

Institution: University of Hertfordshire
Unit of Assessment: Panel B (11): Computer Science and Informatics
Title of case study: Industrial applications of Automatic Differentiation and advanced methods in compilation technology
<p>1. Summary of the impact (indicative maximum 100 words)</p> <p>Graph-theoretic and mathematically rigorous algorithmic methods developed at the University of Hertfordshire have improved the applicability of compiler technology and parallel processing. A compiler developed in the course of a ten-year research programme at the university has been successfully applied to a number of commercial problems by re-purposing the research tool. NAG Ltd has adapted the tool into a commercial product now used operationally by investment banks. Numerous applications of the mathematical methods (such as type-flow graphs used conjointly for correctness and optimisation) have been deployed by industry (including SAP, SCCH, German Waterways Board) working closely with the university.</p>
<p>2. Underpinning research (indicative maximum 500 words)</p> <p>Compilation relies heavily on graph theory as an enabler. Our research uses mathematical approaches to unify seemingly disparate aspects of compilation technology.</p> <p>Numerical simulation of complex real-world phenomena represents major challenges for science and engineering, particularly when it is desired not merely to model a process (such as airflow across the surface of a car or turbine) but to optimize it (reshape the surface to minimize drag). Doing this efficiently for large problems requires numerical derivatives (sensitivities of outputs to inputs) that are accurate (free from rounding and truncation error) and computationally cheap. Automatic Differentiation (AD) techniques transform numerical modelling code mechanically to calculate numerical sensitivities of model values to the same precision as the values themselves. The adjoint mode of AD is uncannily efficient: it calculates a complete output gradient vector at a cost of five evaluations of the model, regardless of the number of input parameters, which may be several million. AD techniques developed significantly from research carried out at the university by Professor Laurence Dixon and his collaborators, the adjoint mode being studied intensively from 1994.</p> <p>The idea to place AD functionality inside the compiler originated at the University of Hertfordshire. The CompAD Fortran compiler developed out of a research collaboration (beginning in 2004) between the university's Centre for Computer Science and Informatics Research (CCSIR), the Centre of Excellence for Computational Engineering Science at RWTH-Aachen, and NAG Ltd, an Oxford company specializing in numerical software and compilers for the financial, business analytics, and engineering fields. The CCSIR team, led by Professor Bruce Christianson, was funded by a series of EPSRC grants, and the bulk of the work building the compiler was carried out by Hertfordshire researchers. The primary objective of the CompAD project is to investigate almost-optimal algorithms for trading off storage against recalculation in extremely large computational graphs. Finding the optimum is NP-hard, so an almost-optimal heuristic is sought.</p> <p>Parallel processing has moved from niche market to prevalent challenge in industrial software development. The university is pioneering programming technologies that address these challenges. Our type-flow analysis enforces software security and, in real-time software, performs accurate flow and timing analysis. Consequently, it can verify and enforce quality for real-time and embedded software. Type-flow analysis also statically identifies compute-intense segments of code more accurately than traditional control and data-flow oriented techniques, by considering the concrete data-types operated on by compute-intense pieces of code. This affects the value of moving the segment to hardware as custom instruction or IP core.</p>

Re-implementing existing software is prohibitively expensive, so we have developed programming systems that allow transparent re-use of existing software components. New components are implemented using our high-productivity, high-performance language SAC (Single Assignment C), a purely functional array programming language designed for computational intensity. S-Net is a declarative coordination language for streaming networks of asynchronous components. SVP (System Virtualization Platform) is a common hardware virtualization layer and execution substrate for higher-level languages, including S-Net and SAC, providing system services and language interfaces for exploiting concurrency on many-core processor chips. SVP can be used with general-purpose multi-cores, but is aimed towards heterogeneous platforms with per-core hardware multithreading.

Underlying all this work is the mathematical analysis and exploitation of computational and flow graphs, for optimisation and correctness.

3. References to the research (indicative maximum of six references)

– University of Hertfordshire authorship is indicated by bold type; the top three publications are indicated by an asterisk.

