the surface of the slab of tissue during activation and repolarisation can be seen in Figure 1b. Our simulations show that the presence of a border zone, in combination with fluorescent photon scattering, distorts the optical signal. Future work will involve applying this technique during chaotic and heterogeneous excitation patterns, such as conduction block and reentry, to characterise sub-surface activity.

References

1. Bishop, M., Rodriguez, B., Eason, J., Whiteley, J., Trayanova, N., and Gavaghan, D. Synthesis of voltage-sensitive optical signals: application to panoramic optical mapping. *Biophys. J.*, 90 (2006), 2938–2945.
2. Carmeliet, E. Cardiac ionic currents and acute ischemia: from channels to arrhythmias. *Physiol. Rev.*, 79 (1999), 917–987.
3. Fiolet, J., Baartscheer, A., Schumacher, C., Terwelle, H., and Krieger, W. Transmural inhomogeneity of energy metabolism during acute global ischemia in the isolated rat heart: dependence on environmental conditions. *J. Mol. Cell. Cardiol.*, 17 (1985), 87–92.
4. Papadakis, M., Sharma, S., Sheppard, M., Panoulas, V., and Behr, E. The magnitude of sudden cardiac death in the young: a death certificate-based review in england and wales. *Europace*, 11 (2009), 1353–1358.
5. Pitt-Francis, J., and et al., P. P. Chaste: A test-driven approach to software development for biological modelling. *Comput. Phys. Commun.*, 180 (2009), 2452–2471.
6. Tice, B., Rodriguez, B., Eason, J., and Trayanova, N. Mechanistic investigation into the arrhythmogenic role of transmural heterogeneities in regional ischemia phase 1a. *Europace*, 9 (2007), 47–58.

Inference of Electrophysiological Conduction Parameters Based on Endocardial Point Measurements

Mikael Wallman[1,2], Nic Smith[3], Blanca Rodriguez[1]

[1]Oxford University Computing Laboratory, Oxford, United Kingdom
[2]Fraunhofer-Chalmers Centre, Chalmers Science Park, Gothenburg, Sweden
[3]Division of Imaging Sciences, King's College London, United Kingdom

Heart failure is a disabling and potentially deadly condition. In the western world, the incidence of heart failure lies between 6% and 10% for persons over the age of 65. The cost for the health care related to heart failure in Great Britain alone was estimated to 905 million for the year 2000 [5, 4].

Over the recent decades, computer modelling and simulations of the heart have emerged as a complement to experimental investigations for understanding the complex mechanisms underlying cardiac function. Simulations allow virtual experiments that can aid clinicians in optimising clinical interventions such as pacemaker implantations [3]. By personalizing the models, these benefits can be further increased.

In order to construct a patient specific model, parameters governing the cardiac electrical conduction need to be estimated. The clinical electrophysiological data used for this estimation is often sparse, typically consisting of point measurements of the electrical potential on the cardiac surface, along with a standard electrocardiogram. This presents two problems: available data must be utilized optimally in order to gain as much information as possible, and a measure of the associated uncertainty must be obtained, to minimize the risk of potentially high-cost errors when simulations are applied in a clinical setting.

In this work, these problems have been simultaneously addressed using Bayesian inference [2], coupled with a very fast method for approximating cardiac electrical activation times. The approximation method relies on modelling the cardiac tissue as a connected graph, with costs of traversing the edges of the graph assigned based on the conduction parameters that are being estimated. Activation times are approximated by finding paths through the graph, from the point of initial activation to the points of measurement, using the A* algorithm [1]. The approximated activation times are used in the Bayesian inference procedure to relate electrophysiological measurements to model parameters. The method has been validated against realistic simulated data sets, and shows high degree of correlation between parameters used for the simulation and those inferred using our method.

References

1. Hart, P., Nilsson, N. J., and Raphael, B. A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on 4*, 2 (1968), 100–107.
2. Jaynes, E. T. *Probability Theory: The Logic of Science*, 1st ed. Cambridge University Press, 2003.
3. Kerckhoffs, R. C., McCulloch, A. D., Omens, J. H., and Mulligan, L. J. Effects of biventricular pacing and scar size in a computational model of the failing heart with left bundle branch block. *Medical Image Analysis 13* (2009), 362–369.
4. McMurray, J. J. V., and Pfeffer, M. A. Heart failure. *Lancet 365* (2005), 1877–1889.
5. Stewart, S., Jenkins, A., Buchan, S., McGuire, A., Capewell, S., and McMurray, J. J. J. V. The current cost of heart failure to the national health service in the uk. *European Journal of Heart Failure 4* (2002), 361–371.

Impact of tissue microstructure on a model of cardiac electromechanics based on MRI data

Valentina Carapella
Computational Biology Group

Cardiac motion is a highly integrated process of vital importance as it sustains the primary function of the heart, that is pumping blood. For this reason cardiac motion abnormalities are often associated with severe pathologies. Clinical non-invasive techniques can assess this fundamental connection between motion aberrant behaviour and pathology only to a certain extent. Computational models of heart function would thus be of great help in linking local to global motion abnormalities and to pathology [2]. Such models are based on the description of both cardiac electrophysiological and mechanical behaviour and their coupling, but, while the electrophysiology of the heart is already well characterised from single cell to whole organ level, the mechanical behaviour is still far from being successfully reproduced. No current model of the heart is able yet to realistically simulate motion patterns in the healthy or diseased heart and therefore prediction of clinically relevant parameters such as ejection fraction, stroke volume, wall thickening and wall motion are not fully reliable.

