Exploring Complexity Metrics for Artifact-Centric Business Process Models

Mike A. Marin[1,2]

[1] University of South Africa
330 Preller St, Muckleneuk, Pretoria, 0002, South Africa
[2] IBM Corporation, Hybrid Cloud
3565 Harbor Blvd, Costa Mesa, CA 92626, U.S.A
mmarin@acm.org

Abstract. This study explores complexity metrics for business artifact process models described by Case Management Model and Notation (CMMN). Over the last few decades business artifacts with Guard-Stage-Milestone (GSM) has emerged in the Business Process Management (BPM) literature as a new way of describing data intensive business processes. The work on GSM influenced the CMMN standard, which was created to fill a market need for more flexible case management processes for knowledge workers.
Complexity metrics have been developed for traditional BPM models, such as the Business Process Model and Notation (BPMN). However, traditional BPM is not suitable for describing GSM or CMMN process models. Therefore, complexity metrics developed for traditional process models may not be applicable to business artifact process models such as CMMN. This research addresses this gap and advances literature in the areas of method complexity, complexity metrics for process models, declarative processes, and research on CMMN by characterizing CMMN method complexity, identifying complexity metrics for CMMN, and exploring the relationship between CMMN and GSM.

Keywords: Business Artifacts · Business Process Management · BPM · Business Process Model · Case Management · Case Management Model and Notation · CMMN · Guard-Stage-Milestone · GSM · Complexity Metric · Process Model Complexity · Method Complexity

1 Introduction

The purpose of this study was to explore complexity metrics for artifact-centric business process models. Since 1996, after Daneva et al. [2] published the first report on business process model complexity, research on the topic has focused on imperative process models. Opportunities for research into data-centric processes were opened with the introduction of case handling by van der Aalst and Berens [7]. In addition, the introduction of Business Artifacts (BA) by Nigam and Caswell [5] expanded data-centric research opportunities into declarative processes.

The publication of Case Management Model and Notation (CMMN) [6] by the Object Management Group (OMG) in 2014 introduced a standard specification for declarative processes. However, complexity metrics for declarative process models have not as of yet been studied. To close this gap, this study explores complexity metrics for the artifact-based declarative case management modeling standard, CMMN.

1.1 Motivation

Extensive research has been conducted on the complexity metrics of imperative process modeling languages. Between 1996 and 2015, starting with the work produced by Daneva et al. [2] and ending with that done by Kluza [3], more than a hundred metrics for imperative process modeling languages were proposed. However, as far as this researcher is aware, no research has been done on the complexity metrics for declarative process model languages.

With the introduction of the CMMN standard based on GSM the use of declarative processes is becoming mainstream, and the need for complexity metrics will increase. Unless research on complexity metrics for declarative processes is advanced our knowledge of process technology complexity will be incomplete. A better understanding of complexity metrics for declarative processes will improve our understanding of these processes, and will help vendors and users of the new CMMN standard.

1.2 Objectives and Contributions

The main goal of this study were to identify and validate complexity metrics for artifacts based process models. The study addressed the following objectives:

1. To formalize CMMN as the basis to identify metrics.
2. To formalize the relationship between CMMN and GSM.
3. To assess the method complexity of CMMN.
4. To analyze the applicability of Business Process Management (BPM) complexity metrics to CMMN.
5. To identify complexity metrics for CMMN.
6. To validate the identified complexity metrics.

This research makes contributions in the areas of formalizing CMMN using first-order logic, proposing bidirectional transformations between CMMN case types and GSM artifact types, proposing complexity metrics for CMMN, identifying and comparing the model complexity of CMMN against other process notations, and clarifying the relationship between CMMN and GSM.

2 Method Complexity and Transformations

The calculated method complexity of CMMN seems to indicate that it is more complex than Event-driven Process Chain (EPC), but less complex than Business

Process Management and Notation (BPMN). This may also indicate that BPMN is more expressive than CMMN. The results are encouraging as they may indicate that CMMN should be easier to learn than BPMN.

This study also contributes formal transformations between CMMN case types and GSM artifact types to the broader BA knowledge base. The transformation from a GSM artifact type to a CMMN case type is relatively straightforward and simple to describe. The resulting case type modeled using CMMN is visually similar to the original artifact type, allowing for it to be easily understood by human beings. The transformation of a case type into an artifact type is more complex because CMMN extends GSM by introducing new constructs and defining a life cycle with a set of standard events for those constructs.

3 Systematic Literature Review

A systematic literature review (SLR) was conducted. The review was designed to identify complexity metrics for process models that have been proposed in the last 20 years (from January 1996 to June 2016 inclusive), and how these were validated. The goal of this review was to identify complexity metrics for process models with the specific purpose of identifying metrics that could be relevant to CMMN. In addition, it was conducted in order to identify the research methods, present in the literature about process models, used to validate complexity metrics that could be adapted to validate CMMN complexity metrics. The principal findings of this review were follows:

- Large number of proposed metrics without validation.
- No complexity metrics for declarative process models have been proposed.
- There is no agreement on research methods for empirical validation.
- Some consensus on theoretical validation is emerging.
- Insufficient guidance on sample size for empirical validation was uncovered.
- Low rigor and relevance of most empirical validations.

