Benefits of Provenance in Home Care

Ákos HAJNAL, Tamás KIFOR, Gianfranco PEDONE, László Zsolt VARGA
System Development Department, Computer and Automation Research Institute, Hungary

Abstract. Electronic information processing is more and more common in healthcare applications and involves the handling of patient healthcare records, medical procedures and decisions. The distributed and heterogeneous nature of healthcare institutions results in disconnected information islands which hinder the treatment of patients. Better and user centred healthcare services require that the information islands are connected and medical professionals are able to view the treatment process as a whole. Healthcare record exchange standards help pulling together the data. The concepts of provenance and process documentation help to analyse the way results are produced. Given the technologies of patient record exchange and process documentation, we are investigating how these technologies can help to improve healthcare services provided in the home care environment of the K4Care EU project. We present the service oriented informatics environment of the K4Care platform and show that making this system provenance aware enables medical professionals to easily find the reasons why certain procedures were followed in a treatment and find out how the procedures can be improved.

Keywords. Home care, service oriented architecture, provenance, reunification of treatment process, analysis and improvement of treatment process.

Introduction

It is more and more common in healthcare applications that practitioners cooperate and share patient data electronically. Electronic information processing in this context involves three major fields: the patient healthcare record, the medical procedures carried out on the patient, and the crucial events in the medical procedures. These are handled by several heterogeneous and distributed information systems, possibly connected into a healthcare grid. These information systems form disconnected *islands of information* and are under the authority of different healthcare actors such as general practitioners, hospitals, hospital departments, etc. Although the different information systems provide services to each other and their clients in order to allow the sharing of information and distributed decision making, the distributed nature of healthcare institutions sometimes hinders the treatment of patients, because the documentation of the healthcare history and therapy of a patient cannot easily be collected from the distributed independent healthcare institutions.

The requirement to provide better, user-centred healthcare services demands that the different pieces of patient treatment are viewed as a whole. This means not only pulling together the different pieces of the healthcare record of a patient, but also discovering the logical links between the different pieces of the treatment processes executed at different places and time. From the informatics point of view the treatment processes are supported by software services connected through the network which can

be viewed as a service oriented architecture forming a grid. In this informatics infrastructure the individual services record information on the patient locally and then later when this information is needed by another service, the electronic healthcare record (EHCR) of the patient is transmitted using some standard format like ENV13606 [1] or HL7[2]. However these data exchange standards do not solve the reunification of the treatment process completely, because they support the data exchange, but they do not give support for locating the relevant information and finding the causal relationship between the record items. Therefore, it is not uncommon for doctors to depend on the patients themselves in order to find and include relevant data from previous treatments and tests.

The Grid Provenance EU project investigated the concept of *provenance* and developed an architecture to support provenance in information systems, especially in grid systems [3]. Making electronic systems *provenance aware* enables users to trace how a particular result has been arrived at by identifying the individual and aggregated services that produced a particular output. Answering questions related to the provenance of a particular output can be done with the help of the *process documentation* which is created by the provenance aware services participating in a distributed process execution. In healthcare systems the process documentation is the basis of an integrated view of the execution of treatment processes in order to a) analyze the performance of distributed healthcare services, b) to be able to carry out audits to assess that, for a given patient, the right decisions were made and the right procedures were followed and c) to help medical professionals to determine the relevance of the patient's medical record to the current treatment.

Given the above mentioned technologies of EHCR exchange and process documentation, we are going to investigate how these technologies can help to improve the healthcare services provided in a home care environment. Our investigations are within the framework of the K4Care EU project [5]. The goal of the K4Care project is to develop an informatics platform to manage the information needed to guarantee improved home care service for the new EU society and for the increasing number of senior population requiring personalized home care assistance. Within this environment we show the possible benefits of provenance awareness and how medical professionals can use process documentation to improve their activities in home care services.

In section 1 we introduce the K4Care home care environment and the service oriented informatics concepts that we developed. In section 2 we summarize provenance in service oriented architectures and grid systems. In section 3 we describe how provenance can be used in the K4Care home care environment. Finally in section 4 we conclude that making the K4Care environment provenance aware enables medical professionals to easily find the reasons why certain procedures were followed in a treatment and find out how the procedures can be improved.

