Flexible and Efficient Retrieval of Haemodialysis Time Series

Stefania Montani, Giorgio Leonardi, Alessio Bottrighi, Luigi Portinale,
and Paolo Terenziani

DISIT, Sezione di Informatica, Università del Piemonte Orientale, Alessandria, Italy

Abstract. The problem of retrieving time series similar to a specified query pattern has been recently addressed within the Case Based Reasoning (CBR) literature. Providing a flexible and efficient way of dealing with such an issue would be of paramount importance in medical domains, where many patient parameters are often collected in the form of time series. In this paper, we describe a novel framework for retrieving cases with time series features, relying on Temporal Abstractions. With respect to more classical (mathematical) approaches, our framework provides significant advantages. In particular, multi-level abstraction mechanisms and proper indexing techniques allow for flexible query issuing, and for efficient and interactive query answering. The framework is currently being applied to the hemodialysis domain. In this field, experimental results have shown the superiority of our approach with respect to the use of a classical mathematical technique in flexibility, user friendliness, and also quality of results.

Tests in other application domains, as well as further enhancements, are foreseen in our future work.

1 Introduction

Several real world applications require to capture the evolution of the observed phenomenon over time, in order to describe its behavior, and to exploit this information for future problem solving.

This issue is particularly relevant in medical applications, where the physician typically needs to recall the clinical history that led the patient to the current condition, before prescribing a therapy; actually, the pattern of the patient's changes is often more important than her/his final state [19]. The need for capturing the phenomenon's temporal evolution emerges even more strongly when a continuous monitoring of the patient's health indicators is required, such as in chronic diseases [3], or when control instruments that automatically sample and record biological signals are adopted, such as in hemodialysis [25] and Intensive Care Units [30]. In these applications, (many) process features are naturally collected in the form of **time series**, either automatically generated and stored by the control instruments (as e.g. in hemodialysis), or obtained by listing single values extracted from temporally consecutive situations (as e.g. the series of glycated hemoglobin values, measured on a diabetic patient once every two months).

Interpreting time series features on screen or on paper can be tedious and prone to errors. Physicians may be asked to recognize small or rare irregularities in the series itself, or to identify partial similarities with previous situations, such as retrieving the past occurrence of a similar trend behavior in a given feature, independently of its actual values. Such an identification may be extremely relevant for patient care, but may also require a significant amount of expertise in the specific field [29]. A user-friendly and flexible retrieval support is therefore strongly desirable in the medical domain.

The problem of retrieving time series similar to a given query pattern has been recently addressed specifically within the reasoning paradigm known as Case Based Reasoning (CBR) [1]; CBR exploits the knowledge collected on previously experienced situations, known as *cases*, and has been recognized as a valuable decision support methodology in time dependent domains, as testified by the wide number of works in the field (see section 2).

However, in time dependent domains, the need for describing the process dynamics strongly impacts both on case representation and on case retrieval, as analyzed in [24]. In particular, representing a case requires to properly represent (one or more) time series features; then, similarity-based time series retrieval has to be addressed and optimized, and becomes one of the most relevant issues. In the literature, most of the approaches to similarity-based time series retrieval are founded on the common premise of dimensionality reduction, which also simplifies knowledge representation (see the survey in [13]). Dimensionality is often reduced by means of a mathematical transform able to preserve the distance between two time series (or to underestimate it) (see section 2).

However, mathematical transforms have several limitations, as they can be computationally complex, and usually work in a black box fashion with respect to end users, lacking of flexibility and interactivity.

In this paper, we propose: (i) to exploit **Temporal Abstractions** (TA) [35,4,31,23], a different technique for dimensionality reduction and time series retrieval, and (ii) to use TA as the basis for a flexible and efficient retrieval of medical time series data, given a query. Indeed, to the best of our knowledge, TA techniques have not been exploited yet in the context of retrieval of time series. The main challenge of our approach is to fill such a gap, extending and adapting TA techniques, in order to achieve the first time series retrieval approach providing users with **flexibility** and user-friendliness. Moreover, our retrieval process has been designed to be computationally **efficient**, since in our framework query answering takes advantage of *multi-dimensional orthogonal index structures* for focusing and early pruning. We are currently applying the approach to the hemodialysis domain.

In the next sections, we describe the main characteristics of our framework, and we provide some experimental results we have obtained in hemodialysis. The paper is organized as follows. Section 2 discusses related works and introduces some preliminaries needed to present the approach (basically, Temporal Abstractions). Section 3 introduces the hemodialysis domain. Section 4 presents

the approach. Section 5 provides the experimental results. Section 6 reports on our conclusions and future work directions.

