BECKHOFF

Review of
Doctoral Thesis
by M. Eng. Dipl.-Ing. Michael Steindl

Univ.-Prof. Dr.-Ing. Frank Schiller
Scientific Safety & Security
Ostendstr. 196
D-90482 Nuremberg, Germany
Phone: +49 911 54056-244
Fax: +49 911 54056-29
f.schiller@beckhoff.de

14 January 2013

Embedded software systems are getting more and more complex. The demand for new features and functions lead to an increasing complexity in the design and development of these systems. One answer to handle this complexity is component-based development, in which systems are built of individual, logically independent software components. These components collaborate in order to provide the complete set or a specified subset of the capabilities the final system is supposed to provide.

One important aspect of the component-based development approach is software integration. The individual components have to be put together, and their interactions have to be verified with respect to the specification. The crucial point of verifying a stepwise integration is the order in which the components are combined. State-of-the-art approaches (e.g. top-down or bottom-up integration) are only coarse guidelines and rely strongly on integrators expertise. More elaborate methods in which an algorithm is used to derive an integration order are only available for object-oriented software and cannot be directly used in procedural programming languages, e.g. in the language C. To deal with these challenges, parameters the software integration process is subjected to are identified and metrics are developed in order to evaluate a certain integration order in the verification phase. Furthermore, an optimization approach based on simulated annealing is presented which is used to derive an integration order with respect to the proposed parameters in a powerful and reliable manner.

The subject of the doctoral thesis by Mr Steindl is significant in various branches, although he refers mainly to examples of automotive software systems. The more complex software systems are, the more they are prone to errors. Therefore, efficient verification and tests are necessary. Instead of verifying the comprehensive software system at once, verification is focused on components and subsystems, respectively. This approach does not only reduce the number of test cases but allows stimulation as well as evaluation of internal data. This kind of data access is usually complicated or even impossible using the overall software system. In this way, a strategy is necessary for the stepwise integration of components in order to identify ones with erroneous

behavior unambiguously and as early as possible. The reviewer would have expected some definitions of that field, especially regarding the terms *test* and *verification*.

Mr Steindl analyzes important approaches of literature after he introduced a helpful categorization (formalized and non-formalized approaches, architectural and functional approaches, cf. Fig. 2.6). He elaborates advantages and disadvantages of the approaches in a convincing way (Sec. 2.3) and motivates the overall thesis correspondingly.

Since Mr Steindl finally handles the problem of finding the appropriate integration strategy as an optimization problem (in Chapter 4), he identifies necessary parameters in Chapter 3. Again, known algorithms in the literature are used as the basis for analysis and discussion. The links to graph theory are remarkable.

In Chapter 4, Mr Steindl states the optimization problem to be NP-hard. The reviewer would have expected some more explanations about this statement. The optimization is finally executed by *simulated annealing*. The cost functions are introduced and results and preliminary results are explained. All difficulties related to a numerical solution of the optimization problem are described in an understandable way.

Chapter 5 (Conclusion) contains the main results of the thesis. Especially in Section 5.2, it is emphasized that the developed approach is scalable. Here some more explanations with respect to the statement of being NP-hard might be necessary. The conclusions are mainly based on two case studies from automotive software. Since the challenges there are much higher than in other branches, the results of the thesis are of great value to the industrial software community.

The results of the thesis are

- valuable literature research,
- formulation of the appropriate optimization problem,
- solution to the optimization problem,
- evaluation of the approach by means of two challenging case studies of automotive software, and
- clear formulation of further work.

The thesis by Mr Steindl constitutes an original contribution to the area of testing of complex software.

The introduced approach has been systematically developed and clearly demonstrated. Form and language of the thesis are appropriate.

Mr Steindl shows an impressive record of nine contributions in international journals and conference proceedings.

The reviewer recommends the thesis for defense and proposes the predicate "cum laude" (good).

Frank Schiller

(Prof. Dr. Frank Schiller)

Beckhoff Automation GmbH
Eiserstr. 5
33415 Verl
Germany

Postfach 11 42
33298 Verl

Phone: +49 (0) 5246/963 -0
Fax: Reception: -149
Sales: -198
Service: -479
E-Mail: info@beckhoff.com
www.beckhoff.com

General manager:
Dipl. Phys. Hans Beckhoff
Arnold Beckhoff
Register court: Gütersloh HRB 1803
Ust.-Id.-Nr.: DE 126787444
Finanzamt Wiedenbrück
St.-Nr. 347/5819/0016

Kreissparkasse Verl
BLZ 478 535 20
Kto.Nr. 4 000 766
SWIFT WELADEDD
BL/47853520
IBAN DE114785352
00004000766

Deutsche Bank Gütersloh
BLZ 480 700 43
Kto.Nr. 371/7014
SWIFT DEUTDE3B480

IBAN DE854807004
00371701400

Review
of the Dissertation Thesis

Evaluation and Determination of Integration Orders In Component-Based Embedded Systems

of Michael Steindl

The Thesis to be reviewed describes the results of the candidate when applying the *Simulated Annealing (SA)* approach to solve a number of variant problems related to optimization of software integration. According to the rules valid for Doctoral Thesis evaluation at the University of West Bohemia in Pilsen, this Review is split into the following six parts.