It is well known that cardiac tissue microstructure, that is cardiac cells organisation into fibres and sheets, has an important role in cardiac motion. Fibre orientation provides the direction along which contraction takes place. This type of information is usually embedded into the model by using prescribed, that is mathematically defined, orientations, although in recent years the use of orientations extracted from data of *ex-vivo* diffusion-tensor magnetic resonance imaging (DT-MRI), also called realistic fibre orientations, has become more frequent. Tissue organisation into sheets has also been long experimentally proven and it is considered to greatly contribute to the even distribution of stress and strain through the tissue. Nonetheless, from the modelling point of view, it is not as well characterised [1]. Sheet orientation can be as well prescribed or obtained from histology or from DT-MRI. Both methods have strong limitations and particularly there is not a full understanding of the relation between DT-MRI values and laminar structure.

The hypothesis driving my DPhil is that current models fail to reproduce realistic features of cardiac mechanics because in most cases they do not embed sufficient information about local tissue organisation into fibres and sheets. Tissue structure presents a high level of heterogeneity, it varies in fact regionally, transmurally from endo- to epicardium, temporally over a cardiac cycle and between subjects [3]. All these aspects of tissue structure deeply affect the anisotropy of cardiac tissue, influencing the local patterns of electrical excitation, mechanical deformation and stress generation. Therefore a more realistic representation of tissue structure within an electromechanical model of the heart, with fibre and sheet orientation extracted from data rather than mathematically defined, together with a more careful definition of tissue material properties, would better take into account the high heterogeneity of tissue structure, thus improving the predictive power of the model.

The aim of my DPhil is to investigate how different settings of tissue structural arrangement affect the motion prediction of an electromechanical model applied to a rat left ventricular geometry obtained from magnetic resonance imaging. My research relies on the integration of cardiac imaging data and mathematical modelling in the belief that realistic models of cardiac function need to reach the best compromise between the level of modelling detail and the amount and quality of information that can be actually obtained from data and used either to instruct or validate such models [4].
My research aims at assessing the impact on electromechanics of the sole fibres, and the

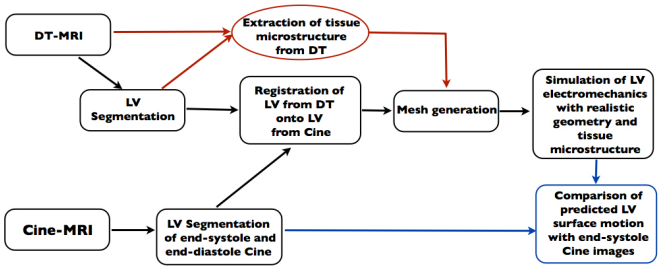


Fig. 1. Pipeline for model construction, simulation and comparison.The middle sub-diagram (black arrows), shows the geometry extraction from images, the upper chart shows the tissue microstructure extraction from DT-MRI (red), the lower chart shows the comparison step of simulated vs realistic motion patterns (blue).

combined effect of fibre and sheet information. In both cases, the different response due to prescribed or realistic fibres and sheets will be investigated. Finally, the direct comparison will be carried out between the simulated motion patterns and those extracted from *in-vivo* Cine-MRI scans. This imaging technique provides with a series of images at progressive points of the cardiac cycle which can then be used to form a cine loop. The key point is that the DT-MRI scans, from which the geometry and the microstructure used to instruct the model are extracted, and Cine-MRI scans, from which the motion patterns are estimated, have been performed on the same subject.

Figure 1 summarises the pipeline that goes from image processing of DT- and Cine-MRI data, necessary to obtain the input for the computational model, to the electromechanical simulations and finally to the comparison between predicted and data-extracted motion patterns. All the image processing part has already been performed during the first year of my DPhil, while the simulation and comparison phases will be carried out in the successive two years. To the best of my knowledge this is the first attempt to perform a comprehensive study of the importance of tissue microstructure in an electromechanical model and to define a framework that combines electromechanical modelling with the use of imaging data from the same subject to first instruct the model and then validate it.

References

1. Gilbert, S., et al. Regional localisation of left ventricular sheet structure: integration with current models of cardiac fibre, sheet and band structure. *European Journal of Cardio-Thoracic Surgery 32* (2007).
2. Kerckhoffs, R., et al. Computational modeling for bedside application. *Heart failure clinics 4* (2008).
3. LeGrice, I., et al. Laminar structure of the heart: ventricular myocyte arrangement and connective tissue architecture in the dog. *American Journal of Physiology - Heart and Circulatory Physiology 269* (1995).
4. Vadakkumpadan, F., et al. Image-based models of cardiac structure in health and disease. *Wiley interdisciplinary reviews. Systems biology and medicine 2* (2010).

