Artificial Intelligence in Medicine 62 (2014) 33–45

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

journal homepage: www.elsevier.com/locate/aiim

# Improving structural medical process comparison by exploiting domain knowledge and mined information

Stefania Montani[a,*], Giorgio Leonardi[a], Silvana Quaglini[b], Anna Cavallini[c], Giuseppe Micieli[c]

[a] Dipartimento di Scienze e Innovazione Tecnologica, Computer Science Institute, Università del Piemonte Orientale, Viale Michel 11, I-15121 Alessandria, Italy
[b] Department of Electrical, Computer and Biomedical Engineering, Università di Pavia, Via Ferrata 1, I-27100 Pavia, Italy
[c] Istituto di Ricovero e Cura a Carattere Scientifico Fondazione "C. Mondino", Via Mondino 2, I-27100 Pavia, Italy[1]

ARTICLE INFO

*Article history:*
Received 26 June 2013
Received in revised form 12 May 2014
Accepted 12 July 2014

*Keywords:*
Process mining and comparison
Graph edit distance
Stroke management

ABSTRACT

*Objectives:* Process model comparison and similar process retrieval is a key issue to be addressed in many real-world situations, and a particularly relevant one in medical applications, where similarity quantification can be exploited to accomplish goals such as conformance checking, local process adaptation analysis, and hospital ranking. In this paper, we present a framework that allows the user to: (i) mine the actual process model from a database of process execution traces available at a given hospital; and (ii) compare (mined) process models. The tool is currently being applied in stroke management.
*Methods:* Our framework relies on process mining to extract process-related information (i.e., process models) from data. As for process comparison, we have modified a state-of-the-art structural similarity metric by exploiting: (i) domain knowledge; (ii) process mining outputs and statistical temporal information. These changes were meant to make the metric more suited to the medical domain.
*Results:* Experimental results showed that our metric outperforms the original one, and generated output closer than that provided by a stroke management expert. In particular, our metric correctly rated 11 out of 15 mined hospital models with respect to a given query. On the other hand, the original metric correctly rated only 7 out of 15 models. The experiments also showed that the framework can support stroke management experts in answering key research questions: in particular, average patient improvement decreased as the distance (according to our metric) from the top level hospital process model increased.
*Conclusions:* The paper shows that process mining and process comparison, through a similarity metric tailored to medical applications, can be applied successfully to clinical data to gain a better understanding of different medical processes adopted by different hospitals, and of their impact on clinical outcomes. In the future, we plan to make our metric even more general and efficient, by explicitly considering various methodological and technological extensions. We will also test the framework in different domains.


## 1. Introduction

Process model comparison and similar process retrieval is a key issue to be addressed in many real-world situations. For example, when two companies are merged, process engineers need to compare processes originating from the two companies, in order to analyze their possible overlaps, and to identify areas for consolidation. Moreover, large companies build over time huge process model repositories, which serve as a knowledge base for their ongoing process management/enhancement efforts. Before adding a new process model to the repository, process engineers have to check that a similar model does not already exist, in order to prevent duplication. Particularly interesting is the case of medical process model comparison, where similarity quantification can also be exploited in a conformance checking perspective. Indeed, the process model actually implemented at a given healthcare organization can be compared to the existing reference clinical guideline, to check conformance, and/or to understand the level of adaptation to local constraints that may have been required. As a matter of fact, the existence of local resource constraints may lead to

* Corresponding author. Tel.: +39 0131 360158.
*E-mail address:* stefania.montani@unipmn.it (S. Montani).
[1] On behalf of the Stroke Unit Network (SUN) collaborating centers.

http://dx.doi.org/10.1016/j.artmed.2014.07.001
0933-3657/

differences between the models implemented at different hospitals, even when referring to the treatment of the same disease (and to the same guideline). A quantification of these differences (and maybe a ranking of the hospitals derived from it) can be exploited for several purposes, like, e.g., administrative purposes, performance evaluation and public funding distribution. The actual medical process models are not always explicitly available at the healthcare organization. However, a database of process execution traces (also called the "event log") can often be reconstructed starting from data that hospitals collect through their information systems (in the best case by means of workflow technology).

In this case, *process mining* techniques [1] can be exploited, to extract process models from event log data. Stemming from these considerations, in this work we present a framework, which allows the user to:

1. extract the actual process model from the available medical process execution traces, through process mining techniques;
2. perform medical process model comparison, to fulfill the objectives described above.

