

2013 International Conference on Computational Science

Introducing: Computational Science

Valerie Maxville¹

iVEC, Perth, Western Australia

Abstract

A deep understanding of computational science takes years to develop, however a basic grasp can be achieved in a short workshop. This can create and nurture the spark of interest in computational approaches to solving science problems. A series of outreach activities has been developed for this purpose at iVEC in Western Australia. The activities include simulation and analogy to give an accessible introduction to computational science and supercomputing concepts. This paper describes the developed lesson plans and reflects on the results of delivering workshops to a range of audiences.

© 2013 The Authors. Published by Elsevier B.V.

Selection and peer review under responsibility of the organizers of the 2013 International Conference on Computational Science

Keywords: computational science, education, outreach

1. Introduction

Many of the key challenges society is facing require innovations in scientific research fields including: climate, water, disease, health, energy and sustainability. These areas are increasingly reliant on computational techniques and it is important that we grow the future researchers and technologists to be part of the necessarily diverse research teams. However, in Australia, as in many parts of the world, we continue to struggle to attract students into science and technology, and, for the last decade, computer science. As Western Australia is a resource-rich state, engineering is very popular. It is the author's view that computational science will be more attractive to young people than the way they imagine science is undertaken. Thus it is important to make them aware of the reality of science research, and the technology that enables it.

The author is the Education Program Leader at iVEC, an organization supplying technology and expertise to support research in Western Australia and supercomputing nationally. iVEC is a joint venture funded by its partners (local universities and CSIRO) and the state and federal governments. Our focus areas are supercomputing, data storage, visualisation and eResearch. The iVEC Education Program facilitates training,

¹ Corresponding author. *E-mail address:* maxville@ivec.org

internships and outreach across the iVEC partners and the wider community. This paper describes some of the outreach activities from iVEC over recent years. For this paper, we begin by considering approaches to outreach, then describe the resources developed. This is followed by examples of workshops that have been delivered and a reflection on the effectiveness of the materials and workshops.

2. Approach

Engaging learners requires accessible activities to include all students, but also have enough depth to hold the attention of the more advanced students. Examples of resources in this style include Interactivate [1] and the Computer Science for Fun Project [2]. The author has attended and learnt from workshops on computational thinking [3] and computational life sciences [4], along with activities from Scitech in Western Australia [5] and similar organisations internationally.

With these initiatives in mind, we set out to develop approachable materials, based on fun, meaningful activities. The topics and activities needed to be suited to a full range of ages and backgrounds to provide maximum value for our effort. As stated by Bob Panoff of Shodor [3], there is greater learning value in providing models for learners to play with, than just explaining a theory or single solution to a problem. So the ability for learners to test ideas and get feedback was intrinsic to the activities developed.

Our resources make use of analogy, simulation and roleplay to involve students. The limitations of the analogies and simulations are pointed out in the workshops to avoid leading students to believe it really is this simple. Once a concept has gelled, we are able to link it to the real world as an extension for them to ponder. For example, an analogy for discussing data locality may be that the desk you are sitting at has a processor and some “local memory/RAM”, the desk across the room is “disk” and the table upstairs in the hallway is “tape”. There is also a focus on group work to encourage peer learning. This includes roleplay where the students may be taking the role of a researcher, or may be taking the role of a piece of computer hardware. We assume face-to-face sessions for the resources, which could also be modified to be delivered as online resources. The materials have been reviewed after each session with modifications and extension occurring continuously.

3. The Resources

One of the aims of the outreach activities that have been developed is to create resources for teachers to use in their classrooms. These have evolved across the workshops described in Section 4 and were made available as a package to teachers in October 2012. The materials are split into three lessons, focusing on supercomputing concepts, data movement and an example of computational science (fluid dynamics).

3.1. Lesson 1: Supercomputing

Each lesson aims to attain a specific learning outcome. With Lesson 1, the expected outcome is: *By the end of this session you [learners] will have an understanding of some concepts and challenges in supercomputing.* There are three activities in the lesson that take approximately one hour to complete as a set.

