

Title Text

Subtitle Text, if any

Dan Szymczak

McMaster University
szymczdm@mcmaster.ca

Jacques Carette Spencer Smith

McMaster University
cchette@mcmaster.ca / smiths@mcmaster.ca

Abstract

This will be filled in once we're done drafting. For now I'm putting in a solid block of text that will let me ramble on for a while when it comes down to the final abstract. Programmers are like gods of tiny universes. They create worlds according to their rules. If we think of software artifacts (that is, all documentation, build instructions, source code, etc.) as their own tiny worlds, a universe takes shape. Now think about what would happen if the laws of that universe were inconsistent between worlds. What if the laws of thermodynamics fundamentally changed as soon as you left Earth? Say goodbye to space travel. Keeping artifacts consistent with each other is hugely important in the software world, and Drasil aims to ensure it...

Categories and Subject Descriptors CR-number [*subcategory*]: third-level

Keywords Literate software, knowledge capture, traceability, software engineering, scientific computing, artifact generation, software quality

1. Introduction

We want to make better software. In particular, we are interested in increased maintainability, traceability, reproducibility, verifiability, and reusability. Our approach is to invest (much) more in the short-term to provide outstanding long-term benefits. This is the fundamental trade-off in our work: we expect our methods to work very well for domains which are well understood with known (but potentially enormous) design spaces. We do not claim to tackle areas of software development where methodologies such as agile are well-suited. But we firmly believe that there are domains – such as safety-critical applications – where agile is not only ill-

suited, it would be downright unprofessional to use such a methodology in that setting.

Rather than talk in vague generalities, we will pick Scientific Computation (SC) for illustrative purposes throughout this paper. But the fundamental ideas (and, in fact, our framework) should be applicable to any software domain which has well-established theoretical underpinnings, and a well understood translation of the theory into effective code. SC is particularly well-suited as it is also replete with *program families*.

We believe that Knuth's Literate Programming (LP) (Knuth 1984) contains some fundamental insights, but is too restricted (i.e. just to code). Rather than restrict LP to just source code, we want to apply it to all software artifacts (requirements and design documents, source files, test cases, build instructions, user manuals, etc.) from a singular common knowledge base.

We want to break away from the LP idea of "one source", instead introducing the idea of "common knowledge" libraries for problems that come up regularly. For example: the theory behind heat transfer including the conservation of energy equations would be in a "common knowledge" library. We could also see the idea of the *Système International* (SI) Units as another piece of "common knowledge".

We also very much want to keep the idea of "chunks" and assemble everything from chunks. How we do that will be discussed in more detail in Section 2.

We intend to make working on program families trivial for the cost of losing the capability to create "local hacks" to "just get things working".

Another disclaimer before we go further: none of the ideas presented in this paper are new. We are building on a very longstanding history of work, and not being particularly creative with the ideas therein. Knuth himself even said that "I have simply combined a bunch of ideas that have been in the air for a long time" when he coined LP (Knuth 1984). We are, however, taking old ideas and breaking them down into something practical that we can work with. We want to *do* something, not spend years stuck in the design phase.

Others have done similar work (which we will get into in Section 3), but they did not achieve the results we are

envisioning. They either set their aims on other targets, spent too long creating a grand design and ended up without any real, practical results, or they simply were not brave enough to break things down into the smallest necessary chunks.

We believe that we have assembled the right ideas to achieve our vision. The overall success or failure of our approach hinges on the idea of a stable, well understood knowledge base. Without that we might as well give up and go home. If we do not understand the fundamentals behind the software we intend to create, then we will be unable to properly encapsulate that fundamental knowledge for use in our framework (effectively getting us nowhere).

As it stands, we are on a (not so) humble, practical route to achieving our goals and improving the overall quality of software.

2. Drasil

Our framework is called Drasil (shortened from *Yggdrasil* from Norse Mythology, which is also known as the *world tree*). It is currently being developed through a practical, example-driven approach in an effort to bring our vision to life. There are three main ideas driving Drasil’s development:

1. Organize the knowledge base – We want a knowledge base that we can structure conceptually (i.e. keep knowledge for a certain class of problems together). This is where chunks come into play: each chunk encapsulates a single piece of knowledge like a concept or a quantity. We want to break our knowledge base down into the smallest possible pieces.
2. Use recipes to create artifacts – We can think of each artifact in our software project as a different view of our knowledge base. We want to use recipes to specify exactly what information from the knowledge base is necessary for each artifact, and how that information should be displayed. For many artifacts we would like to have a standard recipe which can be quickly customized for the current problem and that is how we will avoid duplicating knowledge.
3. Remove technology constraints – We want to be able to create our software without worrying about the underlying technical constraints of our display or specification technology, programming language(s), etc. Anyone using Drasil should be able to work with the knowledge base and their recipes, then simply set their output technology and have the generator take care of all the technical details.

