Report on PhD thesis "Component-based engineering of Smart Cyber- Physical Systems"

Candidate: *Michał Kit*
Reviewer: *Jan Carlson, Mälardalen University, Sweden*
Date: *March 31, 2016*

Content and Contributions

The submitted PhD thesis addresses the challenges of developing software for cyber-physical systems where a large number of loosely coupled embedded systems, interacting with each other and their environment, collaborate to provide the desired functionality. Combining methods from the two research areas component based software engineering and agent-oriented software development, it presents a novel approach to design, implement and test such systems. The material in the thesis is based on six peer-reviewed publications in conferences and journals, and one technical report.

The first chapter introduces the area and the specific challenges and problems of developing this type of systems, and establishes the overall research goals. This gives a good basis for understanding the focus of the work presented in the thesis and the value of its contributions. Following the introduction, the author presents a survey of existing work in the area. In particular, component-, agent- and ensemble-based software development are described and compared, and their respective disadvantages and limitations are presented in order to motivate combining them.

Chapter 3 summarizes the thesis contributions, and Chapter 4 contains the seven papers that constitute the main content of the thesis. For each paper, a description is given of how it relates to the overall scope of the thesis and to the specific goals defined in chapter 2. Moreover, a brief statement is provided that explains the contribution of the thesis author in relation to the other co-authors of the paper.

The contributions can be divided in three parts: component model, runtime environment and simulation framework, respectively addressing three important activities in the software development process. The component model, DEECo, combines typical component characteristics, such as strong encapsulation and explicit dependencies, with concepts from agent-oriented features related to implicit component bindings and a high level of dynamicity. Component synchronization and orchestration is defined by means of ensemble definitions specifying criteria for ensemble membership as well as the resulting data sharing. The resulting component model is very interesting, with several novel concepts, and clearly extends beyond similar existing approaches.

The runtime environment is a concrete realization of the component model concepts, based on the Java programming language. The solution uses a layered architecture where component knowledge is separated from the definition of component and ensemble functionality, which allows for transparent choice of knowledge distribution methods. The runtime environment provides a proof of concept implementation of the component model concepts but also constitute a valuable contribution on its own; in particular, the use of a Gossip protocol to distribute knowledge globally without severe performance penalty. In addition, specific optimizations are proposed to further reduce the communication overhead by explicitly limiting communication to the components that are potential candidates to participate in an ensemble.

Finally, the third part of the contribution is a simulation framework where components can be executed in an environment simulating different network topologies and component mobility. The scientific impact of this part is less significant than that of the other parts, but the simulation framework facilitates experimentation and evaluation of the overall approach.

Presentation and writing style

The thesis is very well organized, well-written and the line of reasoning is easy to follow. The presentation suffers a bit from keeping the original formatting of the included papers, which results in different styles in different sections, and making some of the text unnecessarily small, but this is just a minor inconvenience.

Concerns and Questions

Since the thesis is a collection of papers co-authored by 4-6 persons, it is not straightforward to pinpoint the exact contribution by the thesis author. Most of the paper presentation texts state that the author has participated in the formulation of the main idea, but it remains unclear in which aspects of the work he has been the main contributor or driver. Regarding the implementation and evaluation aspects, it is more clear that the author has taken a more leading role.

The problem formulation is clear, as is the presentation of the solution, but there is only a rather short paragraph in Chapter 5 discussing the validation of the proposed approach. The individual papers contain validation and evaluation of their respective contribution, often in terms of a smaller case study, but it would have been interesting to see a more in-depth discussion about validity and of the limitations and disadvantages of the overall thesis contribution.

The state-of-art section gives a good introduction to the relevant areas and examples of existing methods, but I also expected a more in-depth discussion towards the end of the thesis, comparing the results of the thesis against these existing lines of work.

Detailed questions:

1. Section 2.1 describes the limitations of CBSE for the addressed domain, and later shows how agent-based development can help overcome them. Similarly, the identified issues of agent-based development are alleviated by ensemble-based architectures. Which concepts from CBSE do you find most valuable to bring into the agent- and ensemble-based development methods, and what benefits do they bring?
2. The research goal G3 is to "deliver a method for assessing the correctness of a developed SCPS – i.e. correspondence to its both functional and non-functional requirements". Is this goal really achieved by the simulation framework?
3. In Section 3.2, you state that "DEECoSim itself does not provide any guarantees over the critical aspects satisfiability of a developed SCPS. However, it can support any formal analyzes by delivering their validation through an experimental evidence." What do you mean by "delivering their validation"?
4. The exact rules for ensemble construction are not entirely clear, partly because ensemble definitions differ somewhat between the papers. Paper 4.1 states that bigger ensembles are preferred over smaller, which indicates that the mechanism on purpose disregard some possible ensembles (e.g. subsets of other ensembles), but this is not formalized. Does the mechanism guarantee that once the bigger ensemble

is created, the smaller ones will not be constructed? Similarly, when ensembles are defined in a way that allows any member to be the coordinator, could multiple ensembles be created or does the creation of the first ensemble invalidate the other alternatives consisting of the same components but with different coordinator?

5. Is there any difference between establishing that a previously created ensemble still exists, compared to discovering a new? Is the mapping executed also on previously established ensembles, and if so, how often?
6. Paper 4.3 (section 3.1 and 6.2) discusses the possible consequences of outdated belief leading to inconsistent views of ensemble membership, but couldn't this happen also with perfect knowledge distribution, as a result of periodic evaluation of membership criteria?
7. Communication bounds should be a safe approximation of ensembles. Would it be possible to statically determine from the definitions if this is the case?
8. Considering the overall DEECo approach, what are the main drawbacks and limitations when developing systems in this way, compared to, e.g., more traditional component-based development?

Overall Judgment

Smart cyber-physical systems is a rather new and very interesting area to research, with applications in a wide and increasing range of application domain, and the thesis' focus on architectural support for the dynamical nature of these systems is definitely relevant. The combination of component-, agent- and ensemble-based development has resulted in significant and novel contributions but also provided new insights in the respective areas. The thesis could be more precise in isolating the specific contributions by the author, and would benefit from a more in-depth validation and evaluation of the proposed approach, but in summary, the thesis clearly shows the candidate's capability of creative scientific work, and I recommend the thesis for a defense and judge the candidate worthy of a PhD degree.

Jan Carlson, Associate Professor, Mälardalen University, Sweden
March 31, 2016