Activity 1: Lolly Processors - The first activity looks at how processing and data are spread across multiple processors in a grid-based algorithm, and the communication that is required. Four students initially arrange themselves into a 2x2 grid with the other students watching on. The students are assigned a processor number and given two types of lollies (candy): Lolly A that takes a while to chew (e.g. Mentos) and Lolly B that goes quickly (e.g. Tic-Tac or M&M). To begin an iteration, chewing Lolly A represents each processor working through its part of the grid. Once a processor has finished processing its data, it then exchanges halo data to the neighbouring grid cells. The halo data is represented by Lolly B, and will be passed and received in two

directions for each grid cell. Once the halo data is processed, a new iteration begins and the processors can eat another Lolly A. The 2x2 grid and pseudocode for four iterations is shown in Fig. 1.

```

For i=1 to 4
  while lolly A is uneaten
    chew lolly A
  pass lolly B's to neighbor(s)
  while lolly B's are uneaten
    chew lolly B's

```

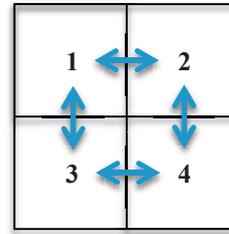


Fig. 1. Pseudocode for four iterations of the Lolly Processors with a 2x2 grid

In the second part of this activity, students form a 4x4 grid (if numbers allow) and work through some iterations. The activity can raise a range of discussions and any mistakes become food for instruction. An extension could be to consider the actual layout of a machine – that they won't be next to all their neighbours, which could lead to discussion of nodes and interconnects. Section 4.1 describes the activity in action.

Activity 2: Parallel Problem Solving – In this activity, students work in groups and are given an allocation of jigsaw puzzle pieces. Initially they can receive a random allocation across the whole puzzle. The time to solve can be recorded. They should then discuss how to improve the solving time by changing how they allocate pieces – data decomposition. Options could include dividing the puzzle into a grid and assigning pieces, or they might allocate by content – sky, trees, water, etc. A set of different puzzles should be used to avoid improved times due to familiarity with the puzzle. After a few tries they can review the effect of the different decomposition approaches on their efficiency. The approaches can be related to geographic or climate models where data is gridded into cubes, but contents may be treated differently, so air, water and land can be segmented to allow uniform processing.

Activity 3: Balloon Workflow – This activity begins with four students preparing to blow up balloons as a pipeline process. There are four sub-tasks: 1) choose balloon; 2) inflate balloon; 3) tie balloon; and, 4) draw a smiley face on the balloon. There is a mixture of trivial and fiddly tasks to create obvious bottlenecks. As the students process the balloons, the rest of the group can observe and talk through the bottlenecks and inefficiencies. Suggestions for improvements can be proposed and implemented (if time/numbers permit). For example: a cylinder of gas might speed up inflation; more people could be added to inflate balloons; or a tool could be used to tie the balloons. These could be considered equivalent to accelerators, multiple-processors and library routines, respectively. The aim is to load-balance across the processors, which may not be able to do the entire workflow independently (some people can't tie balloons). Although the pipelined balloon blowing may not be efficient, most students would agree that the distributed workflow in a fast-food restaurant is very efficient. This activity is included in the second workshop (Section 4.2).

3.2. Lesson 2: Dude, where's my Data?

In Lesson 2, learners will: *Have an understanding of some of the concepts and challenges relating to the movement of data on supercomputers*; and: *Know some strategies and techniques for efficient data movement on supercomputers*. In this lesson, we focus on three broad options for data storage – local memory (RAM), disk and tape. Students may not be aware of tape as storage media, so an explanation may be required – including its energy efficiency and suitability for very large datasets. The discussion could also include/focus on categories for local disk versus global disk, and local memory versus remote memory.

We aim to get across the cost of moving data, the constraints on capacity and how we might manage getting data to the processors. As background, we provide a diagram indicating the time required to shift data between locations and explain that data can only be **processed** when it is in RAM.