We argue that by implementing Drasil around these ideas, we can see drastic improvements in software quality. In fact, using a generative approach we can avoid certain problems altogether. One obvious and recurring problem that comes to mind when upgrading software is “feature creep” or “software bloat” (Amsel et al. 2011). With Drasil, a software up-

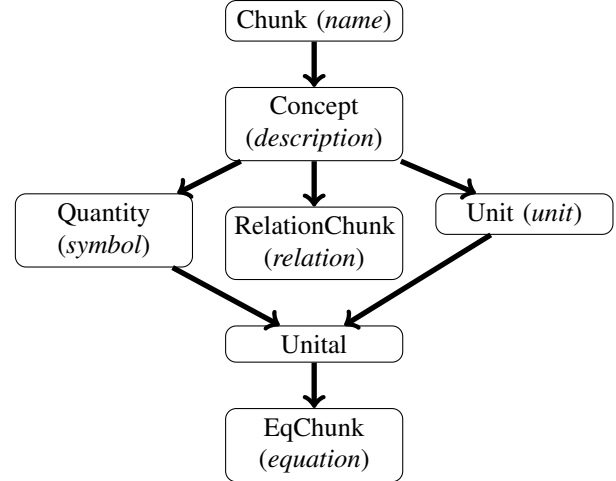


Figure 1. Our current chunk design

grade will actually be a completely new piece of software that includes the previously desired functionality as well as the new upgrades. We will discuss further improvements in more detail after our example in Section 2.2.

2.1 Drasil’s current implementation

Fundamentally, Drasil must be able to capture knowledge and produce different views of that knowledge. With our current implementation we have each individual piece of knowledge as a named *chunk*. We are then able to manipulate our chunks through the use of a *recipe*. Our *generator* interprets the recipes to produce the final desired view. This view represents one of the many software artifacts mentioned in Section 1.

As mentioned, we capture knowledge into named chunks, of which there are several varieties. Actually, we have a hierarchy of chunk types, where each new chunk encapsulates more than the previous one(s) (see Figure 1 for an idea). The most basic *chunk* represents a named piece of information. Above that we have named *concepts* which introduce some descriptive information.

A *quantity* is a concept that has a symbolic representation, we can refer to the quantity by either its name or symbol. In a similar vein are *units*, and *relation chunks* (which add the idea of a relation between some other pieces of knowledge).

In an SC context, most of the knowledge we work with is represented as a quantity with some units, in other words a *unital* chunk. Expanding on that, we can actually calculate values for many of these unital concepts. As such, we have *eqchunks* which allow us to capture the equation along with everything else included in a unital chunk.

Now with the means to encapsulate knowledge, we can turn our attention to our recipes. The recipes are implemented using a combination of embedded Domain Specific Languages (DSLs) in Haskell. We have currently implemented the following DSLs:

1. Expression language – A simple expression language that allows us to capture knowledge relating to equations and mathematical operations. It includes, but is not limited to, operations such as addition, multiplication, negation, derivation, and exponentiation.
2. Expression layout language – A micro-scoped language for describing how expressions should appear. Expressions may need to use subscripts, superscripts, concatenated symbols, etc. to be properly displayed.
3. Document layout language – A macro-scoped language for describing how large-scale layout objects (tables, sections, figures, etc.) should appear.
4. C Representation Language – A DSL for representing parts of the C programming language inside the Drasil framework. This allows the generator to produce working C code.
5. \LaTeX Representation Language – A DSL for representing \LaTeX code inside of Drasil.
6. HTML Representation Language – A DSL for representing HTML within Drasil.

Of the DSLs mentioned, we only actually have to write our recipes using three of them. Each of the representation languages are used strictly by the generator as an intermediary in the production of our desired views. We write our recipes using the document layout language, expression layout language, and expression language. We will discuss the particulars of each of these in more depth during our example (Section 2.2).