1. The K4Care Home Care Environment

It is increasingly important to develop informatics eHealth applications to support people involved in providing basic medical care like physicians, nurses, patients, relatives, and citizens in general. The care of elderly people, who usually have chronic illnesses and sometimes are disabled, involves life long treatment under continuous expert supervision. In case of elderly people with chronic illnesses it is widely accepted that being cared for in hospitals or residential facilities may be unnecessary and even

counterproductive, while their treatment at home increases their life quality and reduces costs, therefore it is important to develop informatics support for home care.

While the healthcare information system of a single medical centre may be rather centralised, medical assistance in home care naturally needs distributed service oriented architecture. In order to support home care, the K4Care project develops a platform to manage the information needed to guarantee an ICT Home Care service. The K4Care Home Care platform will integrate information of different types and from different sources; be integrated with information and communication technologies whilst ensuring private and customized data access; use ontologies to define the profile of accessing subjects (e.g. physicians, patients) and objects (e.g. disease, case study); have a mechanism to combine and refine the ontologies to personalize the system taking into account the way a physician works and the individual patient characteristics; incorporate 'know-how' from geriatric clinical guidelines as Intervention Plans (IPs); generate IPs from the healthcare centre databases if clinical guidelines do not exist or are inappropriate for a particular situation; configure a knowledge-based decision support tool that can supply eServices to all subjects involved in the Home Care model; extract evidence from real patients and integrate it with published evidence derived from randomised control trials.

While the K4Care platform will contain a generic model for home care as well, in this paper we focus on the informatics aspect of the platform which is outlined in the following.

1.1. The K4Care Informatics Service Model

The K4Care platform will provide services to its users. Typical user categories are patients, family doctors, physicians in charge, nurses, head nurse, social workers, etc. Each user will achieve its goals with the help of a set of services specific to his or her user type. A set of services specific to a user type will be incorporated in a software component which we call *agent*. The agents may be distributed in the computer network. The services will invoke other services and thus the K4Care platform will have distributed service oriented architecture. Some of the services will correspond to medical processes and their execution processes will be based on medical guidelines, while other services will correspond to administrative or technical processes related to the operation of the platform or the home care centre.

The K4Care service model is derived from the processes of home care and is based on the following concepts:

- Service: it is an abstract notion of a complex activity which is typically accomplished in collaboration with several actors (see also task below). A service is identified by a unique name (or service id) and it may have several instantiations which are called *procedures*. Different procedures instantiating the same service may be for example different localisations of the same service, e.g. in different countries or medical centres. In a given K4Care platform installation each service has one and only one procedure instantiation.
- Procedure: it is a formal description of a set of *tasks* organized in some workflow (sequential, parallel, preconditions, etc.). The procedure may be the instantiation of a medical or any general process in the medical centre. If the procedure is the instantiation of a medical process, then we call it *Intervention Plan* (IP). The workflow control structures of the procedure are described in

some formal medical guideline language like Proforma [6][7], Asbru [8] or a specific language developed for the K4Care platform. Tasks can invoke the services of another agent in the system, therefore a procedure may be some composition of services. Procedures and IPs are created by humans, e.g. medical centre managers or physicians.

- Task: it is an execution step in a procedure and it is usually a request to execute another service. The task is described by an **n-tuple**: **task = (subject, object, service or action, doc, ….)**. The "*subject"* is the type of agent which is expected to execute the *"service"* or the *"action"*. Subject can be e.g. the agent of a specific nurse or the physician in charge. The "*object"* is the actor on which the service is expected to be executed (e.g. a specific patient). The "*doc"* is a document relating to the "*service*". All actors are expected to document their activities in this document. There may be other optional parameters. If the subject corresponds to the same agent that executes the current procedure, then the service is executed internally.
- Action: it is any activity that can be executed by the agent on its own. The set of actions that an agent can perform can be considered as the agent's skills. Each action of the agent is provided as a service for other agents. When an action is executed, no procedure description retrieval is needed. The set of services that a given agent type is capable of carrying out is described by the Actor Profile Ontology (APO) which describes the actors and their agent representation within the medical centre. Actions can be imaged as a piece of Java code that implements the action.