Activity 1: So far away... – Select pairs of students to work as a processor and data fetcher. Other students can be observers. Each pair will have a lotto sheet, from which cards (picture or number) are randomly assigned to three storage places – the home desk, a desk on the other side of the room, and a table in another room (up or down stairs is even better). One participant is the processor and sits at the home desk assembling the data. The other member will collect the cards in sequence, one at a time, and bring them back to the home desk. The task should be timed – perhaps with a scoreboard to encourage competition. As they run around (or walk), it quickly becomes clear that distant data has a huge impact on the processing time. The students can then suggest ways to improve efficiency – leading the conversation to having the data pre-fetched to be closer to the processor as it is needed. Additional team members can be the pre-fetchers, allowing the first fetcher to have a rest! Rules on how much data (cards) can be in each storage space, and how many cards can be transferred at once can add to the challenge. Students can discuss the approaches used and the relative timings achieved, and can then discuss what this would mean for processing large amounts of data on a supercomputer. Data streams could also be mentioned and enacted – along with what happens if the processor gets overwhelmed.

Activity 2: Masterchef – This activity uses the analogy of cooking, which is often used to introduce sequencing and parallelism. In this case we are interested in data – the ingredients and products of the recipe. The three data locations (two desks and far away table) now represent the kitchen bench, the pantry and the supermarket (grocery store). Following a supplied recipe, students begin with an empty pantry and bench, all ingredients are at the supermarket – so the students will have some running around to do! Ingredients need to go to the pantry before they can be on the bench for cooking. Sticky notes or cards can be used to represent the ingredients (and portions) and are set up on the far away table. Students should take time to think about their approach and then be set going through the recipe. Individuals and team may take different approaches. After a few runs through this activity can be reflected on and related to staging data for processing. A potential extension or alternative to this activity could have data coming from different directions – modelling a mash-up application creating new data from sources such as web pages and data repositories.

3.3. Lesson 3: Fascinating Fluids

The wet-lab/dry-lab approach to research iterates between real world observations and experiments, and the dry lab of computer models [2]. This lesson involves a roleplay with the students becoming computational scientists modeling a lava lamp. In the workshop, participants are provided with a set of lab notes, a dry lab of computers with the lesson pages pre-loaded and a wet lab of lava lamps. The notes include an overview of computational science and templates for Wet Labs 1-3 and Dry Labs 1-3. After a verbal introduction, the students begin by observing the lava lamps in Wet Lab 1 – taking notes on what would be part of a computational model. They then view and appraise a pre-made model in Dry Lab 1. The model is a simple javascript animation and after another wet-lab/dry-lab iteration we discuss how a computational model (not animation) would be built. Wet Lab 3 focuses on fluid flow using water, dyes and glitter in containers and a glitter lava lamp (heat and fluid flow). In the final lab the students play with a fluid model and discuss what else would be required to model the lamp – two fluids with a changing boundary and taking it into three dimensions being the key points. The workshop is further described in Section 4.3.

4. The Workshops

The lessons and activities described in the Section 3 have been delivered to a range of audiences. The author facilitated all of the workshops in seminar rooms, auditoriums and computer labs. The only restriction on venue was the need for computer access and power for the activities Lesson 3.

4.1. *Workshop 1: National Youth Science Forum*

The audience for this workshop was a large group of Year 12 high school students. The National Youth Science Forum aims to expand the horizons of the students for study and career choices as they enter university. These students are selected from a national competitive process and have a keen interest in science. The iVEC visit was part of a two-week program of workshops and site visits. After being given an overview of iVEC and a demonstration of 3D stereo scientific visualisation, we took the students through the grid activity from Lesson 1. The initial group of four volunteers were tentative. When the task was explained and processor numbers assigned, they arranged themselves to match the 2x2 grid diagram and were given two types of lollies to “process”. Longer lasting lollies represented processing of their own data, while smaller lollies were the messages passed between processors. As the processing went on, the concept of calculating sub-steps of a problem through iterations then exchanging halo data between cells was explained and discussed with the group. The request for sixteen students for the second activity had an enthusiastic response. These students formed a 4x4 grid and quickly saw the increase in complexity and communication that came from a larger grid.

Discussion of the impact of different types of lollies – akin to changing algorithms – had them thinking. Drawing attention to the plight of the internal grid cells who now had four messages to deal drew pity and envy. Dropped lollies helped illustrate the challenges, as did having to wait for others to finish chewing. After a few iterations they appeared to have taken in the ideas presented thus far, confirmed by the audible gasp as they were asked to consider the grid in three dimensions – as a 4x4x4 volume.

