A SOFTWARE SOLUTION FOR QUALITY ASSESSMENT OF BUSINESS PROCESS MODELS

One of the primary tools of the BPM (Business Process Management) paradigm is business process modeling. Business scenarios can be presented in the form of graphical models that can be easily understood by both Information Technology (IT) professionals and non-IT professionals — business analysts, software customers, department heads, top managers, and other stakeholders interested in business process improvement — using business process modeling. Small and large businesses can capture their operations in the form of graphic diagrams using business process modeling as the fundamental tool of the BPM approach, which can then be brainstormed by business analysts to uncover ways to optimize organizational workflows. Typically, business process improvement is accomplished by the automation of operations that have been recognized as “bottlenecks” following analysis. However, analyzing a business process model is only viable if it is clear and correct in terms of compliance with both the notation used and the actual business process it reflects. Therefore, this work examines the structural measures of the BPMN (Business Process Model and Notation) business process model. It is assumed that business process models that violate business process modeling rules are neither understandable nor suitable for further work with them, which can also lead to various errors occurring during the stage of business process analysis, as well as the stage of its improvement and implementation of proposed changes, i.e., during development, testing, and maintenance of distinct software components, information system modules, and so on. The object of this work is a process of quality assessment of business process models created using the BPMN notation. The subject of this work is a software solution for the quality assessment of business process models. The goal of this work is the evaluation of a probability of errors in business process models to improve their quality. The study of how the error probability of business process models affects their understandability and modifiability is performed.

Keywords: software solution, quality assessment, business process model, BPMN notation.

Introduction

Business process modeling details the steps a business takes to complete a process, such as hiring an employee or ordering and shipping a product. These models are also called business process diagrams or business process flowcharts. Like other types of charts, these models use specific symbols to represent business activity [1].

Business process modeling can be used to document the current business process and model the new one. Its purpose is to get a detailed view of the process, people, inputs, controls, and outputs, and then potentially simplify it all, make it more efficient, and/or improve the results of the process. Business process modeling takes time and discipline, but over time the payoff can be significant. Business process modeling has become a common approach in the business world to standardize procedures, improve efficiency, meet audit requirements, and gain competitive advantage [1].

Business process modeling can be used to prepare for business audits or sales, to reduce costs, to plan for automation, to understand the impact of unfinished changes, to reconfigure related processes, and to measure and reconfigure the efforts of people involved in processes. Often, a business may think it understands its processes, but
then discovers twists and turns during a mapping initiative. When modeling a new business process, the mapping is sometimes referred to as business process modeling or BPM (Business Process Management) [1].

Business process modeling can range from simple hand-drawn diagrams to more complex ones with expandable sections to offer adequate implementation information. BPMN (Business Process Modeling Notation), at its most sophisticated, is carried out by qualified analysts. The Object Management Group (OMG) offers five BPMN 2.0 certifications known as OCEB 2, which stands for OMG-Certified Expert in BPM 2.0. One track is focused on business, while the other is focused on technology. BPMN 2.0, according to OMG, will unify business process modeling in the same manner that Unified Modeling Language (UML) standardized software modeling [2].

BPMN is a flowchart technique that models the stages of a planned business process from start to finish. As a key to business process management, it includes a detailed description of the sequence of business operations and information flows associated with the completion of a process [2]. BPMN business process modeling demands time and effort, but the payback in terms of comprehension and improvement can be enormous. Version 2.0 improves on prior versions by offering a more comprehensive standard set of symbols and notations, allowing for greater detail for those that require it [2].

The goal of Business Process Management is to develop a continuous improvement life cycle. Model, implement, execute, monitor, and optimize are the steps. BPMN diagrams play an important role in this [2]. The purpose of BPMN is to model ways to increase efficiency, take into account new circumstances, or gain a competitive advantage. In the past few years, this method has been standardized and is now often referred to by a slightly different name: business process model and notation, still using the acronym BPMN. Also, it is different from the Unified Modeling Language (UML) used in software development [2]. BPMN is designed to help participants and other stakeholders in a business process gain understanding through an easy-to-understand visual representation of the steps. At a more detailed level, it is aimed at the personnel who will implement the process, providing enough detail to ensure correct implementation. It establishes a standard, common language for all stakeholders, technical and non-technical alike: business analysts, process participants, managers, and technical developers, as well as external teams and consultants. It should ideally bridge the gap between process intention and implementation by giving enough depth and clarity into the sequence of business processes [2].

