Event Correlation in Non-Process-Aware Systems

Ricardo Pérez-Castillo[1], Barbara Weber[2], Ignacio García-Rodríguez de Guzmán[1], Mario Piattini[1] and Jakob Pinggera[2]

[1] Instituto de Tecnologías y Sistemas de Información (ITSI)
University of Castilla-La Mancha
Paseo de la Universidad 4 13071, Ciudad Real, Spain
{ricardo.pdelcastillo, ignacio.grodriguez, mario.piattini}@uclm.es

[2] University of Innsbruck
Technikerstraße 21a, 6020, Innsbruck, Austria
{barbara.weber, jakob.pinggera}@uibk.ac.at

Abstract. Since business processes supported by traditional systems are implicitly defined, correlating events into the appropriate process instance is not trivial. This challenge is known as the event correlation problem. This paper presents an adaptation of an existing event correlation algorithm and incorporates it into a technique to collect event logs from the execution of traditional information systems. The technique first instruments the source code to collect events together with some candidate correlation attributes. Secondly, the algorithm is applied to the dataset of events to discover the best correlation conditions. Event logs are then built using such conditions. The technique has been semi-automated to facilitate its validation through an industrial case study involving a writer management system and a healthcare evaluation system. The study demonstrates that the technique is able to discover the correlation set and obtain well-formed event logs enabling business process mining techniques to be applied to traditional information systems.

Keywords. Process Mining, Event Correlation, Event Model, Case Study.

1 Motivation

Many companies today carry out a vast amount of daily operations through their information systems without having ever done their own enterprise modeling. Business process mining is a well-proven solution used to discover the underlying business process models that are supported by existing information systems. Business process discovery techniques employ event logs, which are recorded by process-aware information systems, as input. However, there are a wide variety of traditional systems without any in-built mechanism to collect events (representing the execution of business activities). To enable the application of process mining techniques to traditional information systems, mechanisms for collecting events from non-process-aware systems have been proposed. Unfortunately, business processes supported by

traditional systems are implicitly defined, correlating events into the appropriate process instance is not trivial. This challenge is known as the event correlation problem.

2 Proposal

The main proposal is a technique to obtain event logs from traditional information systems[1], which consists of four main stages: (1) The first stage is aimed at instrumenting traditional systems so that such systems can collect events during their execution. Experts identify some information such as candidate correlation attributes, whose runtime values will then be collected together with each event. This approach incorporates a set of guidelines for business experts to facilitate the instrumentation of source code. Additionally, it proposes a metric based on package dependencies to suggest experts which parts of source code should be instrumented. (2) In the second stage, the modified system is executed and events are progressively recorded. As a result, events and their respective attributes are then stored in a database in an intermediate format. (3) The third stage applies an adaptation of the algorithm proposed by *Motahari-Nezhad et al.* to the event datasets in order to discover the necessary set of attributes and conditions to correlate events. Finally, (4) the last stage applies an algorithm taking the correlation set into account to correlate each event with its corresponding process instance. As a result, a standard-format event log is obtained from the traditional system.

3 Empirical Validation Results

The most important contribution of this paper is an empirical study conducted to demonstrate the feasibility of the technique as well as its application in the industry. The study was applied with two industrial, traditional information systems: *AELG-Members*, an author management system and *CHES*, a system for collecting patient reported outcome (PRO) data. The study collected thousands of events to be analyzed after and obtain different correlation sets according to different parameters. The correlation sets were then used to generate different event logs, which were in turn used to discover business processes. Finally, the mined business process models were compared with the reference models. The analysis of the study's results shows that the technique is able to obtain event logs from traditional systems. However, the correlation set depend on the amount of collected events and the beta factor (which determines the way in which process instances are built). Although, more collected events can be good to obtain better correlation sets, the time spent on discovering such information grows according to a quadratic function. As a consequence, the work-in-progress is currently focusing on obtaining correlation sets more efficiently (to reduce the response time of the discovering algorithm) as well as the effectiveness (to provide accurate correlation sets with fewer events).

[1] Pérez-Castillo, R., B. Weber, I. García Rodríguez de Guzmán, P. Mario, and J. Pinggera, Assessing Event Correlation in Non-Process-Aware Information Systems. Software and Systems Modeling Journal, 2013. In Press.

