Artificial Intelligence in Medicine 60 (2014) 133–149

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

Temporal abstraction and temporal Bayesian networks in clinical domains: A survey

Kalia Orphanou[a,*], Athena Stassopoulou[b], Elpida Keravnou[c]

[a] *Department of Computer Science, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus*
[b] *Department of Computer Science, University of Nicosia, P.O. Box 24005, 1700 Nicosia, Cyprus*
[c] *Department of Electrical Engineering, Computer Engineering and Informatics, Cyprus University of Technology, 30 Archbishop Kyprianou Street, 3036 Limassol, Cyprus*

ARTICLE INFO

Article history:
Received 6 February 2012
Received in revised form
15 November 2013
Accepted 27 December 2013

Keywords:
Temporal abstraction
Temporal reasoning
Bayesian networks
Temporal Bayesian networks
Medical knowledge-based systems

ABSTRACT

Objectives: Temporal abstraction (TA) of clinical data aims to abstract and interpret clinical data into meaningful higher-level interval concepts. Abstracted concepts are used for diagnostic, prediction and therapy planning purposes. On the other hand, temporal Bayesian networks (TBNs) are temporal extensions of the known probabilistic graphical models, Bayesian networks. TBNs can represent temporal relationships between events and their state changes, or the evolution of a process, through time. This paper offers a survey on techniques/methods from these two areas that were used independently in many clinical domains (e.g. diabetes, hepatitis, cancer) for various clinical tasks (e.g. diagnosis, prognosis). A main objective of this survey, in addition to presenting the key aspects of TA and TBNs, is to point out important benefits from a potential integration of TA and TBNs in medical domains and tasks. The motivation for integrating these two areas is their complementary function: TA provides clinicians with high level views of data while TBNs serve as a knowledge representation and reasoning tool under uncertainty, which is inherent in all clinical tasks.
Methods: Key publications from these two areas of relevance to clinical systems, mainly circumscribed to the latest two decades, are reviewed and classified. TA techniques are compared on the basis of: (a) knowledge acquisition and representation for deriving TA concepts and (b) methodology for deriving basic and complex temporal abstractions. TBNs are compared on the basis of: (a) representation of time, (b) knowledge representation and acquisition, (c) inference methods and the computational demands of the network, and (d) their applications in medicine.
Results: The survey performs an extensive comparative analysis to illustrate the separate merits and limitations of various TA and TBN techniques used in clinical systems with the purpose of anticipating potential gains through an integration of the two techniques, thus leading to a unified methodology for clinical systems. The surveyed contributions are evaluated using frameworks of respective key features. In addition, for the evaluation of TBN methods, a unifying clinical domain (diabetes) is used.
Conclusion: The main conclusion transpiring from this review is that techniques/methods from these two areas, that so far are being largely used independently of each other in clinical domains, could be effectively integrated in the context of medical decision-support systems. The anticipated key benefits of the perceived integration are: (a) during problem solving, the reasoning can be directed at different levels of temporal and/or conceptual abstractions since the nodes of the TBNs can be complex entities, temporally and structurally and (b) during model building, knowledge generated in the form of basic and/or complex abstractions, can be deployed in a TBN.

1. Introduction

Time is an integral aspect of medical problem solving. Various time-oriented systems have been developed focusing on different clinical tasks (e.g. prognosis) and areas (e.g. oncology). An important characteristic of these systems is the methodologies that they have used for the analysis, representation, interpretation and

---

* Corresponding author. Tel.: +357 22 892673; fax: +357 22 892701.
E-mail address: korfan01@cs.ucy.ac.cy (K. Orphanou).

0933-3657/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.artmed.2013.12.007

reasoning with longitudinal data [1]. Temporal abstraction (TA) and temporal Bayesian networks (TBNs) have become two topics of much interest and research in such systems. TA is a crucial process, especially in clinical monitoring, therapy planning and exploration of clinical databases. TA methods deal with the management and abstraction of time-oriented clinical data, providing high-level views of such data under given contexts. Their principal use so far has been in therapy planning and in the summarization and interpretation of patient records.

TBNs are temporal extensions of Bayesian networks, which are graphical models representing explicitly probabilistic relationships among variables. They can perform knowledge representation and reasoning under uncertainty and this makes their application highly suitable for the medical domain which is inherently uncertain. Uncertainty is present in all medical tasks and it can be in the form of uncertainty in the data, measurement uncertainty, as well as uncertainty in the medical knowledge. TBNs have been proposed in the literature to incorporate the explicit or implicit representation of time. TBNs have many applications in medicine in tasks such as medical diagnosis, forecasting, and medical decision making. They have been applied in various clinical domains to interpret and reason with large amounts of longitudinal data [2–6].