2   Preliminaries and Related Works

In recent years, several CBR works dealing with cases with time series features have been published, in the medical domain [25,27,29,28], as well as in different ones (e.g. robot control [32], process forecast [26,34], process supervision [11], pest management [7], prediction of faulty situations [16]). Moreover, general (e.g. logic-based) frameworks for case representation in time dependent domains have been proposed [24,15,21,6].

In these approaches, dimensionality reduction is typically obtained through classical mathematical techniques, such as the Discrete Fourier Transform (DFT) [2] (e.g. [25]) and the Discrete Wavelet Transform (DWT) [8] (e.g. [28]).

However, mathematical transforms have several limitations. First, they can be computationally complex, and often require additional pre-processing steps, such as mean averaging and zero padding (see e.g. [29] for details), or post-processing steps (for eliminating false alarms, i.e. spurious cases). Additionally, they work well with series with relatively simple dynamics, but they can fail to characterize more complex patterns [10]. Moreover, in many cases, they are not easy to use, and operate in a *black-box* fashion with respect to the end users.

On the other hand, if we wish to exploit case based retrieval in a decision support system, we must grant for: *interpretability* of the output results; *understandability* of the retrieval process; and *interactivity* of the retrieval process itself.

To this end, as explained in the Introduction, in this work we propose the use of Temporal Abstractions (TA). TA is an Artificial Intelligence (AI) methodology able to solve a data interpretation task [35], the goal of which is to derive high level concepts from time stamped data. Operatively, the principle of *basic* TA methods is to move from a *point-based* to an *interval-based* representation of the data [4], where the input points (*events*) are the elements of the time series, and the output intervals (*episodes*) aggregate adjacent events sharing a common behavior, persistent over time. Episodes are identified by symbols. Basic abstractions can be further subdivided into *state* TA and *trend* TA. *State* TA are used to extract episodes associated with *qualitative levels* of the monitored feature, e.g. low, normal, high values; *trend* TA are exploited to detect specific *patterns*, such as increase, decrease or stationary behavior, from the time series. Through TA, huge amounts of temporal information can be effectively mapped to a compact symbolic representation, that not only *summarizes* the original longitudinal data, but also *highlights* meaningful behaviors in the data themselves, which can be *interpreted* by end users as well.

Rather interestingly, TA have been extensively resorted to in the literature, especially in the medical field, from diabetes mellitus [36,3], to artificial ventilation of intensive care units patients [22] (see also the survey in [38]), but typically with the aim to solve a data interpretation task [35], and not as a retrieval support facility. The goal of our proposal is to try to fill this gap, by supporting data

interpretation, as well as case exploration and retrieval; this idea thus appears to be significantly innovative in the recent literature panorama.

It is worth noting that TA are not the only methodology for reducing dimensionality by transforming a time series into a sequence of symbols. Actually a wide number of symbolic representations of time series have been introduced in the past decades (see [10] for a survey). However, some of them require an extremely specific and hard to obtain domain knowledge [17], since they a priori partition the signal into intervals, naturally provided by the underlying system dynamics, which divide the overall time period into distinct physical phases (e.g. respiration cycles in [12]). Rather interestingly, Lin [20] has introduced an alternative to TA, capable to deal with such issues, in which intervals are first obtained through Piecewise Constant Approximation [18], and subsequently labeled with proper symbols. In particular this contribution allows distance measures to be defined on the symbolic approach that lower bound the corresponding distance measures defined on the original data. However, the approach in [20] is not as simple as TA, which allows a clear interpretation of the qualitative description of the data provided by the abstraction process itself. As a matter of fact, such a description is often easier to understand for end users (e.g. clinicians [37]), and easily adapts to domains where data values that are considered as normal at one time, or in a given context, may become dangerously abnormal in a different situation (as in medicine, due to disease progression or to treatments obsolescence). And, of course, the ease and flexibility at which knowledge can be managed and understood by experts is an aspect that impacts upon the suitability and the usefulness of decision support systems in practice.

A further approach, based on a database querying tool, has been introduced in [39]; in this work a symbolic query (in the form of string of symbols, like the ones produced by TA) can be answered over a database of raw time series data, by producing those substrings that best match the query itself, following a set of abstraction rules, operating on a symbol taxonomy and on different time granularities. The data structures we introduce (see section 4) are somehow similar to the ones exploited in such paper, however the work in [39] exploits them only to support roll-up and drill-down operations in a data warehouse context, where the query abstraction level determines the level at which the retrieved data have to be transformed. Instead, we provide a more general and flexible retrieval support framework, in which orthogonal index structures optimize the response time. Details will be provided in section 4.