Understanding the basic symbols used in Business Process Model Notation is essential for learning how to read and build high-quality BPMN diagrams [3]. According to [4], the quality of conceptual business process models is critical for the design of associated information systems. A precise measurement of model properties, in particular, can be useful from a commercial standpoint, allowing for cost savings due to early error discovery. This is also true in terms of software engineering. Models help with stakeholder communication and software system design in the latter instance. From a correlational standpoint, research has looked into numerous proposals for measures for business process models. This is useful for understanding, for example, the general driving forces of error probability, such as size and complexity. However, design decisions must usually be based on thresholds that may reliably indicate that a specific counter-action must be done. Hence, the following recommendations and corresponding size measures with respective thresholds could be used to evaluate the quality of BPMN business process models [4]:

1) avoid inclusive gateways (names as OR-splits below);
2) use one start and one end event;
3) use as few elements in the model as possible.

Consequently, the [4] proposes the following thresholds and probabilities of finding errors in business process models. Despite the variety of suggested measures and corresponding thresholds, we propose to use the size measures (nodes, OR-splits, start-events, and end-events) because they will allow checking a BPMN model without extra calculations in a fast and easy manner.

Problem statement

Small and large organizations can use business process modeling as the primary tool of the BPM strategy to capture their operations in the form of graphic diagrams, which can then be brainstormed by business analysts to identify ways to optimize organizational workflows. Typically, business process improvement is performed by the automation of procedures identified as “bottlenecks” after analysis. Analyzing a business process model, on the other hand, is only feasible if it is clear and correct in terms of compliance with both the notation employed and the actual business process it represents. As a result, the structural measurements of the BPMN business process model are investigated in this paper. It is assumed that business process models that violate business process modeling rules are neither understandable nor suitable for further work with them, which can also lead to a variety of errors occurring during the stage of business process analysis, as well as the stage of its improvement and implementation of proposed changes, i.e., during the development, testing, and maintenance of distinct software components, information system modules, and so on.

The object of this work is a process of quality assessment of business process models created using the BPMN notation. The subject of this work is a software solution for the quality assessment of business process models. The goal of this work is the evaluation of a probability of errors in business process models to improve their quality. To achieve the goal, the following tasks should be solved:

1) existing software solutions for business process modeling should be analyzed;
2) the business process management lifecycle should be studied and an improved business process modeling activity taking into account the suggested software solution should be proposed;
3) the general system architecture design should be proposed, and the software application should be designed and developed;
4) the usage of a software solution should be demonstrated and the experiments to answer the following research question “how the error probability of business process models affects their understandability and modifiability?” should be performed.

**Analysis of the existing software solutions for business process modeling**

Let us make a brief review of existing software solutions for business process modeling [5]. In the list below we consider only the freeware or open-source software tools that can be used by any user without payment from students to experienced engineers. Comparison of reviewed business process modeling software solutions shows that not all diagramming tools are really suitable for BPMN business process design. Table 1 shows that only several of covered business process modeling tools support BPMN 2.0.

**Table 1**

<table>
<thead>
<tr>
<th>Modeling software tool</th>
<th>Web-based</th>
<th>Open-source</th>
<th>Process-centric</th>
<th>BPMN 2.0 support</th>
<th>BPM capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gliffy</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lucidchart</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Open ModelSphere</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ARIS Express</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Modelio</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>ProcessMaker</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Questetra BPM Suite</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>BizAgi Process Modeler</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ArgoUML</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Visual Paradigm</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Adonis</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Creately</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

The list of appropriate software tools includes ARIS Express, Modelio, ProcessMaker, Questetra BPM Suite, BizAgi Process Modeler, and Adonis. Moreover, among these software tools only ProcessMaker, Questetra BPM Suite, BizAgi Process Modeler, and Adonis are BPM-centric software tools that fits the most BPMN process modeling.