1.2. The K4Care Platform and EHCR Systems

In the process of supporting the home care activities by providing home care services as described above, the K4Care platform must have to retrieve data from EHCR systems as well, because the medical data of the patient from previous treatment may be necessary for the current home care activity. If the data was produced in a most recent treatment, for example just after the patient was released from the hospital to his or her home, and the home care centre is in interaction with the institute where the data was produced, then the needed data can be easily located and retrieved through EHCR data exchange standards, because the location and the reference of the data is known. However if the data was produced in a former treatment, even years ago in a hospital, when home care was not envisaged for the patient, then locating the data is difficult, because there is no direct interaction between the hospital and the home care centre. In this case the physician does not know and even the patient might not remember that there is some relevant data in that hospital. This is when provenance awareness can help to realise the importance of some EHCR data and locate that data as described in [4].

Figure 1 shows that agent Bd executes a procedure and in action A1 retrieves some data about patient Oy in the Home Care (HC) centre. This is done through the "HC Centre EHCR Store" which in turn connects to the "Hospital 3 EHCR Store" and retrieves the relevant information. Depending on the retrieved data either service S2 is executed by agent Be or service S3 is executed by agent Bf.

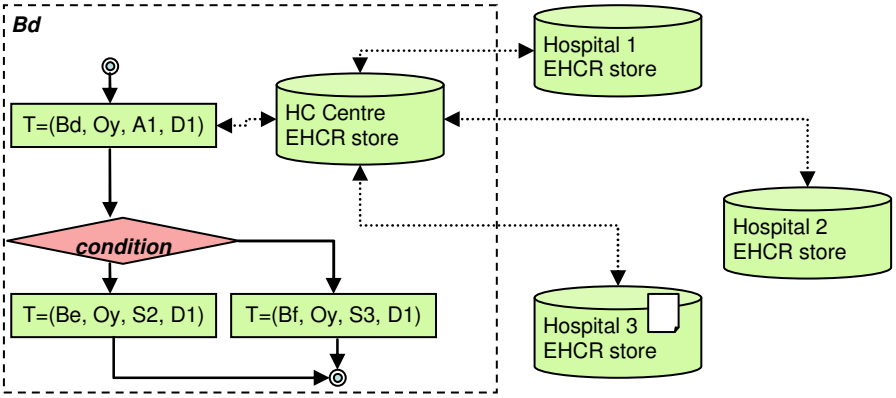

Figure 1. K4Care services and EHCR stores

1.3. K4Care Home Care Environment Summary

In the above we have described the service based informatics environment of the K4Care platform. We have also seen that this platform might use the services of EHCR stores. All together this environment can be seen as a service oriented architecture spanning over a grid of computing nodes inside and outside a home care centre. This operating environment provides the facilities needed for the usual operation of the home care centre together with the EHCR stores, but it does not support per se some analysis and verification functions.

For example if the procedure of Figure 1 is a procedure to be executed by a nurse and later the physician in charge would like to check this procedure and ask the following question "**Why was service S3 invoked?**", then it would be difficult for him or her to find the relevant information, because this is not directly present in the system. One can find out from the procedure description that action A1 was executed and than a decision was made and the execution of service S2 or S3 depends on the decision, but the relationship between these actions, services and data are not explicitly represented anywhere. In the following we will focus on provenance awareness and how provenance awareness can support answering such questions.

2. Provenance and Process Documentation

As described in [3], the concept of *provenance* is well known in fine art where it refers to the trusted, documented history of some work of art. This concept of provenance may also be applied to data and information generated within a healthcare information system. In accordance with the provenance architecture [9] developed by the Grid Provenance EU project [10], the provenance of a piece of data is represented in a computer system by a *process documentation* which is a suitable documentation of the process that produced the data. Provenance is investigated in open, large-scale systems typically designed using a service-oriented approach [11]. In this abstract view, interactions with services (seen as *actors* in the provenance architecture and realized as *agents* in the K4Care platform) take place using messages that are constructed in

accordance with service interface specifications. Such services are brought together by composition into a process to solve a given problem.

The process documentation is represented in a computer system by a set of *p-assertions*, which are assertions made by the actors involved in those processes, documenting some step of the process. The explicit description of the flow of data in a process is recorded by two kinds of p-assertions: *interaction p-assertions* and *relationship p-assertions*. An interaction p-assertion is an assertion of the contents of a message by an actor that has sent or received that message. A relationship p-assertion is an assertion about an interaction, made by an actor that describes how the actor obtained output data or the whole message sent in that interaction by applying some function to input data or messages from other interactions. An interaction p-assertion links together the actions of two actors in a process, while a relationship p-assertion links together multiple actions by a single actor. In addition, *actor state p-assertions* are assertions made by an actor about their internal state in the context of a specific interaction. The *provenance store* is the long-term facility for storing, managing and providing controlled access to process documentation.

