EVALUATION OF BPMN MODELS QUALITY
A Family of Experiments

Elvira Rolón
Autonomous University of Tamaulipas, Centro Universitario Tampico-Madero, 89336, México

Félix García, Francisco Ruiz, Mario Piattini
Department of Information Technologies and Systems, Indra-UCLM Research and Development Institute
University of Castilla-La Mancha, Paseo de la Universidad 4,13071, Ciudad Real, Spain

Corrado Aaron Visaggio, Gerardo Canfora
RCOST – Research Centre on Software Technology, University of Sannio
Palazzo ex Poste, Viale Traiano, 82100 Benevento, Italy

Keywords: Business process models, BPMN, Empirical validation, Family of experiments, Measures.

Abstract: The design phase is of special importance in the development of a business process. This phase refers to the modeling, handling and redesigning of processes, but when maintenance tasks have to be performed, this stage may be rather complicated. It implies a heavy investment of time and resources, since it involves both technical developers and business analysts. Moreover, process modeling should permit not only the production of models which are understandable to the users, but also the early detection and correction of errors. All of this adds to the overall quality of the model. We therefore propose a set of measures with which to assess the structural complexity of conceptual business process models. Our aim is to obtain useful indicators to be used when carrying out maintenance tasks on these models, thus obtaining higher quality models by means of an early evaluation of the model’s given quality properties. With the development of a family of experiments, it has been possible to discover a set of measures which may be useful in assessing the usability and maintainability of conceptual business process models.

1 INTRODUCTION

The objectives of a business process are basically (Multamäki, 2002): a) To improve the understanding of a situation so that it can then be communicated to and among the different stakeholders and b) to use that process as a tool to attain the goals of a process development project. Nevertheless, for business processes to be able to comply with their objective, they are constantly exposed to changes. These come about as a result of organizations’ continuous improvement programmes.

Business process modeling consists of the description and visualization of processes by means of a model which represents them in formal or informal ways or in the form of a graph or diagram. Likewise, the manipulation and redesign process is carried out in the design phase (Smith and Fingar, 2003). Business processes modeling is therefore one of the first steps towards achieving organizational goals. It is an activity which has gained great importance due to the fact that today’s organizations are ever-more focused on their business process (Andersson et al., 2005).

The importance that business process modeling represents has been the springboard for a variety of studies such as that of Bandara (Bandara et al., 2005) in which the authors attempt to identify process modeling success factors and measures. Their empirical evidence comes from the case studies of nine process modeling projects.

Furthermore, business process modeling is of interest in a number of different fields such as that of business and software engineering. This is because its importance lies not only in the description of the
process, but in that it also usually represents a preparatory phase for activities such as (Succi et al., 2000): business process improvement, business process reengineering, technology transfer and process standardization.

High-quality conceptual-modeling plays an important role in carrying out business process reengineering in particular, making it possible to detect errors at an early stage and thus correct them (Wand and Weber, 2002). In addition to this, the analysis of the level of process maturity (Bider, 2005; Francis, 2005), also forces us to have bases which facilitate modeling in the design phase. This is also true as regards the work of future maintenance.

Bearing these factors in mind, and considering the lack of studies on the possible difficulty that business process models may represent in maintenance tasks, our work takes as its main focus the assessment of the structural complexity of business process models (BPMs) at a conceptual level. Our aim is to give support to business process management, allowing an early evaluation of certain quality properties of the models. It also makes the evolution of process models possible, providing as it does so objective information about maintainability, especially in those organizations which have given themselves over to ongoing improvement.

In this work we present the motivation of our research, basing it on how important it is to evaluate business process conceptual models if we are to aid their maintainability (Section 3). In addition, we present the results obtained in five experiments within the context of a family of experiments (Section 4 and 5). With these results we have attempted to obtain a set of measures which will serve as useful indicators towards the usability and maintainability of the BPMs. Finally, in Section 6 some of the conclusions drawn from this work will be put forward.

2 RELATED WORKS

The importance that the subject of business process and its modeling has acquired in the last few years has also generated great interest in the scientific community with respect to its study, analysis and measurement. However, very little can be found in literature as regards the measurement and assessment of business processes, at least at a conceptual level, which is the main topic of our research.

A recent work on measures of complexity for business process models is that presented by (Gruhn and Laue, 2006), in which the authors discuss how ideas that are already a familiar part of research into software complexity might be used to analyse the complexity of business process models.