3   The Hemodialysis Domain

We are currently developing an application of our framework to the hemodialysis domain, within which we have a strong background knowledge in our research group (however, it is worth noting that the framework itself is domain independent, and could be adapted to other fields as well).

Hemodialysis is the most widely used treatment for End Stage Renal Disease, a severe chronic condition which, without medical intervention, leads to death.

Hemodialysis relies on a device, called hemodialyzer, which clears the patient's blood from catabolites, to re-establish acid-base equilibrium and to remove water in excess. On average, hemodialysis patients are treated for four hours three times a week. Each single treatment is called a hemodialysis session, during which the hemodialyzer collects several variables, most of which are in the form of time series.

Considering a case as a hemodialysis session, in the next sections we will show examples of case retrieval, in which cases will be characterized by time series features preprocessed by means of TA.

In particular, we will focus on a single case feature, for the sake of clarity: namely, diastolic pressure. Diastolic pressure is a very powerful indicator for evaluating water reduction from the patient's blood during a session. The reduction of water from the blood during the hemodialysis session causes a constant decrease of the blood pressure. This behavior is correct and, even if it can sometimes cause minor problems to the patient (e.g. light head spinning), it is necessary to achieve a good water and metabolites reduction. However, in certain conditions, the reduction of water is not constant, but can be characterized by stationarity periods and (sudden) increasing or decreasing trend episodes. In particular, problems arise when the pressure remains stationary for almost half of the session, which means that no water reduction takes place. Then (sharp) decreasing or increasing episodes can take place, destabilizing the cardiovascular system of the patient. Specifically, clinical studies state that intradialytic increase in diastolic pressure can complicate the management of hypertension in hemodialysis patients. Furthermore, diastolic pressure increasing trend is associated with greater 2-year mortality in these patients [14]. We are thus particularly interested in studying increasing episodes.

4   A Flexible Retrieval Framework

As explained in the Introduction, we have defined a framework for time series retrieval, which exploits the TA methodology to reduce time series dimensionality, and multi-dimensional index structures to make retrieval efficient.

Specifically, we support *multi-level abstractions* of the original data. Time series values can be abstracted (and queried) at finer or coarser detail levels, according to two dimensions: a *symbol taxonomy*, and a *time granularity taxonomy*.

Working e.g. on trends, the taxonomyin figure 1 can be introduced, in which the symbol I (increase) is further specialized into I_W (weak increase) and I_S (strong increase), according to the slope. In our approach the *symbol taxonomy* is a conventional *isa* taxonomy.

An important property of many symbol domains is *ordering*. Such an ordering naturally emerges from the domain dependent interpretation of the underlying row data, from which bottom symbols in the taxonomy have been abstracted. For instance, D_S (strong decrease) may abstract curve portions with slopes from -90 to -45 degrees, thus preceding D_W (weak decrease) (referring e.g. to slopes

from -44 to -10 degrees), in the symbol domain ordering. Henceforth, we will use the symbols $<^d$ and $\leq^d$ to denote (strict/non-strict) precedence in the symbol domain.

Of course, the symbol taxonomy must respect the ordering (see also [5,33]), if any, as stated by the following axiom:

$$\forall x, y, x', y' \in \Sigma \;\; isa(x, x') \wedge isa(y, y') \wedge x' \neq y' \atop \wedge x <^d y \rightarrow x' <^d y' \tag{1}$$

where Σ is the symbol domain, x is a child of x', and y is a child of y' in the isa taxonomy. For instance, if D_W (weak decrease) precedes I_W (weak increase) in the trend symbol ordered domain, then also D (decrease) must precede I (increase).

A *distance* function may be used in order to measure the distance between symbols in the taxonomy. As regards the distance function choice, any one can be selected. We just enforce the straightforward general constraint that the distance of each symbol from itself is zero.

While we do not impose any further constraint on the distance function for unordered symbol domains, we enforce the fact that distance must be "consistent" with ordering (if any). Specifically, distance monotonically increases with ordering, as requested by the following axiom:

$$\forall x, y, z \in \Sigma \;\; x <^d y <^d z \rightarrow d(x, y) < d(x, z) \tag{2}$$

where $d(x, y)$ denotes the distance between two symbols x and y.

For instance, referring to the trend symbol domain, where the ordering is naturally given by the increasing slope values, axiom 2 states that the distance between D and S (stationary) must be smaller than the distance between D and I.