Importance of the Thesis in the field

Building applications for embedded systems represents a field with rapidly growing importance. The principle reason is that embedded software systems are becoming more and more complex reflecting the possibilities offered by cheap and powerful hardware components. New versions of such systems are built using extensive sets of modules (or components) that already exist mixed with re-designed or brand new software components. In order to obtain a reliable system, its integration apart from being done in time has to succeed in all needed testing. The order in which individual software components are integrated is one of the most important factors in this process.

The Thesis discusses a number of variants of cost functions that can be used by the SA approach to obtain an optimal integration order. Individual cost function variants differ in the parameters they use as optimization measures (test effort, integration schedule, stub number and/or complexity or a combination thereof) and the respective results are compared when applied to real-life examples. As a result, the Thesis extends the spectrum of formalized methods that are worth considering in the process of software integration.

Solution procedure and methods used to meet the Thesis goals

Even though the Thesis uses an already established heuristic method developed to solve global optimization problems, its novelty consists in a systematic exploration of a specific domain of usage of this method and in presenting extensive sets of computed results that make it possible to study properties of different variants of cost functions and their improving impact on real software integration problems.

Roughly half of the Thesis text is devoted to a detailed presentation and discussion of existing integration strategies and parameters that can be used to evaluate the quality of the respective integration process. The second part of the Thesis proposes meaningful variants of cost functions to be used within the Simulated Annealing methodological framework when applied to the problem of specifying the order in software components integration. The candidate concentrates his attention to component-based embedded software systems in order to compare his results with real-life examples taken from automotive industry.

Clearly, there is also a lot of programming and huge amount of computed experimental results that are hidden behind the simplified tables and figures presented in the Thesis.

Thesis results

The goals of the Thesis that are specified on page 18 can be briefly expressed as follows:

- To identify parameters to which software integration is subjected and to define metrics for each parameter in order to evaluate different integration order strategies.
- To develop methods and algorithms for optimizing an order of integration with respect to the defined parameters.
- To validate the developed approaches on real-life examples.

Regarding the first goal, after a detailed description of parameters used so far by other authors, the Thesis presents six different variants of cost functions that use such parameters of component integration as the number of stubs, the overall stub complexity, and integration schedule.

Regarding the second goal, each of the described integration parameters is used either as an isolated optimization criterion or in a mixture with other parameters and the corresponding case study results are presented. All computations are based on a common heuristic method called *Simulated Annealing* that was introduced some 30 years ago, so that the proper work of the candidate is concentrated around the development of meaningful forms of combining integration parameters into a cost function that controls the optimization process

Regarding the last goal, there are two real-life examples taken from the automotive industry that are used as reference systems to run the tests. The first example is an embedded data logger for battery management consisting of 16 components and 23 dependencies, the second example was extracted from the AUTOSAR model and consists of 72 components and 172 dependencies.

Structure, clarity, formal and linguistic quality of the Thesis

Although the Thesis contains a rather extensive description of the state-of-the-art in software integration, I still feel a number of logical gaps that prevent my better understanding of the work described in Chapter 4. Small scale figures make textual parts hardly readable, the captions mostly copy long explanatory text (for the worst example, see p. 53).

I miss a brief explanation of how a sample of 4×10^6 integration orders (used in case studies) has been selected out of 20×10^{12} possible permutation orders of 16 components of the data logger. I would appreciate as well at least a brief algorithmic description (e.g. in a pseudo-code) of how the numbers of needed stubs (both realistic and specific) and other parameters are calculated based on a dependency graph of components. It is not clear why the candidate uses so called *square addition* (and not e.g. an average value) when two or three parameters are mixed to calculate a cost function value. I would appreciate if the candidate clarifies these issues during the Thesis defense.

The most disturbing nuisance – at least to me – is the lack of care paid to the linguistic quality of the text. Usage of some terms is not systematic (e.g. "node" vs. "vertex"), correct "SCC" (strongly connected component) is mostly misspelled as "SSC", while usage of an incorrect abbreviation "eg." instead of correct "e.g." is totally overwhelming. In many cases, a missing suffix "d" or other types of error in past participle hurt the eyes.

Apart from formal defects, the text contains a surprising number of factual mistakes or inconsistencies as well. Does the symbol $d_G(v_x)$ on page 33 represent the number of incoming or outgoing edges or both taken together? What is the role and meaning of formula (3.6) on page 39? Why multiple resources time frames are aggregated using interval intersection (formulae (3.10) and (3.11) on page 40)? What is the meaning of the summation symbol in formula (3.12) on page 42? Wasn't there any better "definition" of configuration (page 46-47) at hand? ... etc.

Author's publications

There are altogether 9 publications, most of them representing papers at conferences of a required international level. The last item of the list represents a paper that has not been published yet, but it was submitted for publication. I consider the author's publication activities adequate.

Conclusion

The Thesis presents a detailed and well structured introduction into the topic of component-based integration of embedded software systems and brings novel methods of evaluation and determination of proper integration ordering of respective components. These novel methods are based on an application of Simulated Annealing heuristics using cost functions proposed and experimentally evaluated by the candidate. Although the results could have been presented in a better formal and factual form,

I recommend the Thesis for defense.

Prague, February 28, 2013

Assoc. Prof. Josef Kolář, CSc.
Czech Technical University in Prague
Faculty of Information Technology