4 Metrics Validation

The validation of complexity metrics for business processes requires both theoretical and empirical validation. This study defined a set of metrics for CMMN. The metrics were theoretically validated using the formal framework for software measurements as defined by Briand et al. [1] and the complexity metrics were further validated using Weyuker's properties [8] for software complexity measures.

This research attempted an approach to empirically validate the metrics based on pairwise comparisons. This study departs from most empirical validations of complexity metrics for process models in several important ways, including the use of power calculations during the experimental design phase, the study targets professional process modelers instead of student, and it uses a within-subjects pairwise comparison experimental design.

4.1 Methodology

The objective of the empirical validation was to compare the calculated complexity metrics against human perceived complexity. This research adopted a quantitative approach using an online survey [4]. The research question was: does calculated complexity correlate with human perceived complexity?

This research hypothesized a positive correlation between calculated complexity and perceived complexity, a negative correlation between calculated complexity and model comprehension, and a negative correlation between model comprehension and perceived complexity.

4.2 Experimental Design

The research methodology followed a within-subjects experimental design. The subjects consisted of a convenient sample of professional process modelers sourced through a set of online forums and mailing lists.

The experiment was implemented using an online survey and tutorial. A short tutorial[3] was included, because the target population might not have been familiar with CMMN. The experiment used a repeated measures design with counterbalancing where each subject was exposed to only two models.

4.3 Findings

Although careful planning, sound experimental design, and efforts to minimize threats to validity were in place, the experiment experienced some problems. In retrospect, the experiment was too complex and required too many subjects to be conducted using an unsupervised online survey.

No evidence was found to support any of the four complexity metrics. This experiment failed to reject the five null hypotheses. Therefore, more research is required to re-test those hypotheses with smaller and more targeted experiments.

Evidence that a metric that was not included in this experiment, for methodological reasons, may predict CMMN model complexity was uncovered. That particular metric was kept constant at 90. The results obtained (inability to reject the five null hypotheses) are consistent with all of the models being equally complex. A post-hoc test designed to identify if subjects found all of the models equally complex, seemed to indicate that there is evidence to support this new hypothesis. However, more research will be required to test this new hypothesis.

5 Conclusion

The purpose of this research was to explore complexity metrics for artifact-centric business process models. The results of the SLR that was conducted for this thesis confirmed that to date, all of the proposed complexity metrics for process models have been for imperative process models rather than for artifact-centric business

[3] Link to CMMN Tutorial at http://cmmn.byethost4.com

process models. This research was based on the assumption that complexity metrics for declarative process models can be derived from the research that has been conducted on complexity metrics for imperative process models. This thesis focused on CMMN because of the potential practical impact that such research may have on vendors and users of this emerging standard.

The goal was to use CMMN as a proxy to fill the gap in the literature concerning complexity metrics for declarative process models. As such, this research contributes new knowledge to the literature by exploring complexity metrics and user comprehension of CMMN models. It attempted to theorically and empirically validate the CMMN metrics. It contributed a novel approach to the empirical validation of complexity metrics using pairwise comparisons. This research also makes contributions in the areas of formalizing CMMN by using first-order logic, identifying and comparing the model complexity of CMMN against other process modeling notations, clarifying the relationship between CMMN and GSM, and proposing complexity metrics for CMMN.

Bibliography

[1] L. C. Briand, S. Morasca, and V. R. Basili. "Property-based Software Engineering Measurement". In: *IEEE Transactions on Software Engineering* 22.1 (1996), pp. 68–86. DOI: 10.1109/32.481535.

[2] M. Daneva, R. Heib, and A. W. Scheer. *Benchmarking Business Process Models*. Research Report 136. Institute for Information Systems, University of the Saarland, 1996.

[3] K. Kluza. "Measuring Complexity of Business Process Models Integrated with Rules". In: *Artificial Intelligence and Soft Computing: 14th International Conference, ICAISC 2015, Zakopane, Poland, June 14-18, 2015, Proceedings, Part II*. Ed. by L. Rutkowski et al. Springer International Publishing, 2015, pp. 649–659. DOI: 10.1007/978-3-319-19369-4_57.

[4] M. A. Marin. *Research data, sources and documents for thesis on Exploring Complexity Metrics for Artifact-Centric Business Process Models*. Dataset on Zenodo. 2018. DOI: 10.5281/zenodo.1241595.

[5] A. Nigam and N. S. Caswell. "Business Artifacts: An Approach to Operational Specification". In: *IBM Systems Journal* 42.3 (2003), pp. 428–445.

[6] OMG. *Case Management Model and Notation, version 1.0*. Standard. Document formal/2014-05-05. OMG, 2014, p. 96.

[7] W. M. P. van der Aalst and P. J. S. Berens. "Beyond Workflow Management: Product-Driven Case Handling". In: *Proceedings of the 2001 International ACM SIGGROUP*. New York: ACM Press, 2001, pp. 42–51. DOI: 10.1145/500286.500296.

[8] E. J. Weyuker. "Evaluating Software Complexity Measures". In: *IEEE Transactions on Software Engineering* 14.9 (1988), pp. 1357–1365. DOI: 10.1109/32.6178.