* **M. Bartholomew-Biggs, S. Brown, B. Christianson and L. Dixon**, Automatic Differentiation of algorithms, *Journal of Computational and Applied Mathematics*, 124 (1–2), 2000, pp.171–190. doi: 10.1016/S0377-0427(00)00422-2

B. Christianson, D. Casanova, R.S. Sharp, **M. Final** and P. Symonds, Application of Automatic Differentiation to race car performance optimisation, in *Automatic Differentiation of Algorithms: From Simulation to Optimization*, Springer, 2002, pp.117–124. ISBN 978-0-387-95305-2

* **C. Grellck and S. Scholz**, SAC: A functional array language for efficient multithreaded execution, *International Journal of Parallel Programming*, 34 (4), 2006, pp. 383–427. doi: 10.1007/s10766-006-0018-x

C. Grellck, S. Scholz, and **A. Shafarenko**, A gentle introduction to S-Net: Typed stream processing and declarative coordination of asynchronous components, *Parallel Processing Letters*, 18 (2), 2008, pp. 221–237. doi: 10.1142/S0129626408003351

U. Naumann, M. Maier, **J. Riehme** and **B. Christianson**, Automatic first- and second-order adjoints for truncated Newton, in *Proceedings of the International Multiconference on Computer Science and Technology (IMCSIT)*, 2007. ISSN 1896-7094, Vol. 2: Computer Aspects of Numerical Algorithms, pp.541–545

F. Rauser, **J. Riehme**, K. Leppkes, P. Korn and U. Naumann, On the use of discrete adjoints in goal error estimation for shallow water equations, *Procedia Computer Science*, 2010, pp.107–115. doi: 10.1016/j.procs.2010.04.013 1(1)

* **J. Riehme**, R. Kopmann and U. Naumann, Uncertainty quantification based on forward sensitivity analysis in SISYPHE, in *Proceedings of the V European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010*, 2010. ISBN 978-989-96778-1-4

Selected Funding

EPSRC Project Funding

2006 (2.25 years). ‘Differentiation-Enabled Fortran 95 Compiler Technology’ (CompAD II). PI: Bruce Christianson. **Award Value** (University of Hertfordshire share): **£317,666**.

2008 (2 years). PI: Bruce Christianson. ‘Differentiation-Enabled Compiler Technology – COMPAD-III’. **Award Value** (Hertfordshire share): **£227,045**.

2012 (1 year). PI: Bruce Christianson. ‘Towards Industrial Strength Automatic Differentiation’ (CompAD Follow-on). **Award Value** (Hertfordshire share): **£101,341**.

EU Project Funding

2005 (3 years). PI: Alex Shafarenko. EU FP-6: Project Aether. **Award Value** (Hertfordshire share): **£456,566**.

2008 (3 years). PI: Sven-Bodo Scholz. EU FP-7: Project Apple-CORE. **Award Value** (Hertfordshire share): **£325,322**

2010 (3 years). PI: Alex Shafarenko. EU FP-7: Project Advance. **Award Value** (Hertfordshire share): **£563,632**

4. Details of the impact (indicative maximum 750 words)

German Waterways Board, NAG Ltd, [text removed for publication]

Despite increasing awareness of the benefits of AD, these techniques are not yet routinely used in modelling and optimisation outside a few niche areas, due to a lack of suitable commercial tools. The CompAD project based at the University of Hertfordshire developed the world's first industrial-strength Fortran compiler with built-in support for AD, based on the NAG Fortran compiler.

Following successful completion of the second phase of the CompAD project in September 2008, the research compiler was re-purposed to meet the requirements of various end-users, as consultancy work carried out by members of the CompAD research project team (including Jan Riehme, then a research fellow at the university, now senior research scientist at Software Tools for Computational Engineering, RWTH-Aachen).