The last piece of the puzzle is the generator. We use it to interpret the recipes, create intermediary representations of our desired views, and then pretty-print them.

2.2 Drasil in action

For this section we will take a look at a simplified version of a Software Requirements Specification (SRS) for a fuel pin in a nuclear reactor (for more information on that particular SRS see (Smith and Koothoor 2016)).

Starting off we will look at one specific term: h_c . In this example, h_c represents the convective heat transfer coefficient between the clad and coolant. The data definition for h_c from the original SRS can be seen in Figure 2.

The data definition of h_c displays some interesting knowledge. It gives us the name for a concept, its description, a symbol to use for easier reference, its (SI) units, and its defining equation. Encapsulating all of this knowledge into a chunk is not very difficult as shown in Figure 3. Note that the equation for h_c is written in our expression language (Expr) and it also includes references to chunks that h_c depends on. These will come into play when we create the description for h_c in the generated view, but we do not need to worry about them right now as that knowledge is stored elsewhere. The

Number	DD2
Label	h_c
Units	$ML^0t^{-3}T^{-1}$
SI Units	$\frac{\text{kW}}{\text{m}^2\text{°C}}$
Equation	$h_c = \frac{2k_c h_b}{2k_c + \tau_c h_b}$
Description	h_c is the convective heat transfer coefficient between clad and coolant k_c is the clad conductivity h_b is the initial coolant film conductance τ_c is the clad thickness

Figure 2. Data definition for h_c from the fuel pin SRS

```

h_c_eq :: Expr
h_c_eq = 2*(C k_c)*(C h_b) /
  (2*(C k_c) + (C tau_c)*(C h_b))

h_c :: EqChunk
h_c = fromEqn "h_c"
  "convective heat transfer coefficient
   between clad and coolant"
  (sub h c) heat_transfer h_c_eq

```

Figure 3. The h_c chunk in Drasil

units for h_c are defined by a piece of common knowledge known as `heat_transfer`.

On the topic of common knowledge, we can show an example alluded to previously, that of the SI unit library. The seven base SI units captured in Drasil are shown in Figure 4. The SI unit library also contains several derived units (for example: centigrade which is derived from kelvin) which we make use of later on in our example. The knowledge behind each unit and (in the case of derived units) their relation to the base units is captured in the relevant chunks.

Now that we've shown a bit of knowledge capture, you may be wondering how to use it. This is where the recipes come in. Figure 5 shows a portion of the recipe for creating our intended SRS view. The body of the simplified SRS is composed of three sections: `s1`, `s2`, and `s3`. We then show the definition of the first section (omitting the others for brevity) which is titled "Table of Units" and includes an introductory paragraph and a table. This table simply extracts the symbol and description information from the SI units to display them in our view. It should be fairly obvious at this point, but these macro-scale layout objects (sections, tables, paragraphs, etc.) are specified using our document layout language.

With our recipe in place, we are now able to run it through the generator and see what it spits out. We've included the

```

metre, second, kelvin, mole, kilogram, ampere, candela :: FundUnit
metre    = fund "Metre"    "length (metre)"    "m"
second   = fund "Second"   "time (second)"     "s"
kelvin   = fund "Kelvin"   "temperature (kelvin)" "K"
mole     = fund "Mole"     "amount of substance (mole)" "mol"
kilogram = fund "Kilogram" "mass (kilogram)"     "kg"
ampere   = fund "Ampere"   "electric current (ampere)" "A"
candela  = fund "Candela"  "luminous intensity (candela)" "cd"

```

Figure 4. The seven funamental SI Units in Drasil

```

srsBody = srs [h_g, h_c] "Spencer Smith" [s1,s2,s3]
s1 = Section (S "Table of Units") [intro, table]
table = Table
  [S "Symbol", S "Description"] (mkTable
    [(\x -> Sy (x ^. unit)),
     (\x -> S (x ^. descr)) ] si_units)
intro = Paragraph (S "Throughout this ...")

```

Figure 5. A portion of our simplified SRS recipe

Table of Units

Throughout this document SI (Système International d'Unités) is employed as the unit system. In addition to the basic units, several derived units are employed as described below. For each unit, the symbol is given followed by a description of the unit with the SI name in parentheses.