The *time granularity taxonomy*, on the other hand, allows one to describe the episodes at increasingly more abstract levels of temporal aggregation. Obviously, the number of levels and the dimension of granules can be differently set depending on the application domain. In hemodialysis, time granularity can be progressively refined from 4 hours to 30 minutes.

Observe that in our approach the time dimension requires that aggregation is "homogeneous" at every given level, in the sense that each granule at a given level must be an aggregation of exactly the same number of consecutive granules at the lower level (while this number may vary from level to level; for instance, two 30-minutes-long granules compose a 1-hour-long granule, while three 10-minutes-long granules compose a 30-minutes-long granule). Such an "homogeneity" restriction is motivated by the fact that, in such a way, the duration of each episode can be (implicitly) represented by the length of the sequence of symbols (differently from e.g. [4]). For example, at the time granularity level of 10 minutes, the string $IIISDD$ represents a 30 minutes episode of I followed by 10 minutes of S and 20 minutes of D. This choice is also useful for indexing purposes.

In order to abstract along the temporal dimension, a function *up* for *scaling up* from two or more values expressed at a specific level in the taxonomy to a

single value expressed at the coarser level must be provided. Abstracting from one granularity to a coarser one is a domain-dependent procedure. In order to retain the maximal generality, our framework allows one to freely define the rule. Once again, however, we impose some general constraints, to grant for the meaningfulness of the function and for its "consistency" with respect to the other knowledge sources. The following axiom grants the fact that up preserves "persistence": the result of coarsening two granules with the same symbol x is a larger granule still labeled as x (here and in the following, for the sake of simplicity and brevity we apply the up function to two granules, but the definitions can be generalized to n-ary up operators):

$$\forall x \in \Sigma \quad up(x,x) = x \tag{3}$$

where Σ is the symbol domain, and $up(x,x)$ denotes the symbol obtained by abstracting two adjacent intervals, both labeled with the same symbol x, at a coarser time granularity.

On the other hand, the following axioms state the relationships between ordering and up, enforcing a sort of "monotonicity": in some sense, they state that ordering is preserved by the up function. In particular:

$$\forall x,y \in \Sigma \quad x <^d y \rightarrow x \leq^d up(x,y) \leq^d y \tag{4}$$

Moreover:

$$\forall x,y,z \in \Sigma \quad x <^d y <^d z \rightarrow up(x,y) \leq^d up(x,z) \tag{5}$$

and

$$\forall x,y,z \in \Sigma \quad x <^d y <^d z \rightarrow up(x,z) \leq^d up(y,z) \tag{6}$$

Given such axioms, some unclear (or, more precisely, meaningless) situations are automatically ruled out. For instance, it can never happen that, if a 1-hour-long episode of D, followed by a 1-hour-long episode of I, abstracts to a 2-hours-long episode of D, it also happens that a 1-hour-long episode of D, followed by a 1-hour-long episode of S, abstracts to a 2-hours-long episode of S.

It is worth stressing that the axioms above code the relationships between the taxonomies and the distance and up functions. As a consequence, the combination of such axioms also fixes the constraints between any "combination" of such primitive notions. For instance, axioms (1) and (2) state that distance "preserves" ordering also in case *isa* relationships between symbols are involved (e.g. the distance between D and S is smaller than the distance between D and I_S).

In order to increase *retrieval efficiency*, in our framework we have also implemented an indexing strategy, exploiting taxonomic indexes built on the basis of the two taxonomies mentioned above.

In particular, we resort to a forest of index structures, providing a flexible indexing of cases at different levels of the symbol and/or time granularity taxonomies. The root node of each index structure is represented by a string of symbols, defined at the highest level in both dimensions. An example, taking as
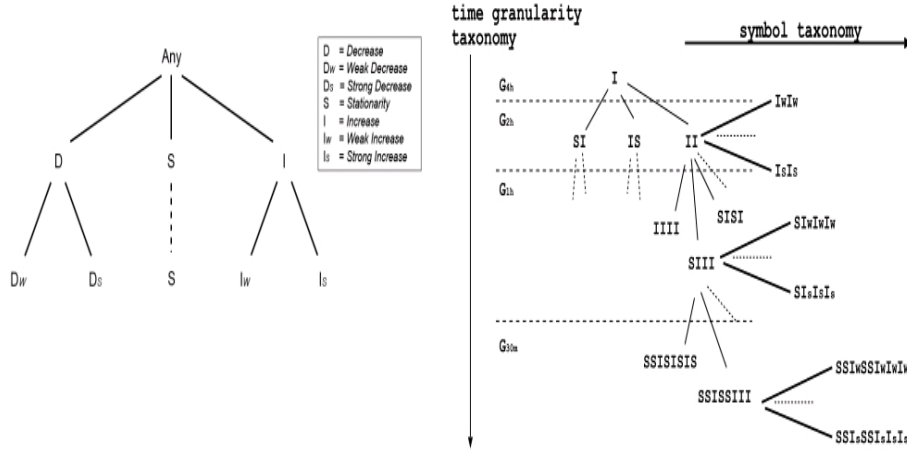
Fig. 1. An example symbol taxonomy (left). An example multi-level orthogonal index structure (right). The symbol taxonomy is articulated into two levels (where e.g. I stands for increase, I_w and I_s for weak increase and strong increase respectively; S stands for stationarity - and cannot be further specialized). Time granularity can be progressively refined from 4 hours to 30 minutes.

