A Protocol for a Systematic Review on Adaptation of Services Choreography

Gustavo Ansaldi Oliva
Fernando Hattori
Leonardo Alexandre Ferreira Leite
Marco Aurélio Gerosa

{goliva, hattori, leofl, gerosa}@ime.usp.br

Department of Computer Science
Institute of Mathematics and Statistics, University of São Paulo
PO-BOX 66281 – 05508-090 – São Paulo – SP – Brazil

December 2011

A Protocol for a Systematic Review on Adaptation of Services Choreography

Abstract- *Services choreography represents a description of the observable behavior of each of the services participating in the interaction, which is defined by public exchange of messages, rules of interaction and agreements between two or more business process endpoints [6]. Despite the recent research on Web services, little is known about approaches and influencing factors in the adaptation (reconfiguration) of choreographies. A systematic literature review evaluates and interprets all available research relevant to a particular question or topic area through a rigorous and auditable methodology [5]. This study presents a protocol for a systematic literature review on adaptation of service choreographies. The protocol structure employed in this study is mainly derived from the guidelines offered by [1], [2].*

1- Research Question Specification

1.1- Question focus

Despite the recent research on Web services, there is an scarce amount of material in the literature regarding implementations of service composition strategies that fit the paradigm of choreographies. In particular, little is known about strategies in the adaptation (reconfiguration) of choreographies. More precisely, by “strategies” we mean “any approach or method for the adaptation of choreographies”.

1.2- Question quality and amplitude

1.2.1- Problem

Service orchestration enables Web services to be composed together in predefined patterns, described using an orchestration language and executed with an orchestration engine. On the other hand, service choreography is more collaborative in nature. A service choreography is a description of the peer-to-peer externally observable interactions that exist between services; therefore, choreography does not rely on a central coordinator. A choreography model describes multiparty collaboration and focuses on message exchange; each Web service involved in a choreography knows exactly when to execute its operations and with whom to interact. A choreography definition can be used at design-time to ensure interoperability between a set of peer services from a global perspective, meaning that all participating services are treated equally, in a peer-to-peer fashion [7].

Choreography description languages specify interactions among a set of services from a global point of view. Generation or discovery of peers (participants) that precisely implement the choreography specification is not always possible: this problem is known as realizability. When peers are being executed, one may want to modify the choreography specification and dynamically reconfigure the system. For instance, interactions between service may be added (choreography extension) or removed (choreography simplification) due to the addition of functionalities to the system at hand or the loss of a service. Also, one may want to modify the choreography as a result of services whose specific attributes violate predefined QoS threshold values. This reconfiguration involves defining a new choreography that replaces the previous one, occasionally leading to the generation of new peers. Therefore, in this work we intend to better understand the state of the art concerning the existing approaches to choreography adaptation.

1.2.2- Question

Staples and Niazi recommend limiting the scope of a systematic literature by choosing clear and narrow research questions [8]. Following this guideline, we state the questions as follows:

Q1) What strategy each selected study uses to deal with service choreography adaptation?

Q2) How each selected study characterizes its adaptation strategy according to the following aspects?

i) *Target*: does the adaptation support functional and/or non-functional requirements changes?

ii) *Required intervention degree*: is the adaptation automatically performed? Or is human intervention necessary?

iii) *Scalability impact*: is the strategy impact on choreography scalability discussed? Is such discussion informal or does it contain formal proofs/experiments?

iv) *Implementations*: is the strategy implemented by a tool or prototype? Is the implementation available to download? If so, is it open-source software?

v) *Underlying models*: which choreography models, representations or standards (WS-CDL, WSCI, BPMN, etc.) are used in the strategy?

