ABSTRACT
Business process management (BPM) is a relatively mature discipline with a large number of practitioners. BPM technology is based on directed graphs used to describe control flow. Over the last decade a new way to describe data-intensive business processes for knowledge workers has emerged, where the focus is changing from control flow to business data. This emerging process technology is referred in the literature as case handling or case management. The Case management modeling and notation (CMMN) is a new process standard that supports case handling. Complexity metrics have been developed and verified for BPM models based on directed graphs, like the business processes modeling and notation (BPMN). But, CMMN processes are not described by directed graphs. Therefore, complexity metrics developed for BPMN may not be applicable to CMMN. This research plans to fill the gap in the literature by exploring and proposing complexity metrics for CMMN. This will have practical implications for commercial products implementing CMMN.

Categories and Subject Descriptors
D.2.2 [Design Tools and Techniques]; D.2.3 [Coding Tools and Techniques]: Standards; D.10 [Design]: Representation

General Terms
Measurement, Design, Human Factors, Standardization, Theory

Keywords
Adaptive case management, Case management, Case handling, CMMN, BPMN, Modeling complexity, Complexity metrics, Process modeling complexity

1. INTRODUCTION
This study will explore complexity metrics for the Case Management Modeling Notation (CMMN) [14], which is an emerging business processes standard. As far as the researcher is aware, this study will be among the first contributions to the understanding of complexity metrics for the CMMN standard.

The purpose of the research is to develop and validate complexity metrics for CMMN. The study will be theoretical and experimental in design and it will be conducted in two phases. The first phase will identify a set of complexity metrics and theoretically validate them. The second phase will empirically validate the proposed complexity metrics.

2. BACKGROUND
An organization uses process modeling to describe the business processes to be automated by organizing the activities that need to be performed to achieve a business goal in the correct sequence. A business process model is described in a visual manner and represents the way that business representatives conduct the operation of a business [1].

Process models for the emerging case management applications created with CMMN and using a declarative approach are more flexible than current BPM prescriptive models. Therefore, the current research on complexity metrics for BPM may not apply to CMMN, and so, there is a need to understand the applicability of BPM research regarding complexity metrics to case management process models and to create and empirically validate specific complexity metrics for case management process models.

2.1 Business Process Management
BPM technology describes and automates business processes. Those processes consist of an organized set of business activities required to achieve a business goal. A business process model is described in a visual manner and represents the way that business representatives conduct the operation of a business [1]. Figure 1 shows an example of a simple insurance claim process described in BPMN [13] by Korherr [7].

Understanding the business models is important for communication between users, and it facilitates the communication and transfer of knowledge throughout the organization.
Table 1: Summary of differences between BPMN and CMMN

<table>
<thead>
<tr>
<th>BPMN</th>
<th>CMMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural</td>
<td>Declarative</td>
</tr>
<tr>
<td>Process centric</td>
<td>Data centric</td>
</tr>
<tr>
<td>Arcs describe the sequence</td>
<td>No predefined sequence</td>
</tr>
<tr>
<td>Guided work (head down workers)</td>
<td>Enables workers (knowledge workers)</td>
</tr>
<tr>
<td>Everything is modeled</td>
<td>Not everything is modeled</td>
</tr>
</tbody>
</table>

Complexity metrics can be used to describe the difficulty involved for a human to understand a business process [5].

2.2 Case Management

Case management [18] was first introduced by Van Der Aalst and Berens in 2001 [20] and by Reijers et al. [15] in 2003 to support the flexibility required by knowledge workers during a business process and to help them better process exceptions that may occur during that process. Unlike workflow, BPM and most other process methods that focus on what should be done in a process, case management focuses on what can be done to achieve the business goal of the process [21].

In 2009, the Object Management Group (OMG) issued a request for proposal (RFP) for the creation of a standard modeling notation for use with case management [12] and to serve as a complement to its BPMN specification. The resulting specification, CMMN [14] is based on business artifacts [9]. Documents, like spreadsheet, word processor documents, pictures, voice recordings, etc. are considered business artifacts. Kumaran et al. [8] arrived to the concept of adaptive documents as domain artifacts, which provided a document centric view of processes using business artifacts. Business artifacts have been found useful to model real-world business processes, including case management [19]. CMMN is a new way to describe business processes, based on business artifacts, using a data-centric perspective [11]. There are important differences between BPMN and CMMN as described in Table 1. Figure 2 shows the simple insurance claim process modeled using CMMN.

2.3 Complexity metrics

Extensive research has been conducted on the complexity of traditional BPM models. For example, Gruhn and Laue [5] adapted several software engineering complexity metrics to business processes models. However, today as far as the author is aware, there is no research on the complexity of case management processes models. Complexity metrics are based on the graphical notation of the models, and as shown in Figure 2, case management is missing the connectors (arrows lines), gateways (diamonds), and events (circles) that describe the execution sequence of the BPMN model (see Figure 1). BPMN and CMMN models are different (see Table 1), therefore, most of the research on complexity metrics may not be applicable to CMMN.