a root the I (increasing trend) symbol, is provided in figure 1. Here, the root node I is refined along the time dimension (which is taken as the leading dimension) from the 4 hours to the 2 hours granularity, so that the nodes II, IS and SI stem from it, provided that $up(I, S) = I$ and $up(S, I) = I$. From each node of the leading dimension structure, another index can stem, built according to the symbol taxonomy dimension, keeping the time granularity abstraction level always fixed. Such an index develops orthogonally with respect to the leading dimension. Observe that indexes may be incomplete, but can be refined on demand.

In summary, time series features in our cases are processed by means of TA, executed at the ground level according both to the symbol taxonomy and to the time granularity one.

Queries on such data can then be issued at any level of detail (see also [9]), according to both dimensions.

For instance, query (7) below abstracts the negative situation introduced in section 3, in which pressure remains stationary in the initial part of the session, and then weak increasing episodes take place (notice that each symbol represents a 1-hour-long episode - thus globally covering the overall 4 hours duration).

$$SI_W I_W I_W \qquad (7)$$

Indexes allow for a quick (and interactive) query answering.

Specifically, to answer a query, in order to enter the index structure, we first progressively generalize the query itself in the non leading dimension. Then,

we generalize the query in the other dimension as well. Generalization steps are then used backwards for index navigation. The index node corresponding to the query, or corresponding to the most specific generalization of the query currently available (if the index is incomplete) is then returned. Finally, all the cases indexed by such a node are shown to the user.

We will now exemplify the query answering process with reference to the diastolic pressure feature. Consider query (7) above. In this example, the query generalization in the direction of the symbol taxonomy generates the sequence $SIII$; starting from the latter, the generalization in time generates the sequences: II (2-hours-long episodes), and then I (4-hours-long episode). The output of the generalization process allows to identify a single index structure in the forest, namely the one whose root is I (i.e. the tree shown in figure 1) as a support for a quick query answering. Matching the steps in the generalization process to the nodes in the index structure (in the time direction), we can descend through the nodes II, and then $SIII$. Now, we can move "horizontally" in the symbol taxonomy direction, to reach the node $SI_W I_W I_W$, which matches exactly our query. As a result, we can retrieve *all* and *only* the cases indexed by such a node.

The query $SI_W I_W I_W$ reflects an uncommon situation to be investigated; therefore, a limited number of cases are retrieved. Users may want to generalize the required behavior, in order to retrieve a larger number of cases. For instance, a user can consider episodes where one hour of stationary pressure is followed by one hour of weak increase, and then by two hours of (weak or strong) increase. Thus, all cases indexed by nodes $SI_W I_S I_W$, $SI_W I_W I_S$, and $SI_W I_S I_S$ (in addition to $SI_W I_W I_W$) can be retrieved, clarifying that their distance from the original query is greater than zero.

Query relaxation (as well as refinement) can be repeated several times, until the user is satisfied with the obtained results.

5   Experimental Results

We executed some experiments in the hemodialysis domain, using a dataset containing data belonging to a set of real patients, for a total of 10388 hemodialysis sessions (i.e. cases), collected at the Vigevano hospital, Italy.

In particular, we compared the approach described in this paper with a more classical approach we implemented in the past within the system RHENE [25]. That approach was based on DFT for dimensionality reduction, and on spatial indexing (through TV-trees) for further improving retrieval performances.

We first completed some quantitative experiments, meant to compare the query answering time required by the two approaches, and their scalability when dealing with a case base progressively growing in size. Specifically, we worked in the range 2000-10388 cases (taking subsets of the totally available real cases).