Although TA methods and many temporal extensions of BNs have been successfully used independently of each other in many clinical systems, they are complementary to each other. TA methods are used to create high-level temporal concepts from data while TBNs are used for reasoning and decision-making under uncertainty. An integration of these two methodologies with respect to clinical tasks, could potentially offer substantial benefits to clinicians. For instance, it could provide them with a concrete understanding of how causal dependencies between temporal abstract concepts influence a particular disease outcome. Knowing how state changes and how trends in variables affect a patient's state or the occurrence of a disease can facilitate treatment choices. Furthermore, the interpretation of high-level data can be effectively achieved by constructing a causal temporal model that would be able to explain the temporal patterns observed in the data. For example, the representation of temporal abstracted concepts in the nodes of a complex network such as TreatBN [7], would have resulted in a simpler and more comprehensible network, where the represented abstracted concepts and the prognostic results are compatible with experts' knowledge and medical literature. Moreover, the network can give a concrete understating of causal dependencies between abstract concepts occurring in a synchronous or asynchronous fashion utilizing temporal relations besides precedence (e.g. meets, during, overlaps). Taking this development into consideration, the current paper provides a broad review of what has been achieved until today independently in these two areas and discusses the anticipated benefits of their integration, which is what motivates our research in the direction of integrating TA with TBNs.

This survey reviews key aspects of TA methods with respect to clinical systems, such as the acquisition of knowledge driving the derivation of temporal abstractions and the methodologies for deriving basic and complex temporal abstractions. The types of temporal abstractions (basic, complex), and consequently the knowledge required for their derivation, constitute the linkage with the Bayesian networks. In other words, the derived temporal abstractions will be given as input to the Bayesian network. Both basic and complex abstractions should be used in the inference process. Acquiring the knowledge for deriving the abstractions (knowledge acquisition) is an essential part of a TA process which is also useful for the construction of a TBN. With respect to TBNs, we have chosen to focus on the categories of those networks that have applications in clinical domains. These categories are reviewed based on the following criteria: (a) representation of time, (b) knowledge representation and acquisition, (c) inference methods and the computational demands of the network, and (d) their applications in medicine.

A simple example, drawn from the domain of diabetes, is used for illustrating the different categories of TBNs. The same example is used in Section 4 to discuss the potential integration of temporal abstraction with each of the surveyed TBNs. The diabetes domain was chosen due to the fact that many TA methodologies and TBN model categories have applications in this domain [8–12].

A comprehensive review on temporal abstraction in intelligent clinical data analysis was published in 2007 [13]. That review provided information on temporal abstraction techniques applied to clinical systems and their main features in general. However, it did not address the potential integration of TA methods with respect to a particular problem solving engine such as the TBNs, which is the focus of our study. As such, the aforementioned survey is broader in context with respect to its coverage of TA methods and their features (information about the data that are abstracted, complex temporal abstractions, dimensionality of TA, etc.) in comparison to our survey, but it addresses TA methods in isolation of any higher decision making per se. Our discussion of TA is on the one hand from a narrower perspective, namely how to deploy effectively TA methods in conjunction with TBNs, but on the other hand is from a specific focus. As expected, there is some overlap between the two surveys regarding their discussion of TA methods, but our focus is distinct and is based on the possible integration of TA with TBNs.

The rest of the paper is organized as follows: Section 2 provides a comparative analysis of temporal abstraction techniques (basic and complex TA) and knowledge acquisition methods for creating abstractions in the most well known clinical systems. Section 3 presents those temporal extensions of Bayesian networks that have been applied in medical domains. Section 4 discusses the possible integration of TA with each of the presented TBN models using the diabetes example, and pointing out the advantages and disadvantages of each pairwise integration. Finally, Section 5 concludes the discussion by outlining the benefits of the proposed integration and giving pointers for further research.

2. Temporal abstraction

Time can be represented either as time points or as time intervals. Time points (e.g. instants) are related to distinct events whereas intervals are related to situations lasting for a period of time. The most influential, and by now classical, logical formalisms for representing temporal information are those developed by Allen [14], McDermott [15] and Kowalski and Sergot [16]. Another formalism proposed by Shoham [17] is an improvement of those three formalisms, having precise syntax and semantics.

In real life, data (and knowledge), especially in medical domains, are riddled with uncertainty and incompleteness, or are unduly low-level and voluminous for direct reasoning with. Various forms of abstraction are therefore called for. A basic abstraction mechanism is used for mapping data using a particular temporal granularity (or time unit) or a set of temporal granularities, thus having multiple granularities and means of traversing from one to the other. When there are multiple temporal granularities, these are often hierarchically organized, thus talking about finer/coarser granularities, but there can be structural relations between distinct granularities. A granularity functions to convert a dense time-axis to discrete chunks (granules) of time, enabling data to be expressed with respect to conceptually meaningful units of time, instead of infinitesimal references to time. A temporal granularity is therefore used to specify the temporal qualification of a set of data. Temporal qualifications are always chronologically ordered and relate the occurrence time of an event to another event [18]. As mentioned,