For example, to predict flow and profile of sediment along the river Danube, the German Waterways Board/Bundesanstalt für Wasserbau has been using the AD-enabled NAG Fortran compiler since August 2009, with a view to making its dredging operations more effective – the Waterways Board is responsible for 7,350km of inland waterways and 23,000 square km of shipping lanes. Instead of calculating just a single value for the evolution at one point in time and space, the compiler-based Automatic Differentiation approach allows a probability distribution of the evolution to be automatically derived from the morphodynamic model. Reliability analysis then allows the uncertain input parameters to be ranked according to their level of uncertainty. The new modelling approach shows good agreement with the Monte Carlo approach previously used, but (for first order reliability analysis) takes only 22 minutes instead of 5 hours. The Waterways Board recognizes this advantage:

It cannot be hoped to overcome this [exponential computing time] limitation of the Monte Carlo method only by increasing computer power. Hence non-linear methods based on algorithmic differentiation seem the more promising way. [Sect. 5, Ref. 1]

Meanwhile, NAG Ltd has made the CompAD compiler available on a commercial basis to its customers, including engineering and financial sector organisations. In response to customer requests, NAG took the commercial decision to develop AD versions of a number of mathematical library routines, using an enhanced version of the CompAD compiler. [text removed for publication]

In particular, NAG is already providing AD software and consultancy services based on the CompAD research [text removed for publication] The C++ tools involved were developed in partnership with RWTH-Aachen, building on the research pioneered by the CompAD project. [text removed for publication]

Thales Group, SAP and SSCH

Other compilation technology tools, including SAC, S-Net, and SVP – which were developed at the university by Professor Alex Shafarenko and Dr Sven-Bodo Scholz (formerly Reader at Hertfordshire, now Professor at Heriot-Watt) in conjunction with EU grant academic and industrial partners – have been taken up by industry. Following completion of the Aether project in 2009, Thales group successfully applied our technology to a pre-existing signal processing application

[Sect. 5, Ref. 5]; SAP (the largest software consultancy company in Europe) has been using our tools since 2010 to solve NP-hard optimisation problems; and SCCH benefits from significant speed-ups of their classification algorithms. An expert from SAP headquarters in Karlsruhe, Germany, wrote in November 2012 to say that:

We found that S-Net gives us an easy way of changing the parallel implementation of the ACO algorithm. Using the S-Net technology, we could quickly try different approaches to modeling concurrency simply by changing a few lines in S-Net code, . . . S-Net imposes only very little overhead compared to another standard parallelization approach and it offers the opportunity to efficiently utilize the available number of parallel processor cores.

Further tribute comes from the Head of Knowledge-Based Technologies at SCCH, who has said that, thanks in large part to the EU FP-7 project Advance (2010–12), which supported this aspect of the university's research and knowledge transfer, 'high performance computing has become an established technique at SCCH', and 'we are very thankful to be integrated in your active and motivating research team'.

Impact via export of former project personnel includes Stephan Herhut who, following his PhD (2012), was headhunted by Intel Research in Silicon Valley, to develop proprietary compilation technology.

5. Sources to corroborate the impact (indicative maximum of 10 references)

German Waterways

1. R. Kopmann, U. Merkel and J. Riehme. Using reliability analysis in morphodynamic simulation with TELEMAC-2D/SISYPHE, *Proceedings of the XIXth TELEMAC-MASCARET User Conference*, Oxford, UK, 18–19 October 2012, 119–125; 124. ISBN 978-1-898485-15-5
 <http://www.opentelemac.org/images/clubu/2012/TELEMAC-MASCARET_Proceedings-2012-R1.pdf>

Numerical Algorithms Group (NAG)

2. Information about the company can be confirmed at: <www.nag.co.uk/>
3. An outline of the NAG Fortran compiler, including University of Hertfordshire involvement, is at: <www.nag.co.uk/nagware/research/ad_overview.asp>

[text removed for publication]

Thales Research

5. Penczek, F. et al, Parallel signal processing with S-Net, *Procedia Computer Science*, 1 (1), 2079–2088, 2010. <<http://dx.doi.org/10.1016/j.procs.2010.04.233>>

Corroboration: Beneficiaries of the Research

Four written statements from individuals who can corroborate the impact within German Waterways, NAG Ltd, Thales Group, SAP AG and SCCH, as outlined above, are available. These include the SAP AG and SCCH quotes in section 4. (Contact details supplied separately.)