The *provenance lifecycle* is composed of four different phases. First, actors *create p-assertions* that are aimed at representing their involvement in a computation. After their creation, *p-assertions are stored* in a provenance store, with the intent they can be used to reconstitute the provenance of some data. After a data item has been computed, users or *applications can query* the provenance store. At the most basic level, the result of the query is the set of p-assertions pertaining to the process that produced the data. More advanced query facilities may return a representation derived from p-assertions that are of interest to the user. Finally the provenance store and its contents *can be managed* (subscription management, content relocation, etc).

In the case of a healthcare information system, by recording all the medical processes related to a given patient one can get an explicit representation of the distributed processes that take place and re-construct the treatment history of the patient. Therefore, making a healthcare information system *provenance-aware* provides a way to have a unified view of a patient's medical record with its provenance.

3. Provenance in the K4Care Platform

In the following we describe how the procedure of Figure 1 is executed in a provenance enabled K4Care platform. The same procedure of Figure 1 and the graphical representation of p-assertions are shown on Figure 2. The provenance aware execution of the procedure consists of the following steps:

a) When agent Bd executes action A1 and invokes the "HC Centre EHCR store", then stores interaction p-assertion IPA1 in the provenance store. On Figure 2 the interaction p-assertion is represented by an arrow between the interacting parties: agent Bd of the first task of the procedure and the EHCR store.
b) The "HC Centre EHCR Store" uses EHCR system provenance information to find out that there is relevant data in "Hospital 3 EHCR Store". The way how provenance information is created and used by the EHCR system is described in [4], and this is not shown on Figure 2. The EHCR provenance information was created when the patient was treated in Hospital 3 possibly long time before the current retrieval.

c) The "HC Centre EHCR Store" retrieves the relevant data from "Hospital 3 EHCR Store" and stores interaction p-assertion IPA2 in the provenance store. This interaction p-assertion is about the interaction between the two EHCR stores.
d) The "HC Centre EHCR Store" stores a relationship p-assertion RPA1 in the provenance store. This relationship p-assertion is between IPA1 and IPA2, and represents that the data retrieval from "Hospital 3 EHCR store" is the consequence of the request from agent Bd.
e) When the requested EHCR data is returned to agent Bd, then the agent makes a decision based on the returned data.
f) As a consequence of the decision, agent Bd asks agent Bf to execute service S3 and stores interaction p-assertion IPA3 in the provenance store about the interaction between agent Bd and Bf.
g) Agent Bd stores a relationship p-assertion RPA2 in the provenance store. This relationship p-assertion is between IPA1 and IPA3, and represents that the invocation of service S3 is the consequence of the EHCR retrieval from "HC Centre EHCR store".