1.2.2.1- PICOC acronym:

** Medical guidelines recommend considering a question about the effectiveness of a treatment from three perspectives: population, interventions and outcomes. More recently Petticrew and Roberts suggest using the PICOC criteria to frame research questions [3]. These criteria extend the original medical guidelines with: comparison and context. The next subsections describe these criteria from the viewpoint of software engineering.*

- Population

** In software engineering experiments, population might be any of the following: a specific software engineering role (e.g. testers, managers), a type of software engineer (e.g. a novice or experienced engineer), an application area (e.g. IT systems, command and control systems), an industry group (e.g. telecommunication companies, small IT companies), etc.*

Service choreographies

- Intervention

** The intervention is the software methodology/tool/technology/procedure that addresses a specific issue, for example, technologies to perform specific tasks such as requirements specification, system testing, or software cost estimation.*

Strategies and tools for the adaptation of choreographies.

- Comparison (optional)

***** *This is the software engineering methodology/tool/technology/procedure with which the intervention is being compared.*

Not applicable

- Outcome

** Outcomes should be related to factors of importance to practitioners such as improved reliability, reduced production costs, and reduced time to market. All relevant outcomes should be specified. For example, in some cases we require interventions that improve some aspects of software production without affecting another, e.g. improved reliability with no increase in cost.*

Aspects of adaptation strategy: target, intervention degree, scalability impact, implementations, and underlying models. Study quality aspects will also be observed.

- Context

** For Software Engineering, this encompasses the context in which the comparison takes place (e.g. academia or industry), the participants taking part in the study (e.g. practitioners, academics, consultants, students), and the tasks being performed (e.g. small scale, large scale).*

Both academy and industry studies will be considered, with no restrictions regarding the participants or tasks scale.

1.2.3- Keywords and synonyms

Question keywords and synonyms/alternative spellings:

- Population: choreography, decentralized composition, decentralized service composition, distributed composition, distributed service composition, decentralized interacting services
- Intervention: {adapt*, reconfigur*}, {self-configur*, auto-configur*, self-healing}, {custom*}
- Comparison: <None>
- Outcomes: <None>
- Context: <None>

Strategy, tools and influencing factors have not been considered as keywords as they would unnecessarily restrict the returned results. This information will be captured from the studies that fit inclusion/exclusion criteria (see session 3.1). Analogously, no words from “Outcomes” were selected.

Keywords wildcard expansion:

- adapt* -> adaptable, adapting, adaptation
- reconfigur* -> reconfigurable, reconfiguring, reconfiguration
- configur* -> configurable, configuring, configuration
- custom* -> customizable, customizing, customization

2- Sources Selection

2.1- Sources selection criteria

The sources will be selected based on the following criteria: broad coverage in software engineering and distributed systems areas, full text availability, boolean operators support for query construction, result export capability, and academic perceived quality of content.

2.2- Studies languages

All studies written in English will be considered, as this is the internationally accepted language for writing and publishing scientific work. Besides, we believe that cultural factors don’t influence the results in this particular study.

2.3- Sources identification

2.3.1- Sources search methods

- Search through web search engines

2.3.2- Search string

The strategy employed to construct the search string is derived from [2]:

1. A structured question SQ in the form of a PICOC statement was derived from research questions Q1 and Q2 (see Section 3.1).
2. The identified synonyms and alternative spellings for each of the SQ keywords (see Section 1.2.3) were linked using the boolean operator OR. As *comparison*, *outcome* and *context* had no associated keywords, this was done only for *population* and *intervention*.
3. The two OR lists obtained from previous step were linked using the boolean operator AND.

Search String:

```
(
        //Population
        (choreography OR "decentralized composition" OR "decentralized service composition"
        OR "distributed composition" OR "distributed service composition" OR "decentralized
        interacting services")

        AND

        //Intervention
        (
        (custom*)
        OR
        ("self-configurable" OR "self-configuring" OR "self-configuration" OR
        "auto-configurable" OR "auto-configuring" OR "auto-configuration" OR
        "self-healing")
        OR
        (adapt* OR reconfig*)
        )
)
```

***=truncation**

2.3.3- Sources list

According to the criteria defined on section 2.1, as well as the suggestions given by [2] and review authors experience, the following initial list of sources was derived:

- IEEE Xplore
  - http://ieeexplore.ieee.org
- ACM Digital Library
  - http://portal.acm.org/dl.cfm?coll=portal
- CiteSeerX
  - http://citeseerx.ist.psu.edu
- SpringerLink:
  - http://www.springerlink.com
- SciVerse Scopus:
  - http://www.scopus.com/home.url
- SciVerse ScienceDirect:
  - http://www.sciencedirect.com/science
- Web of Science (ISI Web of Knowledge):
  - http://isiknowledge.com/?DestApp=WOS

2.4- Sources evaluation

- IEEE Xplore, Scopus, ScienceDirect and Web of Science: These sources fit all established criteria and will be selected. In particular, Scopus claims to be the largest database of abstracts and citations [2].

- ACM Digital Library and CiteSeerX: These sources present difficulties in exporting the results. However, given the quality and the amount of content available in such databases, they will be selected.

- Springer: This search engine presents crucial limitations, such as a limit for the query size. Moreover, IME-USP account on Springerlink does not have access to export facilities and it is not eligible to download some articles and books. As a result, this source will not be selected.

Therefore, the list of selected sources comprises: IEEE Xplore, ACM Digital Library, Scopus, ScienceDirect, Web of Science (Web of Knowledge), CiteSeerX.

2.5- Sources list checking by experts

The following systematic review experts evaluated the list obtained from the previous subsection and they agreed the selection:

- Igor Steinmacher: MSc professor with vast experience in conducting systematic reviews on Software Engineering.
- Marco Aurélio Gerosa: Assistant Professor of the Department of Computer Science at IME-USP.

3- Studies Selection

3.1- Studies definition

3.1.1- Inclusion and Exclusion Criteria

The defined criteria based on the research questions are:

i) Papers must be available to download
ii) Papers must propose or discuss a strategy for the choreography adaptation problem.
iii) Reconfiguration must be motivated from one of the following:
- Modifications in the existing choreography description (WSCI, WS-CDL, custom UML, BPMN2, etc);
- Modifications in the internal processes (e.g. orchestrations) that affect/disturb the choreography model (specification)

3.1.2- Procedures for study selection

Procedures for study selection in search engines (derived from [10]):

1. Query strings will be built according to the specific syntax of each selected source (see sections 2.3 and 2.4) and will be submitted*. Results from all sources will then be grouped in a single spreadsheet.
2. Duplicated and invalid results will be excluded.
3. All clearly irrelevant results will be discarded, i.e. papers that neither address any aspect of the research questions nor relate to Distributed Systems (or even Computer Science
4. The title of each remaining study will be read and a new selection will be done based on inclusion/exclusion criteria.
5. The abstract of every preselected work from the previous stage will be read and another new selection will be done based on inclusion/exclusion criteria. As suggested in [2], if reading the abstract is not sufficient to clearly understand the objectives or the problems being addressed, the review authors will also download the full article and check the study conclusions.
6. In case multiple versions of a study exist, only the most complete version will be included.
7. Finally, the selected studies will be fully read by at least one of the review authors, who will be responsible for writing a structured abstract (executive summary) of the study based on an already defined template.

Important notes:

- An inclusive approach will be carried out for the initial selection (stages 3 to 5), i.e. for one article to be preselected , it is sufficient that at least one of the authors decide in favor of the inclusion. The Kappa coefficient of agreement (which corrects for chance agreement [9]) will also be computed for this initial selection.
- All rejected studies will be adequately grouped in rejection categories (except for irrelevant studies)
- All stages of the whole selection process will be supervised by a more experienced researcher.

** For each search engine, a form containing the following fields will be elaborated: search engine name, query string, date of query submission, number of results and number of distinct and valid results.*

3.1.3- Primary Studies Quality Assessment

The following questions will be answered with Yes/No/Partially, corresponding to scores 1.0, 0.0 and 0.5 respectively.