Symbol	Description
m	length (metre)
kg	mass (kilogram)
s	time (second)
K	temperature (kelvin)
mol	amount of substance (mole)
A	electric current (ampere)
cd	luminous intensity (candela)
°C	temperature (centigrade)
J	energy (joule)
W	power (watt)
cal	energy (calorie)
kW	power (kilowatt)

Figure 6. Section 1 of the generated SRS (HTML version)

three sections of this simplified SRS in Appendix A, and what you see there is the typeset version of the generated LaTeX code. We are also able to output our view as an HTML document (Figure 6).

Now if you recall the data definition from Figure 2, we will show the painstaking amount of work it takes to create (almost) that exact same table:

```
s3_dd2 = Definition (Data h_c)
```

Again looking to the appendix, or to the HTML output (Figure 7 for the data definition), it should be easy to see

Label	DD: h_c
Units	$^{\circ}\text{C}^{-1}\text{s}^{-3}\text{kg}$
Equation	$h_c = \frac{2h_c h_b}{2h_c + \tau_c h_b}$
Description	h_c is the convective heat transfer coefficient between clad and coolant h_c is the clad conductivity h_b is the initial coolant film conductance τ_c is the clad thickness

Figure 7. Generated SRS data definition (HTML version)

that the recipe handled all of the layout details. The recipe actually goes one step further in that it found all of the chunk's dependencies and built the description by finding the appropriate knowledge related to the other chunks. This is configurable in that we can include short descriptions or verbose descriptions at this point in time.

One last thing we'd like to show is the source code generation. The recipe is fairly straightforward in that we have:

```
codesample = CodeBlock (toCode CLang Calc h_c)
```

All that says is `codesample` is a block of code in the C language that calculates the value of h_c . Currently the only implemented output language is C, but we are planning to implement more and are designing around that coming change. The generator output for `codesample` ends up looking like:

```

double calc_h_c(double k_c, double h_b, double tau_c){
    return (2 * k_c * h_b / (2 * k_c + tau_c * h_b));
}

```

This is a fairly trivial piece of code, but it is a good example of how the source can be generated from the knowledge encapsulated in a chunk.

2.3 Advantages

We have already experienced some of the advantages of using Drasil. During the knowledge capture phase of our ex-

ample a seemingly trivial problem came up regarding the simplified SRS. One of the concepts was being referred to with multiple different descriptions. By capturing only the right description, the generator now ensures internal consistency.

With that in mind, look at all of the previous figures (or Appendix A) and how many times knowledge has been copied into different places within one software artifact. Now imagine how many times that knowledge appears across all software artifacts. By using a generator, we avoid the problem of manually duplicating all of that information. It also ensures that should a piece of knowledge need to be updated, those updates will propagate throughout all of the artifacts *automatically*.

Another great benefit to knowledge capture, especially in the SC domain, is that it promotes reuse. The SI units are a trivial example of a completely reusable knowledge library that will come up across many different problem domains (within and outside of SC). If we can build these common knowledge libraries for a multitude of problem domains, we believe it will aid in the creation of higher quality software overall and will allow developers and scientists to spend their time on more important things.

Speaking of more important things, what about software certification? Many types of safety- or security-critical software must be certified. Depending on the certification process and regulatory body that means including different types of high-quality documents along with the source code. As we all know, requirements or algorithm choices can and do change throughout the software development process leading to a need for updated documentation. With Drasil we are able to ensure all of the software artifacts always remain consistent both internally and with each other.

Drasil is meant to design for change. We mentioned previously (Section 1) that we intend to make working on program families trivial, and Drasil does just that. Implementing a likely change is as simple as adjusting a configuration file. If the knowledge has been properly captured, any likely changes will involve trivial modifications. In the event of an unforeseen change, modifying the recipe or (worse) the captured knowledge is still possible and should end up simplifying the development process.

Finally, we need to verify our software and Drasil can help with that as well. We can include so-called “sanity” checks in the knowledge we capture that can be reused throughout the entire development process. A simple example of these types of checks can be seen in the Constraints column of Table 1.

If we encapsulate the knowledge regarding these constraints into the appropriate chunks, we can have Drasil automatically test them for us. For example: a chunk that contains the dimensions of a physical object should also constrain those dimensions to positive values.