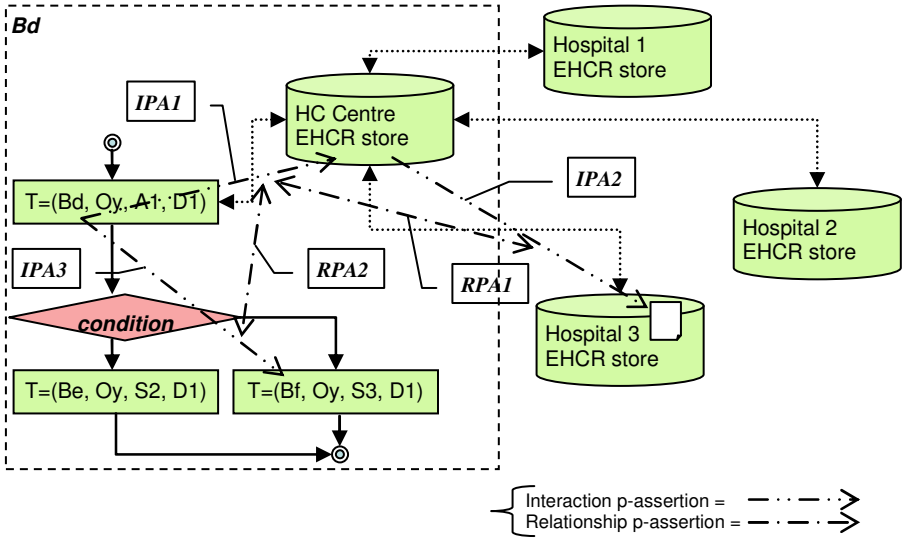

Figure 2. Provenance enabled execution of a K4Care procedure

The p-assertions recorded during the execution of the above procedure are the process documentation which can later be used to analyse the execution of the above procedure. So the question of section 1.3 can be answered by the provenance system. The physician in charge may sit in front of the analysis tool of the provenance system and pose the question. The provenance system can trace the graph of p-assertions and present to the physician in charge that service S3 was invoked because of the EHCR data returned by the "HC Centre EHCR store". The provenance system can also trace that the data returned by "HC Centre EHCR store" in fact originates from a treatment executed long time ago and stored in "Hospital 3 EHCR store".

Based on the information provided by the process documentation, the physician in charge can find out the reason why service S3 was executed and determine whether it was really needed.

The process documentation can also be used for the improvement of the processes themselves by analysing decisions and determining whether the procedure is correct or if it has to be redesigned.

4. Conclusions and Acknowledgements

We have developed and presented here a service oriented conceptual architecture for the K4Care home care environment. The concepts of this informatics environment are based on the home care processes investigated in the K4Care EU project. We have pointed out that the analysis and overview of the medical processes are not supported by the service oriented architecture in itself, however provenance awareness can help to answer analysis questions. We have shown how a provenance enabled K4Care platform works and enables medical professionals to easily find the reasons why certain procedures were followed in a treatment and find out how the procedures can be improved.

The provenance concepts and architecture was developed in the Provenance EU project [10]. The investigations presented here are part of the K4Care EU project [5].

References

[1] *ENV 13606-1:2000 Health Informatics—Electronic Healthcare Record Communication*, European Committee for Standardization, Tech. Committee 251 (Health Informatics), Work Group I (Information Models), 2000.

[2] *HL7 Version 3 Normative Edition, 2006*, Health Level Seven, Inc., 3300 Washtenaw Ave., Ann Arbor, MI 48104, https://www.hl7.org/library/bookstore/

[3] Luc Moreau et al., "The Provenance of Electronic Data," to be published in *Communications of the ACM*, 2007; www.ecs.soton.ac.uk/~lavm/papers/cacm06.pdf.

[4] Tamás Kifor, László Z. Varga, Javier Vázquez-Salceda, Sergio Álvarez, Steven Willmott, Simon Miles, Luc Moreau, "Provenance in Agent-Mediated Healthcare Systems," *IEEE Intelligent Systems*, vol. 21, no. 6, pp. 38-46, Nov/Dec, 2006.

[5] K4Care: Knowledge-Based HomeCare eServices for an Ageing Europe, IST-2004-026968, Specific Targeted Research or Innovation Project funded within Sixth Framework Programme of the European Commission (FP6), http://www.k4care.net/

[6] Jonathan Bury, John Fox, David Sutton, "The PROforma guideline specification language: progress and prospects," *Proceedings of the First European Workshop, Computer-based Support for Clinical Guidelines and Protocols* (EWGLP 2000), Leipzig 13-14 Nov. 2000.

[7] John Fox, Nicky Johns, Ali Rahmanzadeh, "Disseminating Medical Knowledge: the PROforma Approach," *Artificial Intelligence in Medicine*, vol. 14, no. 1-2, pp. 157-182, 1998.

[8] Robert Kosara, Silvia Miksch, "Metaphors of Movement: A Visualization and User Interface for Time-Oriented, Skeletal Plans," *Artificial Intelligence in Medicine*, vol. 22, no. 2, pp. 111-131, 2001.

[9] Paul Groth, Sheng Jiang, Simon Miles, Steve Munroe, Victor Tan, Sofia Tsasakou, Luc Moreau, "An Architecture for Provenance Systems," *Technical report*, ECS, University of Southampton, February 2006; http://eprints.ecs.soton.ac.uk/12023/

[10] Provenance: Enabling and Supporting Provenance in Grids for Complex Problems, IST-2002-511085, Specific Targeted Research or Innovation Project funded within Sixth Framework Programme of the European Commission (FP6), http://www.gridprovenance.org/

[11] M.P. Singh and M.N. Huhns, *Service-Oriented Computing: Semantics, Processes, Agents*, John Wiley & Sons, 2005.