Study quality assessment		**Score**
Q1	Is there a clear description of the strategy objectives?	
Q2	Is there a description of the applicability context and pre-conditions of the employed strategy? ● Applicability context and pre-conditions cannot be identified (score 0.0) ● Applicability context or pre-conditions are informally	

	discussed or can be implicitly inferred somehow (score 0.5) ● Applicability context and pre-conditions are explicitly described (score 1.0)	
Q3	Is there a description of the limitations and drawbacks of the employed strategy? ● Limitations and drawbacks are not explicitly described (score 0.0) ● Limitations are drawbacks are explicitly described (score 1.0)	
Q4	How is the strategy evaluated? ● No evaluation (score 0.0) ● Example of usage/Proof of Concept (score 0.5) ● Experiment/study case/formal proof (score 1.0)	
Q5	Is there a discussion on the strategy scalability? ● No discussion (score 0.0) ● Informal discussion (score 0.5) ● Discussion through formal proof or experiments (score 1.0)	
Q6	Does the strategy support dynamic service reconfiguration (including/removing services at runtime for the sake of adaptability ● No support (score 0.0) ● Partial support (score 0.5) ● Support (score 1.0)	
Q7	Is the strategy implemented in a tool? ● No (score 0.0) ● Yes (score 1.0)	

3.1.4- Structured Abstract Template

A structured abstract with the following topics will be prepared for each primary study: source; paper title; paper type {journal article, conference paper, short conference paper, workshop paper, technical report, PhD thesis}; authors; year; vehicle; paper abstract; research question/issue; choreography adaption strategy description; human intervention degree {manual, automatic, hybrid}; description of the strategy implementation; strategy limitations and drawbacks; study results/conclusion; study assessment; implementation availability; implementation license; additional notes.

3.1.5- Data Analysis and Summarizing

The data extracted from the studies will be tabulated and plotted in order to present basic information about each study. Also, the studies will be cohesively grouped according to the strategies employed and each category will be discussed. Sensitivity analysis will be performed to reason about result robustness, i.e. investigate if there were uncertainties about including or excluding certain studies.

Meta-analysis - the statistical analysis of a large collection of analysis results from individual studies intending to integrate the findings - will also be conducted to help answer the research questions and identify any interesting trends or limitations in current researches. Some basic measures that will be employed in the results summarizing include, but are not limited to*:

- Number of selected studies in each step of the selection process
- Number of selected studies per paper type (conference paper, journal article, technical report, PhD Thesis)
- Number of selected studies per year.
- Number of selected studies by journal/conference/report.

As a final point, we will state concluding comments about the systematic review results:

- Search, selection and extraction bias.
- Publication bias.
- Results application
- Recommendations

** Other descriptive statistics techniques will be employed on demand*

Appendix A - Introduction to Systematic Review in Software Engineering

The term “Systematic Review” refers to a specific methodology of research, developed to gather and evaluate the available evidence concerning a focused topic [Biolchini *et al.*, 2005]. In contrast to the usual process of literature review (non-systematically conducted whenever one starts a particular investigation), a systematic review is developed, as the term denotes, in a systematic way. This means that the research conduction process of a systematic type of review follows a very well defined and strict sequence of methodological steps, according to an elaborated protocol (see Figure 1). This instrument is built upon a central issue or topic that represents the core of the research, and is expressed by using specific concepts and terms. These concepts and term must be addressed towards information related to a specific, predefined, focused and structured question. The methodological steps, the strategies to retrieve the evidence and the focus of the question are explicitly defined, so that other professionals can reproduce the same protocol and be able to judge the adequacy of the chosen standards for the case.

The type of acceptable evidence to be gathered in a systematic review is stated beforehand. In the case of quantitative studies, the evidence data are often normalized to make results from different studies comparable in terms of effect magnitude, even when the studies are presented in diverse ways. Besides comparing results of individual studies, different kinds of syntheses can be done. In particular, meta-analysis is a type of research synthesis where the original individual studies are treated as if they were parts of a larger study, by having their data pooled together in a single and final result that summarizes the whole evidence. By selecting studies that are compatible in their quality level, and by taking strict care of their specific details, this methodological procedure can produce evidence as well as reveal aspects that the original studies are not individually able to elucidate. For instance, meta-analysis may prove that the results are statistically significant when small studies give inconclusive results with large confidence intervals. Besides, when conflicting results arise from different individual studies, meta-analysis may reconcile the data in a synthetic result.