Response to this workshop was very positive and showed the value of shifting to a presentation plus activity-based visit, as compared to the usual dry presentations which would be particularly dull for school students.

4.2. *Workshop 2: Educational Computing Association of Western Australia (ECAWA) Conference Workshop*

This site visit was a workshop as part of the ECAWA annual conference – attracting school teachers from primary and high schools. A short overview of iVEC and computational science was given – asking the question “Why would a scientist use computers?”. We then focused on supercomputing, using the lolly and pipeline activities from Lesson 1. The lolly activity proceeded similarly to the NYSF workshop, with the option of a cup provided for those who preferred not to eat the lollies. Mints were also offered as alternatives. For the pipeline activity, four volunteers were called and given a task each. One would select a balloon from the bag, the second would blow up the balloon, the third tied the balloon and the fourth drew a face on the balloon. These would roughly align to data preprocessing, processing, aggregation and post-processing. The volunteers processed a number of balloons while the rest of the group observed and discussed how the workflow was performing in terms of load balancing. With this group, the balloons were blown up in one breath, and the bottleneck was with the tying of the balloons. Suggestions for improvements were requested and additional volunteers joined on the tying task to relieve the bottleneck. The group noted that the balloons ended up all over the floor and that gave the opportunity to discuss data management and clean up.

The teachers were actively engaged in the activities and a healthy discussion was generated. Strong interest was shown in future materials and potential school visits to iVEC. Conversations continue as to how to introduce computational science and supercomputing into the school curriculum. Suggestions from this

workshop included animations of the activities to allow changing numbers of processors, processing time, communication overheads to allow more feeling for the impact of each of these variables.

4.3. Workshop 3: Scitech Double Helix Workshop

Double Helix is a science-focused club run by Scitech – the Western Australian science exploration centre, whose key audience is in the 8-10 year old range. iVEC offered a workshop on Fascinating Fluids to sixteen students as part of their school holiday program. When the students arrived they were asked to complete an online survey to assess their prior knowledge. The session began with an age-appropriate introduction to iVEC and computational science. The concept of the wet lab-dry lab research process was introduced, and they were given a folder including laboratory notes for them to complete during the workshop.



Fig. 2. Lava lamp wet laboratory with dry laboratory (computer lab) in the background

We then moved to the “wet lab” in an adjoining room (Fig. 2), where eight lava lamps were set up on tables. These were set up an hour before the class to heat up and were put out of easy reach with cables securely out of the way. The students were instructed to observe the lamps and record what they felt would be useful information if they wanted to develop a computer model of a lava lamp. A cold lamp was available for them to hold and look at more closely. The basics of the lamp – glass, fluid, wax, light, air space, coil were noted. Once they had some diagrams and notes, we moved to the “dry lab”. The students clicked through to the first models on the workshop website (Fig. 3a) and turned to their lab notes to record their observations.