**Business process management lifecycle improvement**

![Fig. 1. The AS-IS decomposition diagram of the business process management lifecycle](image-url)
Let us describe the AS-IS (a current state before the software solution for quality assessment of business process models introduced) business process management lifecycle using the IDEF0 functional modeling methodology. The IDEF0 approach is intended to communicate and analyze the functional perspective of a system by modeling the decisions, actions, and activities of an organization or other system [6]. Fig. 1 below depicts the AS-IS decomposition diagram of the business process management lifecycle that includes design, modeling, execution, monitoring, and optimization activities.

Fig. 2 below demonstrates the IDEF3 workflow diagram of the business process modeling activity (Fig. 1).

![Fig. 2. The AS-IS workflow diagram of the business process modeling activity](image)

Process flow modeling (IDEF3) provides a framework for representing a series of events. It depicts the flow, interactions, and objects of a process graphically [7].

Fig. 3 below proposes the TO-BE workflow diagram of the business process modeling activity.

![Fig. 3. The TO-BE workflow diagram of the improved business process modeling activity](image)

According to Fig. 3, the new tasks were added to the AS-IS workflow diagram (Fig. 2). These new tasks assume checking of the created business process model and fixing detected modeling mistakes. Suggested changes made to the TO-BE workflow diagram of the business process modeling activity will affect the whole business process management lifecycle. Introduced tasks require the software solution for quality assessment of business process models to detect modeling errors using quality measures and respective thresholds [4].

Considered measures, thresholds, and error probabilities are the following [4]:

1) measure “OR-SPLITS”, threshold: 0.5, probability: 9%;
2) measure “START-EVENTS”, threshold: 2.5, probability: 7%;
3) measure “END-EVENTS”, threshold: 2.5, probability: 5%;
4) measure “NODES”, threshold: 31.5, probability: 9%.

Therefore, Fig. 4 below demonstrates the TO-BE decomposition diagram of the business process management lifecycle. It introduces the business process model quality analysis tool, quality measures, and thresholds to detect corresponding business process modeling errors.
According to the 3-tier system architecture design [8], the following UML deployment diagram (Fig. 5) demonstrates the nodes and main components of software architecture.

The system architecture (Fig. 5) includes the database server with a DBMS (Database Management System) and a database, the application server with an HTTP (HyperText Transfer Protocol) container and a web application, and the client device with a web browser and an HTML (HyperText Markup Language) web page.

The UML component diagram below (Fig. 6) demonstrates the detailed structure of software components deployed to the system nodes.
According to Fig. 6, the web application is a Java-based software that includes a Servlet to process HTTP requests from the HTML web page and sends responses back to the client. The web application also includes a service layer to implement business logic, a DAO (Data Access Object) layer to interact with the database schema using SQL (Structured Query Language) commands, and Java beans (plain classes with fields, getters, and setters) to reflect database tables.

The service layer of the web application uses a third-party Camunda Java library to work with the BPMN API (Application Programming Interface) [9].

The web page uses the additional JSP (Java Server Pages) markup to interact with the Servlet, CSS (Cascading Style Sheets) styles to make the web design more attractive, and JavaScript scripts to provide interactive elements to the web page.

According to the DB-Engines ranking of database management systems according to their popularity [10], the Oracle DBMS should be used to implement the software solution for the quality assessment of business process models.

The UML sequence diagram (Fig. 7) shows the general lifecycle of certain software elements within the context of the user’s interaction with these elements.
According to this sequence diagram (Fig. 7) users can access the login form to access the application or use the sign-up form to create an account before accessing the software capabilities. Then users can upload BPMN business process models using the analysis form where the measures are also calculated. Before obtaining the report, users can configure settings to define how the business process modeling measures will be interpreted toward the thresholds and error probability values (i.e. the probabilities of detecting errors in business process models). Moreover, the settings form can be opened simultaneously with the analysis form (e.g. as the modal window). Finally, users can save the generated reports with obtained recommendations.

**Quality assessment of business process models using the software solution**

According to [4], chosen quality measures, thresholds, and error probabilities are the following:

1) \( x_1 \) – is the measure “OR-SPLITS” with the threshold \( t_1 = 0.5 \) and probability \( p_1 = 0.09 \);
2) \( x_2 \) – is the measure “START-EVENTS” with the threshold \( t_2 = 2.5 \) and probability \( p_2 = 0.07 \);
3) \( x_3 \) – is the measure “END-EVENTS” with the threshold \( t_3 = 2.5 \) and probability \( p_3 = 0.05 \);
4) \( x_4 \) – is the measure “NODES” with the threshold \( t_4 = 31.5 \) and probability \( p_4 = 0.09 \).