On the other hand, the reference by (Cardoso, 2005) describes a measurement for analyzing the control-flow complexity of Web Processes and Workflows. This measurement is used at the time of design to evaluate the complexity of the design of a process before its implementation.

Having taken into consideration the studies made in the field of software engineering in (Rolón et al., 2006b), we have defined a set of measures for the evaluation of conceptual business process models on the basis of the adaptation and extension of a framework defined for the modeling and measurement of the software process. In (Cardoso et al., 2006) a similar type of compilation of insights from software engineering cognitive science and graph theory is provided, and the authors discuss to what extent analogous metrics of these areas can be defined for business. Finally, in (Latva-Koivisto, 2001) a collection of complexity measures for business process models found in the relevant literature was compared to a set of given criteria.

3 MEASURES FOR BUSINESS PROCESS MODELS

Our interest lies in evaluating the complexity of business processes by starting from the model which represents them at a conceptual level, and in order to do this we have used BPMN (OMG, 2006) as a modeling language. One of the reasons for the use of BPMN in our proposal is that it is one of the most widely recognized standard notations for the modeling of business processes and it is that which is most often used by both business analysts and systems analysts.

Moreover, a variety of business process modeling tools already use the BPMN metamodel and certain studies, such as that of (Mendling et al., 2005), show how BPMN, in comparison with another 14 specifications includes the 15 high-level metamodel concepts defined by the author, almost in their entirety. (Wohed et al., 2006) also provide a comprehensive evaluation of the capabilities of BPMN and its strengths and weaknesses when used for business process modeling. These studies, as is the case of other similar ones found in literature
(Havey, 2005), give us an indication of the importance of using this notation.

In order to attain objective knowledge of the external quality of business process models (BPMs) we have defined a set of measures with which to evaluate their structural complexity which is represented with BPMN. These measures have been placed in two categories:

- **Base measures.** These consist principally of counting the business process model’s significant elements, and a total of 46 base measures have been defined according to the main elements of which the BPMN metamodel is composed (activities, events, gateways, pools, etc).

- **Derived measures.** These have been defined from the base measures, and allow us to discover the proportions that exist between the different elements of the model. This group is made up of a total of 14 measures.

Some of the derived measures defined according to the base measures are shown in Table 1. A more detailed description of all the proposed measures appears in (Rolón et al., 2006b).

![Table 1: Derived Measures.](image)

With the defined base and derived measures, it is possible to evaluate the structural complexity of business process models developed with BPMN. When we analyse the model structurally, it is thus also possible for us to evaluate its internal quality.

The defined measures have been validated theoretically according to the Briand et al. theoretical framework (Briand et al., 1995). As a result, it has been possible to group them in relation to the various properties of structural complexity such as size, coupling and complexity, as regards internal quality (Figure 1).

![Figure 1: Relationship between structural complexity and quality attributes.](image)

Moreover, in line with our objective, which is that of discovering which of the defined measures may provide useful and objective information about the external quality of the BPMs, we will focus on two characteristics of the external quality of the ISO 9126: Usability and Maintainability. These will be evaluated by means of the following two sub-characteristics which are respectively:

- **Understandability.** The ease with which the model can be understood by the user.

- **Modifiability.** The ease with which the model can be modified, by possible errors, by requesting a specific modification or by new requirements.

In order to discover what measures may serve as useful indicators to evaluate the understandability and modifiability of the MPNs, a family of experiments has been carried out, which is described in the following sections.

4 EMPIRICAL VALIDATION – A FAMILY OF EXPERIMENTS

The family of experiments includes the development of 5 experiments which have been carried out in similar circumstances and in the same context.

The experimental design used was the same for all 5 experiments, since the second experiment was planned as a replica of the first, the purpose of this being to corroborate the results obtained, and the fourth experiment was similarly a replica of the third.

The variant of these experiments with respect to both of the first consists of some changes to the
experimental material in the MPNs, with the intention of confirming whether the measures validated in both first experiments might or might not be useful in evaluating the usability and maintainability of the MPNs. A detailed description of the experimental design can be consulted in (Rolón et al., 2006a).

Of the five experiments conducted, it is important to emphasize that the third was carried out with the Masters students of the University of Sannio in Italy, because this implied an additional effort, which was mainly that of considering the language as a barrier since the training session was given to them in English.

In addition, it took longer to carry out this experiment, since the training was more thorough with regard to business process modeling and BPMN modeling notation. Also, subjects such as Business Process Management and tools such as the BPMS (Business Process Management Systems) were mentioned.