A framework for array analysis using separation logic
moving beyond verification of scalar arrays

Abhishek Thakur, Hongseok Yang

Department of Computer Science, University of Oxford

Introduction: Shape analysis is a technique used to prove correctness of programs that involve spatial structures like linked lists. One simple example would be the possibility of accessing a null pointer, or even a dangling pointer, something most of us have done while programming. Shape analysis has come a long way in the last ten years, with seminal work on Separation Logic [1] (which allows us to reason about spatial structures) leading the way. There are tools today [2, 3, 5], based on Separation Logic, that can reason in an automated fashion, not only about linked lists, but more complex spatial structures.

Our work: This work seeks to extend this approach to arrays. Of particular interest to us is reasoning about pointers stored in various fields inside arrays. While reasoning about data stored in array cells has been handled well by predicate abstraction techniques, a complex intertwining of arrays and pointers has not. It is here that the main contribution of this work lies. A very nice example to think of, is a hash table, in which arrays contain pointers to lists, one for each cell, used to store multiple elements, in order to resolve collisions as a result of the hash function. To summarise our work:

1. We describe a first shape analysis framework for arrays and data structures involving arrays.
2. We extend current Separation Logic based analyses [2, 3, 5] to handle contiguous memory cells being accessed using an index.
3. We go beyond verification of arrays with scalar content by providing a framework to reason about arrays containing fields with pointers to other data structures (e.g. lists).
4. We illustrate the success of our ideas by verifying the file *hash.c* from the linux kernel. This is the first full shape analysis of this file, to the best of our knowledge.

Array points-to: We define a basic array points-to predicate to be of the form $a^m[x] \mapsto \overline{f}{:}\overline{E}$, which represents an array a with m as an upper bound on the length, and with ownership of index x, containing information in fields $\overline{f}$ given by respective expressions $\overline{E}$. An example would be $a^{12}[4] \mapsto \{num{:}4, sq{:}16, div{:}p\}$, where p could be a pointer to a list of divisors of 4. It is important to note here that m need not be the known length of the array, but the best known bound on the length at that point in the analysis.

The $\star$ operator: Now that we know how to speak about one cell, we would like to speak about two. We do this using the $\star$ operator of Separation Logic (called the separating conjunction). So

$$a^{m_i}[i] \mapsto \overline{f}{:}\overline{E_i} \;\star\; a^{m_j}[j] \mapsto \overline{f}{:}\overline{E_j} \quad \ldots(1)$$

speaks about two cells of the array, indexed i and j, with the constraint that $i \neq j$. But now, we would want to speak about a set of contiguous array cells using a simple formula. The $\star$ operator is used to create array segments of the form $a^m[x, y)$, which represents array a^m with ownership of cells from x to y, not including y. Logically, this is defined as an iterated separating conjunction

$$a^m[x, y) \;\triangleq\; \circledast_{i=x}^{y\text{-}1} \left(a^m[i] \mapsto \overline{f}_i{:}\overline{E}_i\right) \quad \ldots(2)$$

where one can think of $\circledast$ as the $\star$ operator used y-x-1 times. This construction is the basis of our analysis, and is coupled with a join (abstraction) operation for these segments, which derives from the rewrite rule

$$a^{m_1}[x,y) \star a^{m_2}[y,z) \quad \rightsquigarrow \quad a^{m_1 \sqcup m_2}[x,z) \quad \ldots(3)$$

Materialisation: Next, we would like to extract a cell from a segment to reason about it. This is called materialization (some call it focus). The two rules for materialization are extracted from the following equality in the metatheory ($w', \overline{z}'$ are the set of auxiliary variables introduced):

$$a^m[x,y) \triangleq x{=}y \wedge \texttt{emp} \bigvee \exists w', \overline{z}'.\big(a^m[x,w') \star a^m[w'] \mapsto \overline{f}{:}\overline{z}' \star a^m[w'{+}1,y)\big) \quad \ldots(4)$$

This says that either the array segment is of length zero (when $x{=}y$), or, it has at least one cell (say w') with two possibly empty segments around it. Unlike lists, this array segment can be unrolled to expose one index in the middle, which is what we would like. However, we would like the option of keeping the unroll from the left or the right, which is useful when traversing an array. We use two separate abstraction techniques to present both options to the analysis.

Beyond the scalar: But look at $\overline{f}{:}\overline{z}'$ in the focus rule above. If some $z' \in \overline{z}'$ points to a list, we should include ownership of that list inside our array segment. So we should say something like

$$a^m[x,y) \triangleq x{=}y{\wedge}\texttt{emp} \bigvee \exists w', z'.\big(a^m[x,w') \star a^m[w'] \mapsto f{:}z' \star ls(z', \texttt{nil}) \star a^m[w'{+}1,y)\big) \ldots(5)$$

where $ls(z', \texttt{nil})$ is a null-terminated list segment from z', similar to our array predicate [1]. We would like a more general way of speaking about such linked structures. A cell should not just be one points-to predicate, but a formula in itself, speaking about its pointers and their ownership. For this, we introduce parametrized array segments of the form $a^m\ \Phi\ [x,y)$, which denotes the segment from x to y, with each cell (and the linked data structure) described (and owned) inside Φ. The rules for materialisation are then derived from the following equality in the metatheory:

$$a^m\ \Phi\ [x,y) \triangleq x{=}y \wedge \texttt{emp} \bigvee \exists w', \overline{z}'.\big(a^m[x,w') \star \Phi[a,m,w',\overline{z}'] \star a^m[w'{+}1,y]\big) \quad \ldots(6)$$