Using the probabilities multiplication rule of independent events [11], it is proposed to calculate the reverse probabilities – of not finding errors in a business process model:

\[
q(x_i) = 1 - p(x_i), \quad i = 1, m, \quad (1)
\]

where:
- \( p(x_i) \) – is the probability of finding errors in a business process model if the measure \( x_i \) is greater than the respective threshold value \( t_i \);
- \( q(x_i) \) – is the probability of not finding errors in a business process model if the measure \( x_i \) is greater than the respective threshold value \( t_i \);
- \( m \) – is the number of considered quality measures.

Hence, the reversed probability of having errors in a business process model can be calculated using the following expression [12]:

\[
P(x_1x_2...x_m) = 1 - \prod_{i=1}^{m} q(x_i), \quad (2)
\]

The following screenshot (Fig. 8) shows the homepage of the developed software solution for the quality assessment of business process models. It demonstrates the list of BPMN business process models evaluated toward the quality measures compared to the thresholds and error probabilities.

![Fig. 8. The example of measured thresholds and the error probability](image)

Let us check the calculations and examine the first model (the BPMN 2.0 file has the name “Excercise_3_-_Subrogation_995343cb589ce438e9e0edd816860d54fbpmn”) included into the report in Fig. 8. This model (Fig. 9) and other evaluated models were taken from the open BPMN models repository created by the Camunda company [13].
This model violates the end events threshold since it has 3 end events. According to equations (1) and (2), the error probability of this model is 0.05 (5% chance of having errors).

**Research results and discussion**

Now having the software tool capable of BPMN business process model processing, it is possible to perform the necessary calculations to answer the research question mentioned in the problem statement section of this study – “how the error probability of business process models affects their understandability and modifiability?”. First of all, it is proposed to address the paper [14] where metrics and respective thresholds were proposed to measure business process models from the perspectives of understandability and modifiability. Thus, let us consider two general structural metrics – one for business process models’ understandability and the other – for modifiability.

According to [14], understandability could be estimated using several metrics, among which we chose the coefficient of connectivity:

\[
CNC(G) = \frac{|A|}{|N|} \tag{3}
\]

where:
- \(G\) – is the graph-based representation of a business process model, \(G = (N,A)\);
- \(A\) – is the set of arcs (i.e. sequence flows) that connect business process elements;
- \(N\) – is the set of business process elements (i.e. tasks, events, gateways, etc.).

Whereas among several metrics suggested in [15] for modifiability estimation, we chose the density:

\[
\Delta(G) = \frac{|A|}{|N|\cdot(|N|−1)} \tag{4}
\]

Consequently, the authors of [14] these metrics suggest the respective threshold values:
1) \(CNC(G) \leq 0.6\) – models with values greater than this threshold will consider rather inefficient from the understandability viewpoint;
2) \(\Delta(G) \leq 0.0013\) – models with values greater than this threshold will consider rather inefficient from the modifiability viewpoint.

Therefore, we can conclude that the higher values of metrics (3) and (4), the lower the probability of considering a business process model efficient. Hence, let us estimate the dependency between the error probability (2), coefficient of connectivity (3), and density (4) values using the Pearson correlation coefficient [15]. In addition to metrics (2), (3), and (4), the essential features of business process models, such as the number of sequence flows and the number of process elements, should be used to estimate the correlation.

Table 2 demonstrates the estimated correlation between the mentioned metrics.

To perform these calculations we used the Business Process Management Academic Initiative (BPMAI) [16] dataset of business process model features that includes descriptions of 18813 BPMN models. Interestingly, that obtained Pearson’s values (Table 2) demonstrate a low correlation (below 0.3 [17]) between the error probability and coefficient of connectivity (0.20), as well as a low anti-correlation between the error probability and density (−0.31).