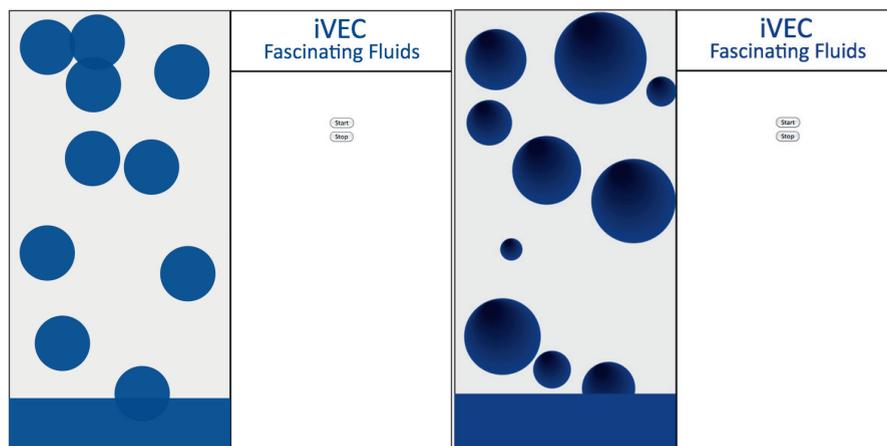


Fig. 3. Animated lava lamps (a) initial model; (b) improved model

The models were simple animations of a lava lamp, written in javascript. The initial model had ten identical semi-transparent circles which were randomly placed and floated up and floated up and down at the same rate with some randomness added for left/right movement. The blobs were constrained to a rectangle. The model could be paused using the start/stop buttons. Students noted: the shape of the model didn't match the lamps; the flatness of the model; uniformity of size and movement; that blobs could be clicked on; and, the whole model could be started and stopped. The interaction with the model was discussed – noting that this can be one of the benefits of a computer model over a real world experiment. With these observations, the group returned to the wet lab to make more observations and notes for improving the computer model.

The group returned to the dry lab and clicked through on to a variety of improved models including Fig. 3b. In their lab notes they wrote observations of changes to the previous model, which included: shading of blobs; variation of size and speed of blobs (smaller move faster); all starting at the bottom and randomly emerging from the goo; and, blobs hanging at the top and bottom of the lamp (not just “bouncing” up and down). A discussion of the limitations of the animated model drew out the visual appeal of the model, but that it did not compute or model the actual properties of the lava lamp. We discussed how we could break up the problem to study different parts of it and build a computational model – perhaps starting with a single fluid and observing and modeling it.

Moving back to the wet lab, rectangular containers had been set up with a small amount water in each. The students split into groups of 2-3 around each container. Students chose food colours (dye) to add to the water and observed how it moved through the fluid, and used chopsticks to influence the mixing (Fig. 4a). More colour was added to make the water quite dark across all groups and then silver glitter was sprinkled onto the water. This made the movement of the fluid more visible when swishing with chopsticks (Fig. 4b). Students wrote some observations before moving to the dry lab. In the dry lab, the students clicked through to a computational fluid model (Fig. 4c), using the model published by Greenwold [6]. The model included the frames per second on the top of the display, which helped to explain that the model calculated changes iteratively through time increments.

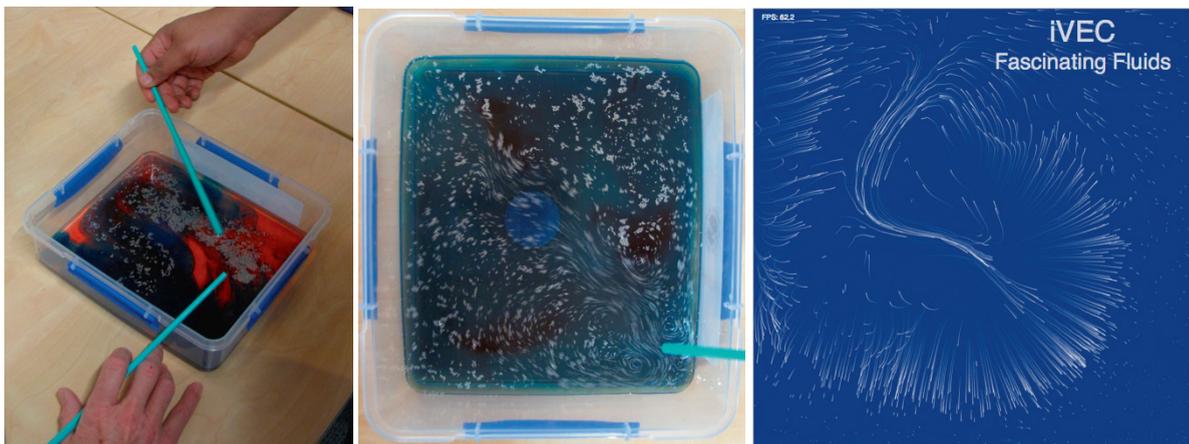


Fig. 4. Exploring fluids (a) coloured dyes, (b) glitter swirls, (c) computational model

In the final discussion we talked about computational fluid dynamics and how what we were looking at had many application areas. Of close similarity to the lava lamps is modeling the flow of blood through arteries as the blobs drew some resemblance to red blood cells. With the activities complete, the students were taken to the nearby iVEC facility where they could see the GPU cluster – Fornax, before being collected by their parents.