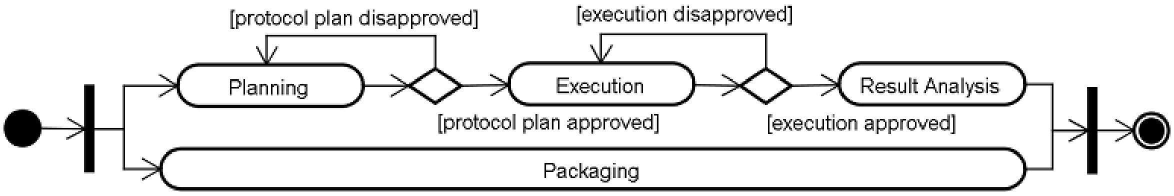

Figure 1: *The systematic review process [Biolchini et al., 2005]*

More information regarding core concepts of systematic review methodology can be found on [1], [2], [3], [4]. In particular, the first two references present such concepts from Software Engineering viewpoint. According to Biolchini *et al.*, applying systematic reviews in software engineering is much more difficult than in other areas, greatly due to the lack of rigor and conscience in reporting the results of the primary studies [1]. It is also hard to make comparisons when we do not have quantitative data, and the lack of standardization on the form of presenting the results is also a difficulty. According to Richard Hamming [12], "perhaps the central problem we face in all of computer science is how we are to get to the situation where we build on top of the work of others rather than redoing so much of it in a trivially different way. Science is supposed to be cumulative, not almost endless duplication of the same kind of things".

References

[1] J. Biolchini, P. Gomes, A. Cruz, and G. Orta, *”Systematic Review in Software Engineering”,* TECHNICAL REPORT. RT – ES 679 / 05. System Engineering and Computer Science Department COPPE/UFRJ, 2005.

[2] Kitchenham, B., & Charters, S. *“Guidelines for performing Systematic Literature reviews in Software Engineering”* version 2.3 EBSE Technical Report. Keele University and University of Durham. 2007.

[3] Petticrew, Mark and Helen Roberts. *“Systematic Reviews in the Social Sciences: A Practical Guide, Blackwell Publishing”*, 2005, ISBN 1405121106.

[4] B. a Kitchenham, S.L. Pfleeger, L.M. Pickard, P.W. Jones, D.C. Hoaglin, K. El Emam, and J. Rosenberg, *“Preliminary guidelines for empirical research in software engineering,”* IEEE Transactions on Software Engineering, vol. 28, Aug. 2002, pp. 721-734.

[5] Beecham S, Baddoo N, Hall T, Robinson H, Sharp H. Motivation in Software Engineering: A systematic literature review. *Information and Software Technology*. 2008;50(9-10):860-878. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0950584907001097.

[6] Qiao Xiaoqiang, Wei Jun, "A Decentralized Services Choreography Approach for Business Collaboration," scc, pp.190-197, IEEE International Conference on Services Computing (SCC'06), 2006.

[7] Barker A, Walton CD, Robertson D. Choreographing Web Services. *IEEE Transactions on Services Computing*. 2009;2(2):152-166.

[8] Staples, M. and Niazi, M. Experiences using systematic review guidelines. Article available online, JSS.

[9] J. Cohen, A coefficient of agreement for nominal scales, Educational and Psychological Measurement 20 (1960) 37–46.

[10] T. Dyba, T. Dingsøyr, Empirical studies of agile software development: A systematic review, Inform. Softw. Technol. (2008), doi:10.1016/j.infsof.2008.01.006 7

[11] J.P.T. Higgins, S. Green (Eds.), Cochrane Handbook for Systematic Reviews of Interventions, Version 5.0.0 (updated February 2008), The Cochrane Collaboration, 2008. Available from: <www.cochrane-handbook.org>.

[12] Richard Hamming 1968 Turning Award Lecture