Table 1. Constraints on quantities

Var	Constraints	Typical Value	Uncertainty
L	$L > 0$	1.5 m	10%
D	$D > 0$	0.412 m	10%
V_P	$V_P > 0$	0.05 m ³	10%
A_P	$A_P > 0$	1.2 m ²	10%
ρ_P	$\rho_P > 0$	1007 kg/m ³	10%

Another need in verification is traceability. With Drasil we have introduced complete traceability: every piece of knowledge can be tracked (through all software artifacts) to its source! We can also check that all chunks are actually necessary, since the recipes can automatically report which chunks they use.

With all knowledge coming from unique chunks, we get one last added benefit (which can also be seen as a disadvantage): pervasive bugs. Any error in the knowledge base will spread through every artifact and even a single bug could break everything. However, we still see this as an advantage since the bugs tend to be fairly shallow, easy to fix, and now much harder to miss.

2.4 Disadvantages

We mentioned one of the disadvantages in Section 1: the inability to create a local hack to get things working. With Drasil we have to take an all-or-nothing approach. Any hacks added to the generated artifact sources will be overwritten the next time anything is updated.

With an all-or-nothing approach we also remove the ability to iterate in a meaningful way. We *must* do everything right (or extremely close to it) the first time. Yes, we can still modify the program family member that we are generating and adapt it as necessary, but adding new knowledge becomes a much more involved process.

Adding new common knowledge libraries is also non-trivially difficult. These must be created by (or with the help of) domain experts to ensure the right knowledge is being captured in the right way.

3. Related work

As we mentioned previously, none of the ideas behind Drasil are new. There has been a lot of work done by a lot of smart people and we are taking advantage of it.

First let us take a look at the current state of SC software development, then we will discuss tools and techniques that came before us.

3.1 The current state of SC software development

Most developers in SC software tend to put an emphasis on their science, not the software they are developing (?). Consequently these developers tend to naturally use an agile philosophy (Ackroyd et al. 2008; Carver et al. 2007; ?; Segal 2005) or an amethodical (Kelly 2013), or a knowledge acquisition driven (Kelly 2015) process. These developers

are scientists first and foremost and they do not view rigid, process-heavy approaches favourably (Carver et al. 2007).

More than half of these scientists do not have a good understanding of software testing (?). In fact, there is very limited use of automated testing for scientific software (?) and quality assurance has “a bad name among creative scientists and engineers” (Roache 1998, p. 352).

Another problem in SC is that knowledge reuse is not fulfilling its potential. For instance, for mesh generators a large number of similar programs have been written. A survey (?) shows different mesh generator packages, with 61 of them generating triangular meshes, and of these using the same algorithm of Delaunay triangulation.

Tool use in SC software development is also limited, especially the use of version control software (Wilson 2006).

Code generation in SC has been successful, but thus far the focus has been on the creation of only one software artifact: the code. Examples include

The lack of emphasis on more artifacts is disadvantageous to SC software developers. The value of documentation and a structured process is illustrated by a survey of statistical software for psychology (??). A case study on the impact of a document driven process for SC software was performed on legacy software for thermal analysis of a fuel pin in a nuclear reactor (Smith and Koothoor 2016; Smith et al. 2013). The redeveloped version discovered 27 issues, ranging from trivial to substantive, within the previous documentation. If rational documentation were used more often, some previous errors in scientific code may have been uncovered. Examples include the Sleipner oil rig, protein retraction, the Patriot missile, and seismic data processing (?).

3.2 Reproducible research

3.3 Literate Programming

Literate programming is a programming methodology that was introduced by Knuth. The LP methodology changes the focus from writing programs which simply instruct the computer how to perform a task. Instead, the focus is now on explaining (*to humans*) what we want the computer to do (Knuth 1984).

In a literate program, the documentation and source code is all kept together in one source. Literate programs are developed by breaking down algorithms into small, easy to understand parts referred to as *chunks* (Johnson and Johnson 1997) or *sections* (Knuth 1984). These chunks are then explained, documented, and implemented in a way that promotes understanding – they are not necessarily written in the order that the computer would read them. To get the working source code for a program written in the LP style, you run the *tangle* process. Similarly, to extract and typeset the documentation you run the *weave* process.

A. Simplified SRS

Here we see the typeset output of the LaTeX code for the simplified SRS example:

Table of Units

Throughout this document SI (Système International d'Unités) is employed as the unit system. In addition to the basic units, several derived units are employed as described below. For each unit, the symbol is given followed by a description of the unit with the SI name in parentheses.