4.4. Workshop 4: Central Institute of Technology (CIT)

This workshop was presented as part of an assessment for a course the author attended for a training qualification. As a result, unlike the other workshops, this group had not opted in to attending a supercomputing and computational science workshop. The group comprised adults in the 25-50 age group, working in areas including occupational health and safety and the transport industry. Unsurprisingly, they were taken aback by the topic and quite daunted. An icebreaker activity was used to help draw attention to the number and variety of computers they use every day. Participants wrote on sticky notes and computers they used and put them on a table (Fig. 5), once a good number were there we sorted them into local and remote (and cloud), and noted the ones they missed – Google, Facebook, Gmail, Twitter, Amazon and so on.

The focus then moved to supercomputers – including the clarification of clusters and datacenters as distinct from supercomputers (fast interconnect being key). Activities 1 and 2 from Lesson 1 were used. Following that we moved onto Lesson 3 – a condensed version made more challenging with local firewalls and blocked javascript. With limited time, we used Wet Lab 1, discussed Dry lab 1 as a group, skipped to Wet lab 3 and discussed Dry lab 3. In all the workshop took an hour, which wasn't enough but the enthusiasm and involvement from participants got us through. A bonus of the venue was the bean bags in the room which helped add to the atmosphere for the lava lamps.



Fig. 5. Daily computers from icebreaker activity a) Central Institute of Technology and, b) ACEC Workshops

4.5. Workshop 5: Australian Computers in Education Conference (ACEC) Workshop

This three-hour workshop was an optional, fee-paying workshop in the pre-conference program. Attendees were computing teachers from schools, interested in supercomputing to add interest to their classes. Participants were given a folder including the compiled lesson plans, a set of Fascinating Fluids lab notes and IVEC promotional material.

The session began with a short introduction followed by the icebreaker activity described in Section 4.4. A few members of the group became quite competitive with their contribution of sticky notes (yellow and green in Fig. 3b). This was followed by Lesson 3: Fascinating Fluids, delivered as described in Section 4.3 (with greater discussion and questioning by the teachers). After a short break we went through activities 1 and 2 from Lesson 1. It was important, particularly with an adult group, to have alternatives and opt-outs for the lolly, balloon and running around activities, offered with discretion.

The final activity was “Dude where’s my data?” from Lesson2. Four animal lotto sheets were used with cards distributed randomly (but equivalently) at the home table (RAM), a table across the room (disk) and a table upstairs and down the hall (tape). Observers made suggestions while the participants struggled to remember which cards were where and did plenty of extra trips. The second run of the task employed helpers to pre-fetch data and required far less exertion. To wrap up the workshop, we visited the nearby GPU cluster before boarding the bus. We were able to drive past iVEC’s next supercomputer facility, the Pawsey Centre building site, before returning the group to their conference venue.

5. Reflection and Feedback

A goal-question-metric approach [7] was developed to assess the workshops. The three goals defined for each workshops and lessons were:

G1: *Communicate concepts*

G2: *Generate interest*

G3: *Be technically correct*

Each of these goals was broken down into a series of questions for which a metric or evaluation can be made. We will use [PASS|OK|FAIL] for the assessment. The questions were different for each lesson, but will be aligned for this paper.

G1Q1: *Does the attendee have more knowledge after the workshop?* In the online post-workshop survey from the Scitech workshop, all participants correctly answered the question “How does the wet/dry lab work?” which they would not have known before the workshop. Some questions were revisited between the pre-workshop and post-workshop surveys. In the post-workshop responses to “What is a supercomputer?” many attendees added the concept of a supercomputer being made of multiple computers (processors) that work together and communicate to solve problems. Similarly for the question “What is computational science?”, initially most responded that it was the science of computers, whereas post-workshop most said it was science using computers. Based on these responses, and assuming similar impact from the other workshops, we assess this question as [PASS].

G1Q2: *Was background material provided?* Some background material has been provided with the lesson plans, and is given verbally at the workshops. There is a need for more background material to support teachers delivering the lessons, so the assessment is [OK], but more work is required.