For example, Cardoso [4] activity complexity (AC) counts the nodes in the process graph. Assuming all the elements in a CMMN model are nodes, then AC can be calculated for both BPMN and CMMN. However, Cardoso [4] control-flow complexity (CFC) is not applicable to CMMN, because...
there are no concepts of AND, OR, or XOR nodes in CMMN. Table 2, shows AC and CFC calculated for the simple insurance claim process in Figures 1 and 2.

### Table 2: Two Cardoso complexity metrics

<table>
<thead>
<tr>
<th>Simple insurance claim process</th>
<th>AC</th>
<th>CFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPMN version (Figure 1)</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>CMMN version (Figure 2)</td>
<td>8</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

3. METHODOLOGY

The study will be limited to the identification and validation of a set of complexity metrics for CMMN. The identified complexity metrics will be theoretically and empirically validated, following the suggestions of Misra et al. [10, p 323] and Briand et al. [2, p 3].

Phase one will identify and theoretically validate potential complexity metrics for CMMN models. It will start with a literature review of BPM complexity metrics to identify potential metrics that can be adapted to be used with CMMN. The literature review will be followed by designing new complexity metrics based on the characteristics of the CMMN models. For each identified metric a theoretical validation will be conducted to guarantee the metrics have good mathematical properties as described by Misra et al. [10, p 323] and Briand et al. [2, p 3]. The identified metrics that pass the theoretical validation will be referred to as the calculated complexity metrics in phase two.

Phase two will empirically validate the identified calculated complexity metrics. An experiment will be conducted to examine differences between the calculated complexity metrics in terms of the perceived complexity of the subjects. A within-subjects experimental design will be used to perform the empirical validation. The independent variable will be the treatment to which the subjects will be exposed. The treatment will consist of a set of ten CMMN models with different calculated complexity. All the subjects will be exposed to all the treatments, however, the instruments will be ordered differently for each subject. The design of the experiment may follow the design used by Canfora et al. [3, p 118-120].

3.1 Research questions

The following research questions are identified for both phases of this proposal.

- **Q<sub>1</sub>** Are there applicable complexity metrics for CMMN [14]?
- **Q<sub>2</sub>** What is the relationship between the calculated complexity of the CMMN process model, as measured by the complexity metrics created for this study, and the ‘subjective evaluation’ of the complexity as measured by an instrument based on the work of Rolón et al. [16, p 59] and developed for this study?
- **Q<sub>3</sub>** What is the relationship between the calculated complexity of the CMMN process models, as measured by the complexity metrics created for this study, and user model comprehension, as measured by an instrument based on the work of Hadar et al. [6, p 2326] and developed for this study?
- **Q<sub>4</sub>** What is the relationship between the perceived complexity of the CMMN process models, as measured by ‘subjective evaluation’ Rolón et al. [16, p 59], and user model comprehension, as measured by an instrument based on the work of Hadar et al. [6, p 2326] and developed for this study?

4. CONTRIBUTION

This study explores complexity metrics for CMMN models, and will contribute to new knowledge by providing an understanding of complexity metrics for case management. In addition, this study will contribute to the literature with identification and validation of complexity metrics for case management models. Furthermore, this study will have practical implications for case management products and practitioners.

5. REFERENCES


APPENDIX

A. CONTROL-FLOW COMPLEXITY

Cardoso [2] control-flow complexity is defined as

\[ CFC(p) = \sum_{i \in \{ \text{XOR-splits of } p \}} CFC_{\text{XOR-split}}(i) + \sum_{j \in \{ \text{OR-splits of } p \}} CFC_{\text{OR-split}}(j) + \sum_{k \in \{ \text{AND-splits of } p \}} CFC_{\text{AND-split}}(k) \]

Where,

- \( CFC_{\text{XOR-split}}(\text{XOR activity}) = n \)
- \( CFC_{\text{OR-split}}(\text{OR activity}) = 2^n - 1 \)
- \( CFC_{\text{AND-split}}(\text{AND activity}) = 1 \)

\( n \) is the fan-out of the split, which corresponds to the number of outgoing arcs.

The calculation of \( CFC \) for Figure 1 takes into account that there are two \( CFC_{\text{XOR-split}} \) (diamonds with an X inside and outgoing arcs) each with a fan-out of 2, and one \( CFC_{\text{AND-split}} \). Therefore,

\[ \sum_{i \in \{ \text{XOR-splits of } p \}} CFC_{\text{XOR-split}}(i) = 2 + 2 = 4 \]
\[ \sum_{j \in \{ \text{OR-splits of } p \}} CFC_{\text{OR-split}}(j) = 0 \]
\[ \sum_{k \in \{ \text{AND-splits of } p \}} CFC_{\text{AND-split}}(k) = 1 \]

\[ \Rightarrow CFC(p) = 4 + 0 + 1 = 5 \]