Symbol	Description
m	length (metre)
kg	mass (kilogram)
s	time (second)
K	temperature (kelvin)
mol	amount of substance (mole)
A	electric current (ampere)
cd	luminous intensity (candela)
°C	temperature (centigrade)
J	energy (joule)
W	power (watt)
cal	energy (calorie)
kW	power (kilowatt)

Table of Symbols

The table that follows summarizes the symbols used in this document along with their units. The choice of symbols was made with the goal of being consistent with the nuclear physics literature and that used in the FP manual. The SI units are listed in brackets following the definition of the symbol.

Symbol	Description	Units
h_g	effective heat transfer coefficient between clad and fuel surface	$^{\circ}\text{C}^{-1}\text{s}^{-3}\text{kg}$
h_c	convective heat transfer coefficient between clad and coolant	$^{\circ}\text{C}^{-1}\text{s}^{-3}\text{kg}$

Data Definitions

Refname	DD:h.g
Label	h_g
Units	$^{\circ}\text{C}^{-1}\text{s}^{-3}\text{kg}$
Equation	$h_g = \frac{2h_c h_p}{2h_c + \tau_c h_p}$
Description	h_g is the effective heat transfer coefficient between clad and fuel surface

Refname	DD:h.c
Label	h_c
Units	$^{\circ}\text{C}^{-1}\text{s}^{-3}\text{kg}$
Equation	$h_c = \frac{2h_c h_b}{2h_c + \tau_c h_b}$
Description	h_c is the convective heat transfer coefficient between clad and coolant

References

- K. S. Ackroyd, S. H. Kinder, G. R. Mant, M. C. Miller, C. A. Ramsdale, and P. C. Stephenson. Scientific software development at a research facility. *IEEE Software*, 25(4):44–51, July/August 2008.
- N. Amsel, Z. Ibrahim, A. Malik, and B. Tomlinson. Toward sustainable software engineering (nier track). In *Proceedings of the 33rd International Conference on Software Engineering*, pages 976–979. ACM, 2011.
- J. C. Carver, R. P. Kendall, S. E. Squires, and D. E. Post. Software development environments for scientific and engineering software: A series of case studies. In *ICSE '07: Proceedings of the 29th international conference on Software Engineering*, pages 550–559, Washington, DC, USA, 2007. IEEE Computer Society. ISBN 0-7695-2828-7. doi: <http://dx.doi.org/10.1109/ICSE.2007.77>.
- A. Johnson and B. Johnson. Literate programming using noweb. *Linux Journal*, 42:64–69, October 1997.
- D. Kelly. Industrial scientific software: A set of interviews on software development. In *Proceedings of the 2013 Conference of the Center for Advanced Studies on Collaborative Research, CASCON '13*, pages 299–310, Riverton, NJ, USA, 2013. IBM Corp. URL <http://dl.acm.org/citation.cfm?id=2555523.2555555>.
- D. Kelly. Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software. *Journal of Systems and Software*, 109:50–61, 2015. doi: 10.1016/j.jss.2015.07.027. URL <http://dx.doi.org/10.1016/j.jss.2015.07.027>.
- D. E. Knuth. Literate programming. *The Computer Journal*, 27(2):97–111, 1984. doi: 10.1093/comjnl/27.2.97. URL <http://comjnl.oxfordjournals.org/content/27/2/97.abstract>.
- P. J. Roache. *Verification and Validation in Computational Science and Engineering*. Hermosa Publishers, Albuquerque, New Mexico, 1998.
- J. Segal. When software engineers met research scientists: A case study. *Empirical Software Engineering*, 10(4):517–536, October 2005. ISSN 1382-3256. doi: 10.1007/s10664-005-3865-y. URL <http://dx.doi.org/10.1007/s10664-005-3865-y>.
- S. Smith and N. Koothoor. A document driven method for certifying scientific computing software used in nuclear safety analysis. *Nuclear Engineering and Technology*, Accepted, 2016. 42 pp.
- W. S. Smith, N. Koothoor, and N. Nedialkov. Document driven certification of computational science and engineering software. In *Proceedings of the First International Workshop on Software Engineering for High Performance Computing in Computational Science and Engineering (SE-HPCCE)*, page [8 p.], November 2013.
- G. V. Wilson. Where’s the real bottleneck in scientific computing? Scientists would do well to pick some tools widely used in the software industry. *American Scientist*, 94(1), 2006. URL <http://www.americanscientist.org/issues/pub/wheres-the-real-bottleneck-in-scientific-computing>.