However, the probability of errors is moderately correlated (between 0.5 and 0.7) [17] with the number of sequence flows (0.67) and process elements (0.71). Hence, we can conclude that large business process models are more likely to have more errors. Also, since the numbers of sequence flows and process elements belong to understandability metrics [14], which growth also negatively affects the business process models’ understandability, we can assume that the higher the error probability of a business process model, the less understandable it is.
Conclusions

In this paper we proposed the software solution for quality assessment of BPMN business process models. It was considered, that small companies and large enterprises can use business process modeling as a key element of the BPM strategy to capture their operations in the form of graphic diagrams, which can then be brainstormed by business analysts to identify ways to enhance organizational workflows.

Therefore, we assumed that business process models that violate business process modeling rules are unintelligible and unsuitable for further work, which can lead to a variety of errors during the stages of business process analysis, improvement, and implementation of proposed modifications, i.e. during the development, testing, and maintenance of various software components, information system components, etc. Hence, this study investigated the structural measures of BPMN business process models.

Therefore, to evaluate the probability of errors in business process models necessary to improve their quality, the following tasks were solved:

1) existing software solutions for business process modeling were analyzed – ProcessMaker, Questetra BPM Suite, BizAgil Process Modeler, and Adonis BPM-centric software tools were defined as the most suitable for BPMN business process modeling;
2) the business process management lifecycle was studied and an improved business process modeling activity was proposed – it depends on the developed software solution for quality assessment of business process models to detect modeling errors using quality measures and respective thresholds [4];
3) the general system architecture design was proposed, as well as the database and the software application were designed and developed – it is a Java-based 3-tier client-server web application that uses the Oracle database;
4) BPMN business process models were evaluated toward the quality measures and compared to the thresholds to calculate the probability of errors;
5) performed experiments demonstrate that business process models with high error probability values are less understandable, while the low anti-correlation between the error probability and density may be explained by the fact that business process models usually do not contain complex structures, which are typical for the adequate diagrams, so it is easier to modify less complex diagrams even if they are fault.

Future work in this field may include the elaboration of the mathematical apparatus on which the proposed software solution is based, by introducing new structural measures of business process models, advanced prediction techniques, and intelligent recommending methods for business process model quality improvement.

References


| Correlation values | Error probability, $P(x_1, x_2, ..., x_m)$ | Coefficient of connectivity, $CN(G)$ | Density, $\Delta(G)$ | Number of sequence flows, $|A|$ | Number of process elements, $|N|$ |
|---------------------|------------------------------------------|----------------------------------------|----------------------|-----------------|-----------------|
| Error probability, $P(x_1, x_2, ..., x_m)$ | 1.00 | 0.20 | -0.31 | 0.67 | 0.71 |
| Coefficient of connectivity, $CN(G)$ | 0.20 | 1.00 | 0.16 | 0.44 | 0.32 |
| Density, $\Delta(G)$ | -0.31 | 0.16 | 1.00 | -0.33 | -0.38 |
| Number of sequence flows, $|A|$ | 0.67 | 0.44 | -0.33 | 1.00 | 0.96 |
| Number of process elements, $|N|$ | 0.71 | 0.32 | -0.38 | 0.96 | 1.00 |
Andrii Kopp
Andriй Конь
PhD, Associate Professor of Software Engineering and Management Intelligent Technologies Department, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine, e-mail: kopp93@gmail.com, https://orcid.org/0000-0002-3189-5623
Доктор філософії, доцент кафедри програмної інженерії та інтелектуальних технологій управління, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна.

Dmytro Orlovskyi
Дмітро Орловський
PhD, Associate Professor of Software Engineering and Management Intelligent Technologies Department, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine, e-mail: orlovskyi.dm@gmail.com, https://orcid.org/0000-0002-8261-2988
Кандидат технічних наук, доцент, доцент кафедри програмної інженерії та інтелектуальних технологій управління, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна.

Ilyass El Arbaouti
Ель Арбауті Ілясс
Bachelor, Graduate of Software Engineering and Management Intelligent Technologies Department, National Technical University “Kharkiv Polytechnic Institute”, Kharkiv, Ukraine, e-mail: mrfixer59@gmail.com
Бакалавр, випускник кафедри програмної інженерії та інтелектуальних технологій управління, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна.