Evaluating The Understandability of Hybrid Process Model Representations Using Eye Tracking: First Insights

Amine Abbad Andaloussi[1], Tijs Slaats[2], Andrea Burattin[1], Thomas T. Hildebrandt[2], and Barbara Weber[1] ⋆

[1] Software and Process Engineering, Technical University of Denmark
2800 Kgs. Lyngby, Denmark
[2] Department of Computer Science, University of Copenhagen
2100 København, Denmark

Abstract. The EcoKnow project strives to promote flexible case management systems in the public administration and empower end-users (i.e., case workers) to make sense of digitized models of the law. For this, a hybrid representation combining the declarative DCR notation with textual annotations depicting the law text and a simulation tool to simulate the execution of single process instances was proposed. This hybrid representation aims to overcome the notorious limitations of existing declarative notations in term of understandability. Using eye tracking, this paper investigates how users engage with the different artifacts of the hybrid representation.

1 Introduction

The Ecoknow project aims at integrating hybrid technologies in public administration as part of the effective digitization of knowledge work processes. In this context, a hybrid representation combining the declarative DCR (Dynamic Condition Response) notation [1] with textual annotations depicting the law and a simulation tool allowing to simulate the possible process executions was proposed (cf. Figure 1). While the use of declarative notations (i.e., DCR graphs) enables flexibility and higher adaptability, their understandability is controversial especially with regards to novice users [3]. This hybrid representation (called "hybrid DCR representation") aims to overcome the understandability limitation by offering a multi-artifact representation to help end-users to make sense of digitized models of the law.

As part of an exploratory study investigating the understandability of the hybrid DCR representation, this paper uses eye tracking to provide some first insights about the way end-users engage with the different artifacts of this hybrid representation. The outcome of this study will provide insights into the use of the

DCR platform. Section 2 presents briefly the related work. Section 3 describes the research method pursued to plan and conduct the exploratory study. Section 4 unveils some first insights into the study results, and finally Section 5 highlights the future work and concludes the paper.

2 Related Work

The literature about hybrid process model representations can be categorized into the following groups: *(a)* The first group comprises hybrid representations that combine two or more notations into a single artifact. For instance, in [2] and [7] the authors combine imperative and declarative notations into a single process model. This part of the literature is beyond the scope of this paper. Instead, this work focuses on a hybrid representation that combines several artifacts. *(b)* The second group considers approaches that combine a graphical model with textual annotations. For instance, Pinggera et al. in [5] proposed the Literate Process Modeling technique (LiProMo). Inspired by the dual coding theory, LiProMo aims at fostering the communication at the process modeling stage by fusing textual descriptions with a process model. Likewise, Wang et al. in [8] proposed a hybrid representation that combines a process model with linked rules expressed as textual annotations. Both approaches demonstrated higher comprehension accuracy and lower mental effort. *(c)* The third group of the literature comprises hybrid representations that combine a graphical model with a tool that allows the execution of single traces. For instance, Zugal et al. in [10] proposed a test driven approach to support modelers in maintaining declarative process models. Hereafter, the authors introduced a tool that allows to test whether the process model complies with a set pre-defined (positive and negative) test cases. The evaluation (cf.[9]) demonstrated an increased process maintainability with reduced mental effort. Similarly, this work combines a graphical model in DCR notation with textual annotations referring to the law. In addition, the scrutinized hybrid representation includes a simulation tool that allows to simulate the execution of single traces and perceive the allowed behaviour.

3 Research Method

To investigate the factors affecting the understandability of the hybrid DCR representation, we have planned and conducted an eye tracking experiment. This section highlights the key aspects considered in the design phase of the experiment and provides insights into the measurements deployed in the analysis.

Research Question. To obtain a better understanding about the way the hybrid DCR representation is used, we formulate the following research question: **"How end-users engage with the different artifacts proposed by the hybrid DCR representation?"** To answer to this question we analyze the following: *(a)* the distribution of attention between the different artifacts, and *(b)* the common reading patterns seen in different groups of end-users.
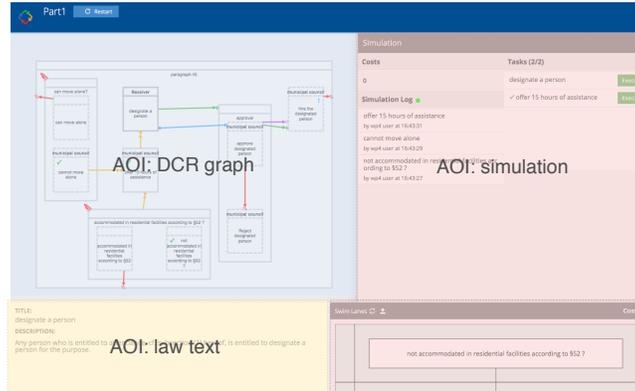
Fig. 1: A view showing the hybrid DCR representation. At the analysis phase, this view is split into 3 areas of interest, each representing a distinct artifact.

Subjects and Objects. The end-users (called "subjects" in an experimental context) who took part in this study have varying levels of expertise in using DCR. They have been recruited among case-workers from Syydjurs municipality in Denmark and students/employees from Technical University of Denmark (DTU) and IT university of Copenhagen (ITU). The DCR model used in this study originates from Section §45 of the "Consolidation Act on Social Services"[3].

Design. The experiment begins with a brief training session where all the subjects receive basic guidance about the hybrid DCR representation. Throughout the experiment eight comprehension tasks are displayed sequentially. The tasks evaluate the subjects' capacity to understand the semantics of the DCR graph as well as engage them to read the law fragments and to use the simulation. After each comprehension task, a set of questions investigating the artifacts used, the cognitive load and the subject's emotional state are prompted. By the end of the experiment, a think aloud session is held. Finally, a post-experiment questionnaire is used to collect data about the subjects' demographics, domain knowledge and experience[4] . For the sake of brevity, this paper will put emphasis only on the gaze data collected from the eye tracker, and the artifacts data. The rest of the data will be investigated in upcoming work. The experiment material used for this paper is available online at http://andaloussi.org/papers/DeHMiMoP2018/Material.pdf

Measures. To investigate the distribution of attention between the different artifacts, we have used the following fixation-derived measures [6]: *(a) fixation count* which quantifies the number of fixations on a specific area of the stimulus [4, p. 412-415], and *(b) total fixation duration* which sums up the duration

---

[3] http://english.sm.dk/media/14900/consolidation-act-on-social-services.pdf (Eng), https://www.retsinformation.dk/Forms/R0710.aspx?id=197036 (Dan)

[4] the post-experiment data is available online at http://andaloussi.org/papers/DeHMiMoP2018/demographicsAndBackground.xlsx

of all the fixations on a specific area of the stimulus [4, p. 377-386]. Indeed, as shown in Figure 1, the stimulus can be divided into several areas called "Areas Of Interest" (AOIs), such that each AOI corresponds to a different artifact (DCR graph, law text, simulation). Finally, we have compared the distribution of attention with the subjective artifact data used by the subjects to answer the comprehension tasks.

To investigate the common reading patterns, we have developed a new approach to analyze the transitions between the different AOIs. Given a timestamped log file containing the sequence of fixations and their corresponding AOIs, we have generated an event log and used the process mining tool Disco[5] to discover and analyze the underlying reading pattern. More insights about the experimental settings and the measures are demonstrated in the video available online at `https://youtu.be/8OsY9PYAs3I`.

4 Early Results

This section provides insights about the understandability of the hybrid DCR representation. Section 4.1 investigates the distribution of attention, and Section 4.2 analyzes the different reading patterns.

4.1 Distribution of Attention

The data used for the analysis and the tables described in this section are available in an online spreadsheet at `http://andaloussi.org/papers/DeHMiMoP2018/Analysis.xls`. By looking at the total fixation duration for all participants, we have noticed that the DCR graph was the most focused artifact (duration: 881.767 sec, proportion: 0.511), followed by the law text (duration: 443.251 sec, proportion: 0.257), then the simulation (duration: 398.954 sec, proportion: 0.231). Comparing these values with the fixation count for all participants, we have noticed a similar distribution of attention with almost similar proportions. The same observation holds with the subjective artifact data. Overall, this comparison between the artifacts shows that the graph caught most of the subjects' attention. However, this observation does not provide enough insights into the usability of the law text and the simulation, since the different AOIs differ in size and content.

As an alternative, we have compared the subjects based on their proportions of fixation duration and fixation count on each artifact (cf. AOI artifacts Sheet). Hereafter, we have observed the following: *(a)* subjects with highest fixation duration and fixation count proportions on the graph (P03, P07, P09 and P10), *(b)* subjects with highest fixation duration and fixation count proportions on the law text (P01 and P05), and *(c)* subjects with highest fixation duration and fixation count proportions on the simulation (P02, P06 and P08). These observations allow distinguishing between three user profiles with varying artifacts preference (DCR graph, law text, simulation), which lead to the question: "Do the three user profiles exhibit different reading patterns?"

---

[5] Available online at `https://fluxicon.com/disco/`

4.2 Reading Patterns

To investigate the reading patterns for the different user profiles, we have analyzed the transitions between the different AOIs using the process mining tool Disco. For sake of brevity, this analysis was conducted on a single question (Question 3, cf. online material). Figure 2 depicts the reading patterns for the different profiles. The activities in the model represent the different AOIs of the stimulus, the digits on the activities refer to the absolute frequency of AOI visits, the arcs between the activities indicate the transitions between the AOIs, and the digits on the arcs count the number of transitions. The arcs looping around the same activity refer to intermediate transitions to areas that are not relevant to our analysis.

Figure 2a shows that the graph profile has high abs. visit frequency on the graph and few transitions from the graph to the law text and simulation. By looking at the gaze video recordings, we have noticed that the graph profile subjects were constantly checking the graph semantics to identify the answering clues. However, since the graph is not providing clear answering clues to all the questions, some subjects have used the law text and the simulation to complement their understanding. Figure 2b shows that the simulation profile has balanced abs. visit frequency on the graph and the simulation, and more transitions between the graph and simulation compared to the transitions between the graph and the law text. By looking at the gaze video recordings, we have noticed that the simulation profile subjects were constantly checking the DCR relations surrounding the activities targeted by the simulation, which explains the high frequency of transitions between the graph and simulation AOIs. Finally, Figure 2c shows that the law text profile has two thirds of the abs. visit frequency on the graph while the remaining is on the law text. By looking at the gaze video recordings, we have noticed that the subjects from this profile have not used the simulation, thus, the fixations on the simulation AOI for this profile were ignored. Hereafter, frequent transitions were between the graph and the law text. This observations demonstrate that each user profile has a distinct reading pattern.

5 Conclusion and Future Work

This paper describes the approach used to investigate the understandability of the hybrid DCR representation. The analysis depicts the distribution of attention between the different artifacts, and shows that each user profile exhibits a different reading pattern. It has to be noted that the reported results are subject to limitations due to the small number of subjects, and the use of domain knowledge by some subjects (i.e., case workers from the municipality) to answer to some of the experiment questions.

As future work, the user interactions and the verbal data transcribed from the think aloud will be used to explain the transitions between the different artifacts, and to identify the circumstance of using each of them. Moreover, the verbal data will be analyzed to spot the typical challenges faced by the subjects,
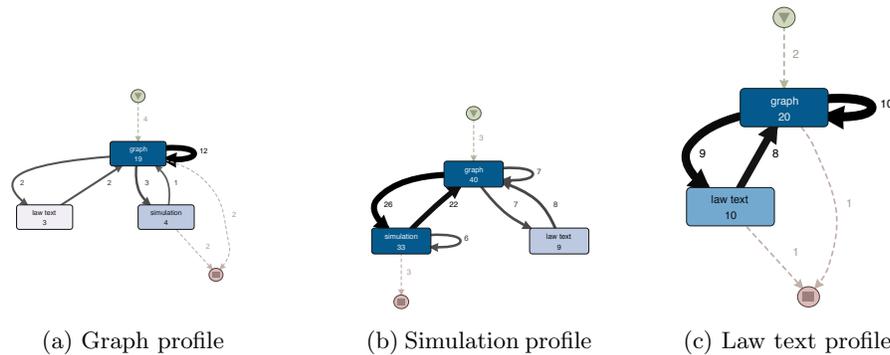
(a) Graph profile   (b) Simulation profile   (c) Law text profile

Fig. 2: Reading patterns for the different user profiles. Higher resolution at `http://andaloussi.org/papers/DeHMiMoP2018/userprofiles.pdf`

and the questionnaire data will be used to measure the comprehension accuracy. Finally, the subjects' cognitive load and emotional reactions will be examined using the pupil data and the Galvanic Skin Response (GSR) data.

References

1. J. De Smedt, J. De Weerdt, E. Serral, and J. Vanthienen. Improving Understandability of Declarative Process Models by Revealing Hidden Dependencies. In *Advanced Information Systems Engineering*, pages 83–98. Springer, 2016.
2. J. De Smedt, J. De Weerdt, J. Vanthienen, and G. Poels. Mixed-Paradigm Process Modeling with Intertwined State Spaces. *Business & Information Systems Engineering*, 58(1):19–29, feb 2016.
3. D. Fahland, D. Lübke, J. Mendling, H. Reijers, B. Weber, M. Weidlich, and S. Zugal. Declarative versus Imperative Process Modeling Languages: The Issue of Understandability. In *Proceedings of EMMSAD*, pages 353–366. Springer, 2009.
4. K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer. *Eye Tracking: A comprehensive guide to methods and measures*. OUP, 2011.
5. J. Pinggera, T. Porcham, S. Zugal, and B. Weber. LiProMo-Literate process modeling. *CEUR Workshop Proceedings*, 855:163–170, 2012.
6. A. Poole and L. J. Ball. Eye tracking in human-computer interaction and usability research: Current status and future. In *Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc*, 2005.
7. T. Slaats, D. M. M. Schunselaar, F. M. Maggi, and H. A. Reijers. The Semantics of Hybrid Process Models. pages 531–551. 2016.
8. W. Wang, M. Indulska, S. Sadiq, and B. Weber. Effect of Linked Rules on Business Process Model Understanding. pages 200–215. 2017.
9. S. Zugal, J. Pinggera, and B. Weber. The Impact of Testcases on the Maintainability of Declarative Process Models. In *Enterprise, Business-Process and Information Systems Modeling*, pages 163–177. Springer, 2011.
10. S. Zugal, J. Pinggera, and B. Weber. Creating Declarative Process Models Using Test Driven Modeling Suite. In S. Nurcan, editor, *IS Olympics: Information Systems in a Diverse World*, pages 16–32. Springer, 2012.