Note the similarity between (4) and (6). This allows us to represent an array where each cell links to a singly-linked list. In turn, Λ can define any data structure linked from each of these lists (e.g. doubly linked lists, see [3] for more on Λ), thus allowing us to materialise, reason about, and put back together many data structures involving arrays. Our abstraction [6], based on join rule (3), 'gobbles up' the entire structure back into the compact cell representation, then cells back into segments, ensuring termination.

Conclusion: There are many files in the linux kernel which use arrays along with some spatial data structures, which can be subject to shape analysis in their entirety using this framework. We refer the reader to a working report [6] for more - including technical details, examples, and a brief discussion on related work.

References

1. P. O'Hearn, J. Reynolds, H. Yang: Local reasoning about programs that alter data structures. In *CSL '01*.
2. D. Distefano, P. O'Hearn, H. Yang: A local shape analysis using separation logic. In *TACAS '06*
3. J. Berdine, C Calcagno, B. Cook, D. Distefano, P. O'Hearn, T. Wies and H. Yang: Shape analysis for composite data structures. In *CAV '07*.
4. C. Calcagno, D. Distefano, P. O'Hearn, H. Yang: Footprint analysis: A shape analysis that discovers preconditions. In *SAS '07*.
5. C. Calcagno, D. Distefano, P. O'Hearn, H. Yang: Compositional shape analysis by means of bi-abdiction. In *POPL '09*
6. A. Thakur, H. Yang: A framework for shape analysis for arrays using separation logic: www.qmul.ac.uk/∼abhi/arrays/arrays.pdf

Analysis Framework For Enhanced Mobile Computing Using Cloud Resources

Andrew J. Paverd, Michael R. Inggs and Simon L. Winberg

Department of Electrical Engineering, University of Cape Town
andrew.paverd@cs.ox.ac.uk, michael.inggs@uct.ac.za, simon.winberg@uct.ac.za

Mobile computing devices such as smartphones and tablet PCs have become ubiquitous. Recently there has also been a significant increase in the use of mobile apps. Through the functionality provided by these apps, mobile devices can be considered to be computing tools. The prevalence of both devices and apps indicates a trend towards mobile computing.

However, in terms of computational capacity, a mobile device will always be resource-poor compared to a static computing element due to the fundamental challenges of mobility as identified by Satyanarayanan [3]. In order to continue the trend towards mobile computing, it is necessary to augment the computational capacity of the mobile device. It has been proposed and demonstrated that this can be achieved through a distributed computing relationship between a mobile device and a static computing element via a network such as the internet [1, 2]. Due to the recent expansion of the cloud computing paradigm, it is now feasible to provide the supplementary computational capacity from the cloud. This emerging area of technology is often referred to as mobile cloud computing.

The context of a mobile cloud computing system is shown in Figure 1. In this figure, the mobile device constitutes the primary interface between the user and the system. The limited computational capacity of this device is augmented by dynamically partitioning and distributing computation to the associated cloud resource which offers significantly greater computational capacity. Other internet end-points such as webservices are both sources and sinks of data but are considered to be external to the core system.

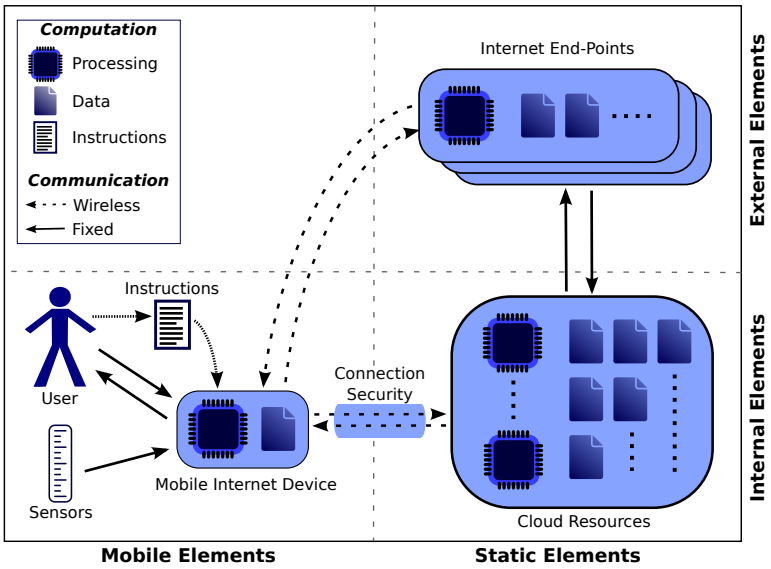


Fig. 1. Overall context showing internal and external elements in a mobile cloud computing system.

The concept of mobile cloud computing has been discussed in recent scientific and technical literature. However, when analysing implementation approaches or applications of this technology, authors often use disparate sets of analysis criteria. This situation makes it difficult to compare different approaches and in some cases the omission of a particular analysis criterion leads to unbalanced conclusions about the overall system.

In order to improve this situation, a theoretical analysis framework is proposed specifically for the analysis of systems based on mobile cloud computing. This framework is a structured consolidation of the salient considerations identified in recent literature. Based on functional similarities, these considerations are grouped into seven major aspects as listed below. A number of interdependencies between aspects are also identified in the framework.

Computational requirements are constituted by the quantity, complexity and type of computational operations performed as well as the volume of data which is processed and/or stored.
Communication requirements include the quantity of data to be transferred over the network as well as performance metrics such as bandwidth and latency.
Mobile network impact analyses the impact on the system given that communication will take place via real-world mobile wireless communication networks.
Energy considerations focus on the amount of energy consumed by the mobile device and by the overall system. Possible improvements in overall energy efficiency are also investigated.
Information security deals with issues of privacy and security arising from the transfer of data over a non-private network between the mobile device and the cloud.
System availability investigates the risk of system failure and how the consequences of such failures curtail the possible applications and thus the overall utility of the system.
Application usability considers any application limitations caused by the inherently constrained physical User Interface (UI) capabilities of mobile devices.

This analysis framework has been used in the investigation of various mobile application domains which will benefit from the mobile cloud computing paradigm. These application domains include mobile scientific computing, mHealth, mobile tools for education, advanced human computer interfaces, multimedia and mobile gaming. The systematic use of this framework in the analysis of these different domains has resulted in the identification of various domain-specific considerations as well as opportunities. It is also envisioned that this analysis framework will form the foundations of future research into the development and application of mobile cloud computing technology.

References

1. Chun, B.-G., Ihm, S., Maniatis, P., Naik, M., and Patti, A. CloneCloud: Elastic Execution between Mobile Device and Cloud. In *Proceedings of the sixth conference on Computer systems - EuroSys '11* (New York, New York, USA, 2011), EuroSys '11, ACM Press, p. 301.
2. Kumar, K., and Lu, Y.-H. Cloud Computing for Mobile Users: Can Offloading Computation Save Energy? *Computer 43*, 4 (Apr. 2010), 51–56.
3. Satyanarayanan, M. Fundamental challenges in mobile computing. In *Proceedings of the fifteenth annual ACM symposium on Principles of distributed computing - PODC '96* (New York, USA, May 1996), ACM Press, pp. 1–7.

A CSP Model of a Process-Oriented File System

James Whitehead II and Thomas Gibson-Robinson

University of Oxford Department of Computer Science
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

Introduction MINIX3 [4] is a Unix-like operating system built on a microkernel architecture. The first two versions of MINIX were designed as teaching tools, but this latest version adds the goal of being a serious option on resource limited and embedded devices. However, in many cases the operating system sacrifices advanced features in favour of clarity of design and simplicity of implementation. For example, the file system only processes a single request at a time, even if servicing the request requires disk I/O. Other requests must wait for it to complete, even if they can be satisfied without I/O. This severely restricts the throughput of the system where there is available multiprogramming. This can cause unresponsive behaviour in the presence of an unreliable device or badly behaved server process.

We have constructed an alternative implementation of the MINIX file system using the Go programming language. This implementations extends the microkernel architectual style into the actual implementation of the server, using process-oriented techniques. This design allows us to introduce concurrency to the file system in a disciplined way without the need for complex layered locking schemes. We have created a CSP [2][3] model of this concrete implementation that illustrates the shortcomings of a sequential file system implementation, and proves the correctness of the patterns of communication in our file system architecture in the presence of multiple concurrent users.

Constructing a Model The MINIX file system format on disk consists of a read-only block containing information about the parameters of the file system, such as the block size, called the *super block*. Files and directories are represented by a data structure called an *inode* which includes information about its size, permissions, etc. The contents of files and directories are stored in uniformly-sized data blocks. Additionally, the file system contains an allocation table that keeps track of which inodes and data blocks have been allocated.

Our model consists of several processes that map to the corresponding components in the file system architecture. The events communicated by these processes correspond to request/response pairs in the implementation. For example, to request a block from the block cache, another process communicates the $call.GetB$ event and then waits for a $ret.GetB$ event in response. We have taken a data-agnostic approach throughout this model: two requests for blocks in the system are indistinguishable. In this way, we model a superset of the actual behaviour of the file system while limiting the state space of the model.

The process network of our model is shown in Figure 1. The *BlockCache* process interfaces directly with the raw disk devices, caching results for commonly requested blocks. The *InodeCache* serves the same purpose for the inodes on the file system, ensuring that access to the open files and directories can be performed quickly. The *Bitmap* is used to keep track of which inodes and blocks are in use on the file system. *OpenFile* and *OpenDir* represent open files and directories on the system, providing ways of interacting with those files. Finally, a *Process* encapsulates a user processes interface with the file system, including the details of how a system call is performed.

The model is configurable in the number of cache slots available for the block and inode caches. Each is implemented as the parallel composition of cache slots, and it is necessary to

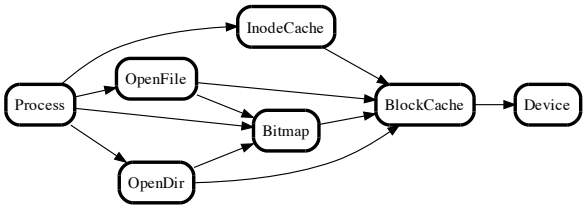


Fig. 1. Process network for MINIX file system

bound this value in order to limit the state space of the model. A cache size of 1 approximates a sequential file system, where only one block and one inode may be accessed at a given time.

The FDR2 [1] model checker provides us with the means to check various properties of the model. Specifically, we have used it to verify that the file system cannot reach a state where it can make no further progress (i.e. it cannot *deadlock*) and, further, that it does not do an infinite amount of internal work to service a particular request (i.e. it does not *livelock*). Additionally, we show that the system call requests and responses can be arbitrarily interleaved, indicating that the concurrent threads are fully independent.

We have also used the model to show that the concurrency policy for access to files and directories is correct. Specifically, a file or directory allows for multiple concurrent readers, but requires exclusive access for a write call. This is perhaps an over-protective restriction, but allows for reasonable concurrency in our data-agnostic model.

Additionally, we have modelled a badly behaved server process that makes a request of the file system but does not accept (or refuses) the response. With a single cache slot, representing the sequential file system implementation, the system deadlocks. Our alternative implementation that allows for concurrent system calls does not deadlock when presented with the same refinement check for varying levels of cache size and user processes.

Conclusions We have presented a process-oriented architecture for a Unix-like file system. This architecture does not require any explicit locking, allowing for the introduction of concurrency in a disciplined manner. The encapsulation and interfaces provided by the architecture make it explicitly clear how components of the file system interact.

To show the correctness of this architecture, we have built a model of the system using the CSP process algebra. We have utilized the FDR2 model checker to prove various properties of the file system such as deadlock and livelock freedom, and shown that the architecture allows file system requests to be handled concurrently. Additionally, we have shown that a faulty device or badly behaved server process does not cause the file system as a whole to block; only the corresponding user process is affected.

References

1. Formal Systems (Europe) Ltd. *Failures-Divergence Refinement—FDR 2 User Manual*, 1997.
2. Hoare, C. A. R. *Communicating Sequential Processes*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1985.
3. Roscoe, A. W. *The Theory and Practice of Concurrency*. Prentice Hall, 1997.
4. Tanenbaum, A. S., and Woodhull, A. S. *Operating systems: design and implementation*. Pearson Prentice Hall, 2006.

Multiphase Modelling of Tissue Engineering

Using mathematical modelling and numerical simulation to improve the quality of engineered bone

K. H. L. Leonard

Department of Computer Science
Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

Human life expectancy is increasing. Failure of organs and tissues as a consequence of age, disease and damage due to injury is increasing the demand for donors. This increasing demand is not being met by an increase in the number of donated organs and there is a chronic shortage; currently around 1000 people per year die in the UK waiting for donor organs[1]. Donor tissues have inherent problems. Engineered tissue has the potential to provide an alternative source of tissue to replace those that have failed due to damage, disease or old age, enabling both length and quality of life to be increased without some of the above mentioned complications that can occur with donor tissues. In addition, engineered tissues with appropriate physiological properties for drug trials can be used as an alternative to animal testing [3].

Tissue Engineering (TE) can be defined as "the application of scientific principles to the design, construction, modification and growth of living tissues using biomaterials, cells and factors, alone or in combination" [8]. This research focuses on one approach to *in vitro* TE, which involves seeding cells onto a scaffold of a suitable material, and placing this within a controlled environment to stimulate cellular growth, the deposition of ExtraCellular Matrix (ECM), and ultimately the development of the tissue required. The term *construct* is used to describe the combination of scaffold, cells and attendant ECM, to distinguish it from the culture medium in which it lies. A bioreactor is normally used for controlling the conditions during the development process; providing the correct nutrient environment via the delivery of a nutrient-rich culture medium and enabling conditions to be set precisely and optimally for cellular proliferation and migration and for ECM synthesis [3], allowing tissue development at an accelerated rate until the tissue is of sufficient functional quality for implantation. For an engineered tissue implant to be functional in the long-term it needs to; withstand the mechanical forces it is subjected to within the body; remodel and thereby integrate into the surrounding host tissue; and respond to biochemical and biomechanical cues. There still exist many challenges that need to be overcome in order to develop engineered tissues that can be clinically useful. In order to improve experimental design, and thereby the quality of the tissue-constructs, the underlying biological processes involved need to be better understood.

Currently engineered bone tissue is not of sufficient quality to be used in widespread clinical practice. Developing an implant that will be robust enough to withstand the forces it would be subjected to within the body is a complex, and as yet, not fully understood process. Bone is mechanosensitive; it undergoes remodelling and restructuring in response to its environment, but the exact cellular response to various mechanical stimuli is not fully understood. To engineer functional implants, the construct needs to be subjected to appropriate mechanical conditioning so that the cells differentiate down the osteogeneic cellular pathway [4].