G2Q1: *Did the attendee find the topic interesting?* The post-workshop survey had most students rating Workshop 3 as “Pretty cool” or “Awesome”, one said it was “OK” and there were no negative responses. Participant evaluation forms from the CIT workshop included the phrases: informative, interesting, fun, held interest, great demonstrations, very interesting. The CIT assessor said “This was an excellent couple of sessions that explained the concept of supercomputers and of the applications in an interesting/interactive and fun way”. This question will receive a [PASS].

G2Q2: *Did the class group get into the activity?* The engagement and appreciation of all of the workshops indicates that we have succeeded in creating accessible and fun materials. Observation and facilitation of the workshops had all learners involved throughout – across all ages and backgrounds, including those who didn’t choose to attend such a workshop (CIT). The Scitech host indicated that she was surprised at how the children stayed on task and contributed, compared to other holiday program workshops she has hosted. Overall this question has been [PASS]ed.

G3Q1: *Are the concepts depicted accurately?* The outreach materials draw upon the iVEC training materials we deliver to university researchers. Those materials are developed with the input of supercomputing experts, so we consider this a [PASS].

G3Q2: *Are the activities accurately illustrating the concepts?* Overall the activities and models are matched to concepts, although they are understandably simplified. In the sessions we state where the analogies don’t

match reality and indicate what happens in reality (where the audience is capable of the more complex information). This is given a [PASS] however it would be valuable to work through the materials more methodically with a supercomputing expert evaluating their accuracy across G3Q1 and G3Q2.

Overall the assessment is five passes and one OK, with some points to follow up on. We would also like to have future groups complete the pre and post-surveys to provide more data. This may draw time away from the workshop and can be difficult with young attendees, however there is value for us and the reflection may enhance retention by the participants.

6. Conclusion and Future Work

This paper described materials and workshops for outreach developed at iVEC. The response to the workshops has been positive with high engagement and evidence that attendees have learnt from the activities. Materials have been distributed to teachers and are available to be reused and the Creative Commons attribution/share-alike licence. The materials are a work in progress and will be extended with more lessons, and more background material. We will continue to run professional development for teachers and offer outreach to schools and the community, building on this positive start.

Acknowledgements

The author acknowledges the contributions and feedback on these materials and workshops from her iVEC colleagues and her children. In particular Dr Rebecca Hartman-Baker and Dr Daniel Grimwood for the jigsaw and masterchef activities. Thanks to all the attendees and the organisations that supported the workshops: NYSF, Scitech, ECAWA, Central Institute of Technology and the ACEC Conference. We are also grateful for the use of the SPICE lab at UWA that was the venue for the Scitech and ACEC Workshops.

For access to the materials described in this paper, please contact the author.

References

- [1] Interactivate – Java-based courseware for exploration in science and mathematics, URL: <http://www.shodor.org/interactivate/>, accessed 5/3/2013.
- [2] Computer Science 4 Fun - A magazine where the digital world meets the real world., URL: <http://www.cs4fn.org/>, accessed 5/3/2013.
- [3] Panoff B., Shiflet A.B. and Shiflet G., 2011. Introduction to Computational Thinking, Melbourne, Australia, URL: <http://messagelab.monash.edu.au/IntroductionToComputationalThinking>, accessed 7/1/2013.
- [4] Shiflet A.B. and Shiflet G., 2011. Computational Workshop for the Life Sciences Classroom, Perth, Australia, URL: <http://messagelab.monash.edu.au/PerthComputationalThinking>, accessed 7/1/2013.
- [5] Scitech, 2013. Scitech, URL: <http://www.scitech.org.au/>, accessed 7/1/2013.
- [6] Greenwold, S., 2011. A Fluid Simulation in Javascript for the HTML5 Canvas, URL: <http://upfork.com/articles/javascript-fluid-simulation-for-html5-canvas/>, accessed 7/1/2013.
- [7] Basili V., Caldeira G. and Rombach H. D., 1994. The Goal Question Metric Approach, New York, USA: John Wiley and Sons, <http://www.wagse-old.informatik.uni-kl.de/pubs/repository/basili94b/encyclo.gqm.pdf>