A HCI Technique for Improving Requirements Elicitation

John W. Castro[1], Silvia T. Acuña[1], Natalia Juristo[2]

[1]Departamento de Ingeniería Informática, Universidad Autónoma de Madrid
Calle Francisco Tomás y Valiente 11, 28049 Madrid, Spain
john.castro@estudiante.uam.es, silvia.acunna@uam.es
[2]Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo s/n, 28660 Boadilla del Monte, Madrid, Spain
natalia@fi.upm.es

1 Summary

Human computer interaction (HCI) uses development processes and techniques to assure that software product usability complies with minimum requirements. The Personas technique [3] gathers, analyses and synthesizes information related to the users that are to interact with the software system. This technique helps to focus software analysis and design on end user features and goals. However, it shares the shortcomings of other HCI techniques: it has no detailed definition of activities and products. These problems make the introduction of Personas into the software engineering (SE) requirements stage overly complex and unclear for developers. In order to design and implement a usable system, there should, according to HCI, be an understanding not only of users' needs and goals but also of their characteristics and capabilities. The understanding of the people that interact with the system should constitute the groundwork for software development. The SE requirements activity could be improved by incorporating Personas technique tasks to understand the user. The goal of our research is to modify Personas to readily build the technique into the requirements stage of regular SE developments.

We studied Cooper's version of the Personas technique [3] and set out to apply this technique in a case study [2]. From the very outset, we had trouble applying the technique. For example, the first step of the technique recommended by Cooper is Identify Behavioural Variables, that is, Cooper assumes that users have already been researched and the gathered data have been roughly organized. This task is not however explicitly mentioned in his description of Personas. The user study necessary to extract behavioural variables is not an altogether straightforward step and should be specified rather than implied as the technique's first activity. Additionally, some technique activities, like Identify Significant Behaviour Patterns and Check for Completeness and Redundancy, fail to specify any output product. Finally, the final technique outputs are not related to the software engineering requirements stage. To be able to build Personas into routine SE developments, it is necessary to define activities and products associated with each activity. For each of the identified limitations, we devised an improvement to be built into Personas. We opted to incorporate these im-

provements into the latest version of the Personas technique published by Cooper et al. [3]. The grounds for this choice were: (i) Cooper made the original proposal; (ii) this proposal was the groundwork for research by other authors; and (iii) this proposal has been successfully used in a number of real projects. In [1], we incorporated these improvements into a SE version of Personas. Our proposal is composed of a group of activities and their associated inputs and outputs that, together, lead to the creation of personas [1]. As part of the first new activity, State Hypotheses for Personas, for example, a List of Hypotheses for the Personas to be created should be generated, and interviews should be designed and held with potential users. The responses from the Transcribed Interviews should then be used to gather the information required to carry out other activities. Second, as part of the Identify Behavioural Variables activity, we propose a new activity for synthesizing each response to the interviews held in the previous activity as behavioural variables. Third, we have defined a new activity that links the user research using Personas with the remainder of the requirements stage: Build Use Cases. This activity should output an Annotated Use Case Diagram. This diagram is based on the traditional use case diagram, to which we add a brief description of each persona involved in the use case. We applied our proposed technique to several case studies for validation [1]. Additionally, we designed and implemented a prototype tool to support the proposed Personas technique.

The improved Personas avoids the obstacles encountered by an average software developer unfamiliar with HCI techniques applying the original Personas. We think it is worthwhile adapting Personas for integration into SE development process. The integration of Personas into the SE requirements stage could improve the understanding of what the software product should do and how it should behave. The Personas technique appears to help focus the software analysis and design activities on end user characteristics and goals. We have enriched the SE requirements process by incorporating Personas activities into requirements activities. Requirements elicitation and requirements analysis are the requirements engineering activities most affected by incorporating Personas.

Acknowledgements. This work has been funded by the Spanish Ministry of Science and Innovation as part of the *Tecnologías para la Replicación y Síntesis de Experimentos en IS* (TIN2011-23216) and *Go Lite* (TIN2011-24139), and by Community of Madrid R&D program *e-Madrid* project (S2009/TIC-1650).

References

1. Acuña, S.T., Castro, J.W., Juristo, N.: A HCI Technique for Improving Requirements Elicitation. Information and Software Technology. 54 (12), 1357-1375 (2012).
2. Castro, J.W.: Extensión de la Técnica Personas para su Incorporación en la Actividad de Análisis del Proceso de Ingeniería del Software (Extension of the Personas Technique for Incorporation into the Software Engineering Development Process Analysis Activity). Master Thesis, Universidad Autónoma de Madrid, Madrid, Spain (2009).
3. Cooper, A., Reimann, R., Cronin, D.: About Face 3.0: The Essentials of Interaction Design. Wiley Publishing, Indianapolis (2007).