Item 2 has required the introduction of proper metrics, in order to quantify process model similarity. We could rely on an extensive literature when studying this topic (see Section 4). In particular, since process mining extracts the process model in the form of a graph, our work is located in the research stream on graph structural similarity, and on graph edit distance-based approaches [2,3]. The state of the art on structural similarity on process models is represented by the work by Dijkman et al. [2]. Specifically, we have extended the work in [2], by:

- exploiting domain knowledge;
- exploiting process mining outputs and statistical temporal information.

We believe that the use of domain knowledge represents a significant enhancement in the metric definition, which, otherwise, would operate in a "blind" and context-independent fashion. Indeed, the original metric in [2] is completely independent of the domain of application. On the other hand, when domain knowledge is available, rich and well consolidated, as is often the case in medicine, its exploitation can surely improve the quality of any automated support to the expert's work – including process comparison (see e.g., [4]). Moreover, the use of additional information extracted from data, and of temporal information in particular, can be a relevant advancement as well, in fields in which the role of time can be very critical, like, e.g., emergency medicine. We are currently applying our framework to stroke management. In this domain, the positive experimental results we have obtained support the statements above. Indeed, our metric has proved to outperform the original metric in [2], and to generate outputs that are closer to those provided by a stroke management expert (see Section 3.1). Having verified the reliability of our tool through the experimental study described in Section 3.1, we have then applied it to address a key, open research question in stroke management, namely: do similar process models (implemented in different stroke units) lead to similar clinical outcomes (e.g., patient survival rate and/or patient improvement rate at discharge)? Some interesting conclusions on this issue were obtained (see Section 3.2), testifying the potential clinical usefulness of our contribution. The paper is organized as follows. Section 2 provides the details of our methodological approach. Section 3 showcases experimental results. Section 4 compares our contribution to related works. Section 5 shows the limitations of our work, and our future research directions, meant to overcome the open issues. Section 6 illustrates our concluding remarks.

## 2. Methods

In this section, we will first introduce process ming and the ProM tool; then we will provide the technical details of our metric.

### 2.1. Process mining and the ProM tool

Process mining describes a family of a-posteriori analysis techniques exploiting the information recorded in event logs, to extract process-related information (e.g., process models). Typically, these approaches assume that it is possible to sequentially record events such that each event refers to an activity (i.e., a well-defined step in the process) and is related to a particular process instance. Furthermore, some mining techniques use additional information such as the timestamp of the event, or data elements recorded with the event.

Traditionally, process mining has focused on discovery, i.e., deriving process models and execution properties from event logs. It is important to mention that there is no a-priori model, but, based on logs, some models, e.g., a Petri net, are constructed. However, process mining is not limited to process models (i.e., control flow), and recent process mining techniques have focused more and more on other perspectives, e.g., the organizational perspective, the performance perspective or the data perspective. Moreover, as clearly stated in [5], process mining also supports conformance analysis and process enhancement. In this paper, however, we will focus on the process perspective.

In our approach, we resorted to the process mining tool called ProM, extensively described in [6], and to ProM's Heuristic Miner [7] plug-in. ProM (and specifically its newest version ProM 6) is a platform-independent open source framework that supports a wide variety of process mining and data mining techniques, and can be extended by adding new functionalities in the form of plug-ins. Heuristic Miner [7] is a plug-in for mining process models from event logs. Heuristic Miner receives in input the log, and considers the order of the events within every single process instance execution. The timestamp of an activity is used to calculate this ordering. Heuristic Miner can be used to express the main behavior (i.e., not all details) registered in a log. Indeed, abstract information, such as the presence of composite tasks (i.e., tasks semantically related to their constituent activities by means of the "part-of" relation), cannot be derived by Heuristic Miner, which will only build a model including ground (i.e., not further decomposable) activities. On the other hand, it can mine the presence of short-distance and long-distance dependencies (i.e., direct or indirect sequence of activities), and information about parallelism, with a certain degree of reliability (see also Section 2.2). The output of the mining process is provided as a graph, known as the "dependency graph", where nodes represent activities, and edges represent control flow information. Heuristic miner does not extract behavioral/causal dependencies. The output can be converted into other formalisms as well. Currently, we have chosen to rely on Heuristics Miner, because it is known to be tolerant to noise, a problem that may affect medical event logs (e.g., sometimes the logging may be incomplete). Testing of other mining algorithms available in ProM 6 is, however, foreseen in our future work, as discussed in Section 5.

### 2.2. Distance definition for process model comparison

In order to compare process models on the basis of their distance, we have introduced a distance definition that extends previous literature contributions [2,3] (see also Section 4) by properly considering the information mined/learned from data, as well as