As shown in figure 2, the TA-based method proved to be extremely more efficient in query answering than the mathematical method. In particular, as expected, the mathematical method was strongly penalized by the post-processing step (see dashed curve in figure 2 - top, which represents the total query time,

including post processing), needed in order to filter out false alarms [2], while TV-tree-based indexing time was sublinear (see dotted line in figure 2 - top, which represents indexing time alone). Index creation times, not shown for space constraints, largely favored the TA-based method as well.
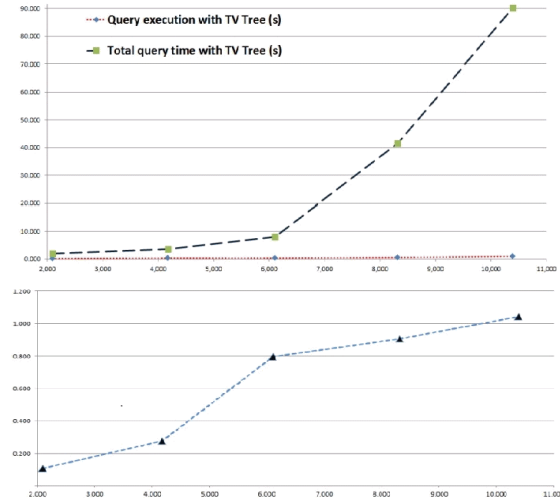

Fig. 2. Quantitative query answering results when using DFT (top) and when using TA (bottom). Performances were obtained on different database sizes (x axis), by calculating the average time required for answering a query (y axis).

We also conducted several experiments, comparing DFT with our TA-based methodology, operating both on trend and on state TA. Experimental results cannot be reported in detail, due to space constraints. However, we would like to summarize the observations we were able to draw:

- TA-based retrieval allows to focus either on trend or on state behavior separately, providing two different retrieval sets, related to the two different viewpoints from which a specific situation can be analyzed. Such a possibility is not available when resorting to DFT, which is unable to separate the two aspects;
- DFT basically works on point-to-point distance between the query and the retrieved time series, which means that we are searching for the best overall alignment of the resulting case over the plot of the query. On the other hand, DFT is unable to specifically consider trend behavior. This is thus an added value provided by TA-based retrieval;
- in DFT retrieval, details matter. With TA, on the other hand, we are able to abstract from unnecessary details, ignoring e.g. (strong) oscillations which do not change the qualitative value of the function. This is the reason why, even

though some cases could be retrieved both by DFT and by state TA-based retrieval, others could not;
– considering the dynamics of the actions performed to answer a query, it is clear that mathematical methods work in a black box fashion. The user selects the query case, sets the parameters of the query (e.g. the value of K for K-Nearest Neighbor, or a maximum distance for a range query), and then the system performs its retrieval task. Using our TA-based method, instead, the user can drive the query in an interactive fashion, taking control of the retrieval task, to find the proper level of generalization or refinement of the query;
– the results of the query performed using our TA-based method are easily interpretable compared to the results obtained by mathematical methods. For each retrieved case, we can obtain its distance from the query, as well as its representation as a sequence of abstractions. This sequence gives an immediate understanding of the shape of the result; therefore, in most cases it is not necessary to look at the original plot.

In conclusion, the use of TA allows the user to perform a flexible and interactive retrieval, with easy-to-understand queries and results. Furthermore, the use of TA allows to explicitly take into account trend information, and to abstract from unnecessary details, thus potentially providing qualitatively better results.

6   Conclusion

In this paper, we have presented a methodology for supporting time series retrieval, in which time series dimensionality is reduced by means of TA. Specifically, we support multi-level abstractions of TA-based time series, both along the time dimensions, and along the symbol taxonomy one, thus increasing the *flexibility* of the retrieval facility, both in query definition and in query answering. Similar cases can be retrieved *efficiently* and *interactively*, relying on proper orthogonal index structures, which can grow on demand (and on a user-friendly graphical interface). In our opinion, flexibility and interactivity represent a relevant advantage of our approach to time series retrieval with respect to more classical techniques, in which end users are typically unable to intervene in the retrieval process, that often operates in a black-box fashion. This observation is particularly relevant for medical applications.

The framework is currently being applied to the hemodialysis domain. Experimental results have clearly shown that our framework is more user-friendly, thanks to the simplicity in query issuing, to the understandability of results, and to its interactivity, with respect to a more classical mathematical approach. The experiments have also shown that TA can provide qualitatively better results when resorting to trend abstractions (by looking for a similar trend behavior, abstracting from quantitative values).

In our current implementation, the user is forced to express only queries that cover the overall temporal duration of the time series. Indeed, substring alignment and matching are not supported. On the other hand, we believe that

substring matching would be a significant improvement in time series retrieval, especially when dealing with medical applications. To this end, in the future we will design a new version of the framework, in order to properly address this additional issue, and, generally speaking, to support the search and the indexing of more complex patterns, also through rule-based abstraction mechanisms like the ones described in [35].

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations and systems approaches. AI Communications 7, 39–59 (1994)
2. Agrawal, R., Faloutsos, C., Swami, A.: Efficient Similarity Search in Sequence Databases. In: Lomet, D.B. (ed.) FODO 1993. LNCS, vol. 730, pp. 69–84. Springer, Heidelberg (1993)
3. Bellazzi, R., Larizza, C., Magni, P., Montani, S., Stefanelli, M.: Intelligent analysis of clinical time series: an application in the diabetes mellitus domain. Artificial Intelligence in Medicine 20, 37–57 (2000)
4. Bellazzi, R., Larizza, C., Riva, A.: Temporal abstractions for interpreting diabetic patients monitoring data. Intelligent Data Analysis 2, 97–122 (1998)
5. Bergmann, R., Stahl, A.: Similarity Measures for Object-Oriented Case Representations. In: Smyth, B., Cunningham, P. (eds.) EWCBR 1998. LNCS (LNAI), vol. 1488, pp. 25–36. Springer, Heidelberg (1998)
6. Bichindaritz, I., Conlon, E.: Temporal knowledge representation and organization for case-based reasoning. In: Proc. TIME 1996, pp. 152–159. IEEE Computer Society Press, Washington, DC (1996)
7. Branting, L.K., Hastings, J.D.: An empirical evaluation of model-based case matching and adaptation. In: Proc. Workshop on Case-Based Reasoning, AAAI 1994 (1994)
8. Chan, K.P., Fu, A.W.C.: Efficient time series matching by wavelets. In: Proc. ICDE 1999, pp. 126–133. IEEE Computer Society Press, Washington, DC (1999)
9. Combi, C., Pozzi, G., Rossato, R.: Querying temporal clinical databases on granular trends. Journal of Biomedical Informatics 45(2), 273–291 (2012)
10. Daw, C.S., Finney, C.E., Tracy, E.R.: Symbolic analysis of experimental data. Review of Scientific Instruments (July 22, 2002) (2001)
11. Fuch, B., Mille, A., Chiron, B.: Operator Decision Aiding by Adaptation of Supervision Strategies. In: Aamodt, A., Veloso, M.M. (eds.) ICCBR 1995. LNCS (LNAI), vol. 1010, pp. 23–32. Springer, Heidelberg (1995)
12. Funk, P., Xiong, N.: Extracting Knowledge from Sensor Signals for Case-Based Reasoning with Longitudinal Time Series Data. In: Perner, P. (ed.) Case-Based Reasoning in Signals and Images. SCI, vol. 73, pp. 247–284. Springer, Heidelberg (2008)
13. Hetland, M.L.: A survey of recent methods for efficient retrieval of similar time sequences. In: Last, M., Kandel, A., Bunke, H. (eds.) Data Mining in Time Series Databases. World Scientific, London (2003)
14. Inrig, J.K., Patel, U.D., Toto, R.D., Szczech, L.A.: Association of blood pressure increases during hemodialysis with 2-year mortality in incident hemodialysis patients: A secondary analysis of the dialysis morbidity and mortality wave 2 study. American Journal of Kidney Diseases 54(5), 881–890 (2009)