[1] According to www.organdonation.nhs.uk

There are a large number of factors to consider when engineering a tissue-scaffold construct within a bioreactor external to the body. To create a functional piece of tissue a number of complex and interacting processes that occur within the body during development need to be reproduced. Such biological systems show complexity on a wide variety of temporal and spatial scales. Mathematical modelling and numerical simulation is an invaluable tool for studying such complex biological systems [2, 9]. It is the conjunction of experimental, mathematical and computational research that will further understanding, and help progress in this interdisciplinary field.

This research is a collaborative effort, in conjunction with the cell engineering group run by Professor Alicia El Haj at the Institute for Science & Technology in Medicine, Keele University. We will consider the effect of hydrodynamic pressure on the development and regulation of bone, in a bioreactor designed specifically for this purpose. This area has not received much attention, and as far as we know, there have been no attempts to model such a system to date. We modify a three phase viscous flow model presented in the series of papers by [5, 6] and [7], to describe the growth of a tissue-scaffold construct within a hydrodynamic bioreactor. The resulting model, a system of Partial Differential Equations, can be split into two; a constrained elliptic velocity-pressure system which is similar to Stokes flow, and a volume-fraction system, which is hyperbolic. We use the Deal.II package [1], which has been specifically designed to solve differential equations using Finite Element Methods, to obtain a numerical solution to the mathematical model.

We present numerical results in two spatial dimensions, for pressure-dependent growth within our collaborator's hydrodynamic bioreactor, to demonstrate how the tissue-construct develops over time. Future work will involve extending the current multiphase viscous flow model to include more realistic modelling terms and geometries. A first step will be to model the ECM as a poroelastic material, so as to allow deformation, and will enable better understanding of the forces experienced by tissue constituents. This will require the use of more advanced numerical techniques. In addition, experimental collaboration will enable experimental validation of the model, which has previously been lacking.

References

1. Bangerth, W., and Kanschat, G. `deal.II` *Differential Equations Analysis Library, Technical Reference*.
2. Díaz-Zuccarini, V., and Lawford, P. V. An in-silico future for the engineering of functional tissues and organs. *Organogenesis 6*, 4 (2010), 245–251.
3. Martin, I. The role of bioreactors in tissue engineering. *Trends in Biotechnology 22*, 2 (2004), 80–86.
4. Mullender, M., El Haj, A. J., Yang, Y., van Duin, M. A., Burger, E. H., and Klein-Nulend, J. Mechanotransduction of bone cells in vitro: mechanobiology of bone tissue. *Medical & Biological Engineering & Computing 42*, 1 (2004), 14–21.
5. O'Dea, R. D., Waters, S. L., and Byrne, H. M. A two-fluid model for tissue growth within a dynamic flow environment. *European Journal of Applied Mathematics 19*, 06 (2008), 607–634.
6. O'Dea, R. D., Waters, S. L., and Byrne, H. M. A multiphase model for tissue construct growth in a perfusion bioreactor. *Mathematical Medicine and Biology 27*, 2 (2010), 95–127.
7. Osborne, J. M., O'Dea, R. D., Whiteley, J. P., Byrne, H. M., and Waters, S. L. The influence of bioreactor geometry and the mechanical environment on engineered tissues. *Journal of Biomechanical Engineering 132*, 5 (2010), 051006+.
8. Rose, F. A. J., and Oreffo, R. O. C. Bone Tissue Engineering: Hope vs Hype. *Biochemical and Biophysical Research Communications 292*, 1 (2002), 1–7.
9. Sengers, B. G., Taylor, M., Please, C. P., and Oreffo, R. O. C. Computational modelling of cell spreading and tissue regeneration in porous scaffolds. *Biomaterials 28*, 10 (2007), 1926–1940.

Visualization of experimental designs and workflows in biological experiments

E. Maguire[1,2], S.A. Sansone[2], J. Davies[1], P. Rocca-Serra[2], M. Chen[2]

[1]Oxford University Dept. of Computer Science, Oxford, OX1 3QD, UK
[2]Oxford e-Research Centre, Oxford, OX1 3QG, UK

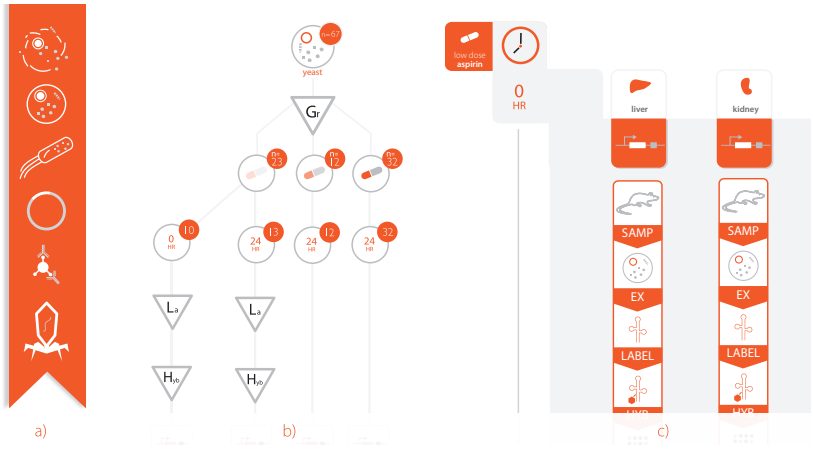


Fig. 1. We present: a) example glyphs for biological experiments; b) a directed acyclic graph for visualizing experimental design and work flows; and c) an alternative tree-based layout scheme with implicit edges.