4.1 Subjects

The participant subjects in all the experiments had similar knowledge as far as process modeling was concerned. All the groups were nevertheless given a training session to ensure that they were conscious of the aspects that we were attempting to evaluate. A summary of the groups who participated in each experiment can be see in Table 2.

![Table 2: Participant groups in the experiments.](image)

4.2 Material

In all cases the material consisted of ten BPMs represented with BPMN, whose structural characteristics and dimensions were different from each other; that is to say, models with different degrees of complexity were selected. These were obtained by varying the value of the measures in each model. Our intention, upon choosing models with different dimensions, was to determine the influence of the complexity of the model upon different subjects such as business analysts and software engineers, who are the main focus of our study.

Moreover, two questionnaires were formulated for each of the aforementioned models. The first consisted of a series of six questions related to the model’s understandability, and the second proposed a series of modifications to be carried out on the model, such as evaluating the complexity of the process models presented. In addition, at the end of each questionnaire a question was included, whereby the subjects were asked to assess the complexity of the models presented in a subjective manner. The material also included an example of a solution which showed how the exercises should be done. An example of the material used can be found in (Rolón et al., 2007).

4.3 Objective

Using the GQM template (Goal Question Metric) the goal in all the experiments is defined as being:

- To Analyse measures of BPM structural complexity
- To evaluate them as regards their capability of being used as indicators of business process model understandability and modifiability
- The researchers’ point of view
- The context of PhD students, research assistants and lecturers in software engineering (Exp. 1); Students of the Masters degree in Information Systems (Exp. 2); Post graduate students (Exp. 3); Administrative staff and health professionals (Exp. 4) and PhD Students (Exp. 5).

4.4 Variables and Hypothesis

Within the context of the family of experiments the same variables have been considered. These are:

- Independent variables: those which correspond with the proposed measures, which is to say the base measures and derived measures already described.

- Dependent variables: those which relate to the understandability and modifiability of the BPMs, which will be measured according to the subjects’ efficiency when performing the tasks, which is calculated as the ratio between the number of right answers and the time.

The dependent variables were measured through the subjects’ response times when carrying out the required tasks, the success rate in the questions relating to the understandability and modifiability
tasks of the model, the subjective evaluation with respect to the complexity of the models, and also the efficiency of the successes in relation to the times.

The hypotheses proposed with respect to the objective of our investigation are the following:

- Null hypothesis, \( H_0u \): There is no significant correlation between the structural complexity measures and the understandability.
- Alternative hypothesis, \( H_{1u} \): There is a significant correlation between the structural complexity measures and the understandability.
- Null hypothesis, \( H_{0m} \): There is no significant correlation between the structural complexity measures and the modifiability.
- Alternative hypothesis, \( H_{1m} \): There is a significant correlation between the structural complexity measures and the modifiability.

5 ANALYSIS OF RESULTS

Once the individual experiments had been carried out, a global analysis of the results took place within the context of the family of experiments in order to determine whether or not they had attained the general objective of the empirical evaluation. In order to do this, a descriptive analysis and a statistical analysis of the data collected in all five experiments were carried out. A description of both analyses is presented in the following sections.

5.1 Descriptive Analysis

Having taken into account that the dependant variables are those which are relative to the understandability and modifiability of the model, a summary of the data obtained from the results of each experiment was carried out.

Upon analyzing the time taken by the subjects to carry out the required tasks and upon obtaining the mean times of the five experiments, it can be observed in Table 3 that, in the case of the tasks relating to the model’s understandability, the subjects took more time with models 5, 7 and 10, whilst they took more time to carry out the requested modifications with models 3, 4 and 7.

These results can be better appreciated in Figure 2, in which the results that appear in Table 3 are grouped according to the average of the results obtained in each experiment in order to discover both which models the subjects took most time to respond to, and the models’ understandability and modifiability.

The descriptive analysis relating to success, subjective evaluation and efficiency was carried out in a similar manner. With regard to success in the required tasks, the results of the five experiments show that models 3, 4 and 7 were those which led to the subjects producing the greatest amount of errors in the responses related to understandability, whilst in the tasks related to modifiability the majority of mistakes were made with models 3, 7 and 10 (Fig.3).
As regards the subjects’ subjective evaluation of the complexity of understandability of the models presented, models 5, 6, 9 and 10 were evaluated as being the most complex, whilst in the case of modifiability the most complex models were 5, 7 and 10 (Figure 4).