15. Jaczynski, M.: A framework for the management of past experiences with time-extended situations. In: Proc. ACM Conference on Information and Knowledge Management (CIKM 1997), pp. 32–38. ACM Press, New York (1997)
16. Jære, M.D., Aamodt, A., Skalle, P.: Representing Temporal Knowledge for Case-Based Prediction. In: Craw, S., Preece, A. (eds.) ECCBR 2002. LNCS (LNAI), vol. 2416, pp. 174–188. Springer, Heidelberg (2002)
17. Kadar, S., Wang, J., Showalter, K.: Noise-supported travelling waves in sub-excitable media. Nature 391, 770–772 (1998)
18. Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S.: Dimensionality reduction for fast similarity search in large time series databases. Knowledge and Information Systems 3(3), 263–286 (2000)
19. Keravnou, E.T.: Modeling Medical Concepts as Time Objects. In: Wyatt, J.C., Stefanelli, M., Barahona, P. (eds.) AIME 1995. LNCS (LNAI), vol. 934, pp. 67–90. Springer, Heidelberg (1995)
20. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proc. of ACM-DMKD, San Diego (2003)
21. Ma, J., Knight, B.: A Framework for Historical Case-Based Reasoning. In: Ashley, K.D., Bridge, D.G. (eds.) ICCBR 2003. LNCS (LNAI), vol. 2689, pp. 246–260. Springer, Heidelberg (2003)
22. Miksch, S., Horn, W., Popow, C., Paky, F.: Utilizing temporal data abstractions for data validation and therapy planning for artificially ventilated newborn infants. Artificial Intelligence in Medicine 8, 543–576 (1996)
23. Montani, S., Bottrighi, A., Leonardi, G., Portinale, L.: A cbr-based, closed loop architecture for temporal abstractions configuration. Computational Intelligence 25(3), 235–249 (2009)
24. Montani, S., Portinale, L.: Accounting for the temporal dimension in case-based retrieval: a framework for medical applications. Computational Intelligence 22, 208–223 (2006)
25. Montani, S., Portinale, L., Leonardi, G., Bellazzi, R., Bellazzi, R.: Case-based retrieval to support the treatment of end stage renal failure patients. Artificial Intelligence in Medicine 37, 31–42 (2006)
26. Nakhaeizadeh, G.: Learning Prediction from Time Series: A Theoretical and Empirical Comparison of CBR with some other Approaches. In: Wess, S., Richter, M., Althoff, K.-D. (eds.) EWCBR 1993. LNCS (LNAI), vol. 837, pp. 65–76. Springer, Heidelberg (1994)
27. Nilsson, M.: Retrieving short and dynamic biomedical sequences. In: Proc. 18th International Florida Artificial Intelligence Research Society Conference–Special Track on Case-Based Reasoning. AAAI Press (2005)
28. Nilsson, M., Funk, P., Olsson, E., von Scheele, B., Xiong, N.: Clinical decision-support for diagnosing stress-related disorders by applying psychophysiological medical knowledge to an instance-based learning system. Artificial Intelligence in Medicine 36, 159–176 (2006)
29. Nilsson, M., Funk, P., Xiong, N.: Clinical decision support by time series classification using wavelets. In: Chen, C.S., Filipe, J., Seruca, I., Cordeiro, J. (eds.) Proc. Seventh International Conference on Enterprise Information Systems (ICEIS 2005), pp. 169–175. INSTICC Press (2005)
30. Palma, J., Juarez, J.M., Campos, M., Marin, R.: A fuzzy approach to temporal model-based diagnosis for intensive care units. In: Lopez de Mantaras, R., Saitta, L. (eds.) Proc. European Conference on Artificial Intelligence (ECAI 2004), pp. 868–872. IOS Press, Amsterdam (2004)

31. Portinale, L., Montani, S., Bottrighi, A., Leonardi, G., Juarez, J.: A case-based architecture for temporal abstraction configuration and processing. In: Proc. IEEE International Conference on Tools with Artificial Intelligent (ICTAI), pp. 667–674. IEEE Computer Society, Los Alamitos (2006)
32. Ram, A., Santamaria, J.C.: Continuous case-based reasoning. In: Proc. AAAI Case-Based Reasoning Workshop, pp. 86–93 (1993)
33. Resnik, P.: Using information content to evaluate semantic similarity in a taxonomy. In: Proc. IJCAI, pp. 448–453 (1995)
34. Rougegrez, S.: Similarity Evaluation Between Observed Behaviours for the Prediction of Processes. In: Wess, S., Richter, M., Althoff, K.-D. (eds.) EWCBR 1993. LNCS (LNAI), vol. 837, pp. 155–166. Springer, Heidelberg (1994)
35. Shahar, Y.: A framework for knowledge-based temporal abstractions. Artificial Intelligence 90, 79–133 (1997)
36. Shahar, Y., Musen, M.A.: Knowledge-based temporal abstraction in clinical domains. Artificial Intelligence in Medicine 8, 267–298 (1996)
37. Stacey, M.: Knowledge based temporal abstractions within the neonatal intesive care domain. In: Proc. CSTE Innovation Conference, University of Western Sidney (2005)
38. Terenziani, P., German, E., Shahar, Y.: The temporal aspects of clinical guidelines. In: Ten Teije, A., Miksch, S., Lucas, P. (eds.) Computer-based Medical Guidelines and Protocols: A Primer and Current Trends (2008)
39. Xia, B.B.: Similarity search in time series data sets. Technical report, School of Computer Science, Simon Fraser University (1997)