Abstract. Evidence based science mandates full disclosure of scientific data and ancillary experimental metadata payload. Only this data provenance can warrant the possibility of reviewing findings and claims. Reports should therefore contain information about biological materials, their treatments and the molecular dimensions being surveyed and the methods to perform those measurements. The ISA-Tab format [4] provides a generic grammar to organize and structure experimental information and linking it to matrices of results. While spreadsheet based representation is popular as a natural interface, devising new means to graph experiments and sample processing workflows can allow for more immediate understanding of overall experiment design. In particular by mimicking the practice of drawing experimental groups and applying the technique of dimensionality reduction, we are working at devising more engaging, compact and informative renderings than techniques currently available. We posit that the same techniques could be used to plan experiments and drive data acquisition (therefore providing game changing means to data management).

1 Experimental Design Overviews

Hypothesis evaluation and testing are at the core of scientific practice. In this world, the notion of design of experiment is key. Declaring variables, sample sizes and procedure should be the norm. Yet, extracting this information from public resources is seldom easy, let alone the task of presenting it in a meaningful way. While making

a judgment call over experimental design choices should remain the realm of scientific debate, providing the means to present essential information about experimental design attributes such as nature of replication, variables and their level is a reasonable endeavor. We have developed a simple view to show distribution of biological replicates across treatment conditions in our ISAcreator tool [4]. Limitations of this current method are that there is no timeline support or concept of treatment intensity. We are expanding on this initial work to create a reader-driven, narrative visualization [2] capable of telling the underlying story behind an experiment. Figures 1b and c show our progress in this work.

2 Experimental Workflows

Workflows are made up of protocol application events occurring over the course of the experiment and enacted on physical objects (biological specimen, populations) or digital objects (data files, images). Recording them properly vouches for reproducibility of experiments and helps in downstream analysis. Some effort has gone into visualizing experimental workflows in the past with the general mechanism for doing so being the directed acyclic graph. Most of these visualizations do not scale well and even simple workflows are difficult to represent and compare [1]. Taking a different direction to current state of the art techniques, we wish to provide a visualization that: 1) better utilizes screen space whilst increasing the information on display through the use of a new visual language for biological experiments; 2) makes comparisons between workflows easier; and 3) builds on previous work in timeline visualization, particularly genealogical timelines which have similar paradigms to sample preparation (e.g. merge and split of samples can be mapped to marriage, divorce in geneology) [3].

3 A visual language for biological experiments

Following work in devising designs for the aforementioned visualisations, we are progressing in the creation of a glyph based language for biological experiments. The aim of such a language is to encapsulate the key materials and processes used in such experiments whilst reducing the visual space required to render this information. Through the creation of such glyphs rather than simply text as in [1], it is possible to make comparisons more readily. This work has yielded encouraging results (see figure 1a) and work is ongoing to validate the designs qualitively and quantitively using formal mathematical approaches and user testing.

4 Acknowledgements

Funding: BBSRC, NERC (BB/I000917/1, BB/I025840/1).

References

1. Experiment E-MTAB-346. http://www.ebi.ac.uk/arrayexpress/files/E-MTAB-346/E-MTAB-346.biosamples.png, Oct. 2011.
2. E.Segal, and Herr, J. Narrative visualization: Telling stories with data. *IEEE Trans. Visualization and Comp. Graphics (Proc. InfoVis)* (2010).
3. K Nam Wook, S K Card, J. H. Tracing genealogical data with timenets. *Advanced Visual Interfaces* (2010), 241–248.
4. Rocca-Serra, P., Brandizi, M., Maguire, E., Sklyar, N., Taylor, C., Begley, K., Field, D., Harris, S., Hide, W., Hofmann, O., Neumann, S., Sterk, P., Tong, W., and Sansone, S.-A. ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level. *Bioinformatics 26*, 18 (2010), 2354–2356.

Posters

Philip Gemmell	Effects of Physiological Parameter Variation in a Computational Rabbit Ventricular Cell Model
Eoin Hyde	Determining Fluid-Related Parameters for Poroelastic Perfusion Modelling
Hsiang-Shang Ko	Solving the Dutch National Flag Problem via Datatype Ornamentation
Philip Maybank	A Novel Method For Automatically Identifying Key Processes In Detailed Biophysical Models: Model Reduction Using A Posteriori Error Analysis
Ganesh Narayanaswamy	On the Homotopy Structure of Abstract Interpretation
Margarita Satraki	Biomechanical Modelling of Soft Tissues Using the FEM Method
Tamás Székely Jr.	Higher Order Numerical Simulation of Stochastic Chemical Reaction Systems

Prizes

Best Talk:	Jamie Frost
Second Place (awarded jointly):	Eamonn Maguire & Mikael Wallman
Best Poster:	Hsiang-Shang Ko