Finally, the efficiency of the successful responses to the tasks in relation to the time taken to carry them out was obtained from the statistical analysis of the dependant variables.

The results of the correla tion analysis of the five experiments to discover understandability and modifiability times obtained that the measures which correlated with the response times for the tasks relating to understandability, and which were validated in at least two of the five experiments were: NIMsE (Number of Intermediate Message Events), NEDDB (Number of Exclusive Decision Data-Based), TNIE (Total Number of Intermediate Events of the model), NSFE (Number of Sequence Events).
Flows from Events) and TNE (Total Number of Events of the model).

With regard to modifiability, the NEDEB (Number of Exclusive Decision Event-Based) and CLA (Connectivity Level between Activities) measures were validated in experiments 2 and 3.

The analysis of the correlations with regard to successes, subjective evaluation and efficiency was carried out in a similar manner. With regard to the correlations of the measures as regards successes in the required tasks, only the TNSE (Total Number of Start Events of the model) was validated in two of the five experiments as far as successes in understandability were concerned. In the case of successes in the modifiability tasks, of the various correlation measurements the NDOIn (Number of Data Object-In of the process) and TNEE (Total Number of End Events of the model) measures were only validated in two experiments.

In the efficiency analysis (Table 4), the measures validated in at least two of the five experiments with regard to understandability were: NIMsE, NEMsE (Number of End Message Events), NEDDB, NSFE, TNE and NSFL. In relation to modifiability, the following measures were validated: NCS (Number of Collapsed Sub-processes), TNCS (Total Number of Collapsed Sub-processes of the model), NEDEB and CLA.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Understandability</th>
<th>Modifiability</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NIMsE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NEMsE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NEDDB</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NSFE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NCS</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNCS</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NEDDB</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NSFL</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NIM</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NEM</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NSFG</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNG</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNA</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNEE</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NT</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNT</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NSFG</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NDOIn</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>TNG</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>NSFL</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

Table 4: Efficiency correlations.

Finally, upon analyzing the correlations as regards the subjective evaluation that the subjects made of the complexity of the models, we obtained the result that measures which were validated in at least two of the five experiments only existed in the case of modifiability.

These measures were: TNE, TNA (Total Number of Activities), NENE (Number of End None Events), NT (Number of Tasks), NSFL, TNEE, TNT (Total Number of Tasks in the model), NEDDB, NSFG (Number of Sequence Flows from Gateways), and TNG (Total Number of Gateways of the model).

6 CONCLUSIONS AND FURTHER WORK

This work shows the results obtained from carrying out a family of experiments. This was done with the objective of analyzing and evaluating the structural complexity of business process models. The analysis took place at a conceptual level of the models and used a set of measures which were defined on the basis of the BPMN standard notation.

As a result of this family of experiments we have obtained a significant set of measures that could serve as indicators towards the maintainability of the business processes models expressed in BPMN. From the 60 defined measures, 22 have been correlated with BPMN model understandability or modifiability in at least one of the experiments. With these 22 measures, which have been validated as results in order to measure the efficiency in the accomplishment of the tasks (dependent variable), it is possible to reject the formulated null hypotheses.

Future experimentation will focus upon evaluating this set of measures which we consider relevant from the results obtained in this first family of experiments. To achieve this, new material will be designed in which the validated metrics will be grouped into three categories (participants and roles, activities and control flows and decision nodes). Specific material will be designed for each group. This will allow us to obtain a higher variation of the complexity in each subgroup of models which may confirm the usefulness of the measures validated in the first family. Other aspects to be tackled in the future are:

- We shall conduct new experiments with the aim of analyzing two further sub-characteristics of the quality of the model, namely analyzability and ease of learning, which are respectively related to usability and maintainability.
- Furthermore, the development of business process models within a company in the health sector is being carried out, which will allow us to use experimental models of real cases.
ACKNOWLEDGEMENTS

This work has been partially financed by the INGENIO Project (Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia, PAC08-0154-9262), MECENAS (Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia, PBI06-0024) and ESFINGE Project (Ministerio de Educación y Ciencia, Dirección General de Investigación/Fondos Europeos de Desarrollo Regional (FEDER), TIN2006-15175-C05-05.

REFERENCES


Mulamaki, M., 2002. Objective-driven planning of business process modeling, Department of Industrial Engineering and Management, Helsinki University of Technology.


