

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300

Open access books available

131,000

International authors and editors

160M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Temporal Knowledge Generation for Medical Procedures

Aida Kamišalić<sup>1</sup>, David Riaño<sup>2</sup> and Tatjana Welzer<sup>1</sup>

<sup>1</sup>University of Maribor

<sup>2</sup>Rovira i Virgili University

<sup>1</sup>Slovenia

<sup>2</sup>Spain

## 1. Introduction

Decision support systems (DSSs) in medicine are designed to aid medical professionals on making clinical decisions about prevention, diagnosis and corresponding treatment. When DSSs are applied to medical procedures, two sorts of predictions are possible: procedural (i.e. indications on what to do), and temporal (i.e. indications on what are the time restrictions). Clinical Practice Guidelines (CPGs) are statements that assist physicians making appropriate medical decisions during patient encounters. They are a set of assertions used to manage patients with a particular disease to improve quality of care, decrease unjustified practice variations and save costs. Clinical algorithms (CAs) obtained from CPGs are introduced to make the procedural knowledge explicit and formal. It is important to enable the latest clinical knowledge to be accessible and usable at the point of care, and therefore make significant contributions to safety and quality in medicine. Medical knowledge is used to assist patients suffering from one or several diseases. CAs could be explicitly given, or obtained with a knowledge management mechanisms. Among these mechanisms, there are some that aim at generating CAs from existing patients' data for a particular disease. However, either explicitly given or generated CAs are atemporal, which means that there is no an explicit time labelling of the elements in the CA. Time plays a major role in medicine and therefore also in medical information systems. It is an important concept of the real world, which needs to be managed in different ways (events occur at some time points, facts hold during time periods, temporal relationships exist between facts and events) (Combi et al., 2010). If we want to overcome the gap of atemporal CAs it is necessary to define a time dimension and make also temporal knowledge (the indications on what are the time restrictions) explicit and formal. It has been proved that obtaining explicit temporal knowledge from physicians is often a difficult and time-consuming task regardless of the knowledge engineering mechanisms or tools employed to simplify the process. As data saved in hospital databases are primarily time dependent, they can be used to obtain temporal constraints to define the time dimension of CAs. We have propose generation of temporal constraints considering patients' data of a particular disease for atemporal CAs. We have defined two types of temporal constraints: macro-temporality and micro-temporality. Macro-temporality is defined as a constraint  $[t_{\min}, t_{\max}]$  on the time required to cross a particular edge of a CA, where  $t_{\min}$  and  $t_{\max}$  are the lower and the upper

bounds of the time-lag, respectively (Kamišalić et al., 2007). Macro-temporality denotes time delays which have to be fulfilled before the treatment of the patient proceeds. For example,  $[1d, 4d]$  assigned to the edge between two actions, would mean that after applying the first action the treatment should wait a minimum of one day and a maximum of four days before proceeding with the next action (Kamišalić et al., 2007). The concept of micro-temporality is defined as a constraint  $[s_t, e_t, f_t]$  on the start time  $s_t$ , the end time  $e_t$ , and the frequency of occurrence  $f_t$  of some medical action (Kamišalić et al., 2007). For example, the term 'prehypertension' can be considered as a state term, where  $[1w, 3d, 6h]$  constraint means that 'prehypertension' was part of the patient condition since one week ago ( $1w$ ), till three days ago ( $3d$ ), and it was observed every six hours ( $6h$ ). In the case of action term, micro-temporality means that the action must start after time  $s_t$ , that should last till  $e_t$ , and that the application of the action has a frequency of  $f_t$  (for example, take beta-blocker agent for two weeks every six hours  $[-, 2w, 6h]$ ).

There are several formalisms which can be used to represent CAs. Some of them are very complete from a medical point of view and therefore difficult to manage by untrained physicians (e.g. Asbru (Seyfang et al., 2002) or EON (Musen et al., 1996)), some are oriented to the description of medical activities as processes rather than as an explicit representation of the treatment as a procedure (e.g. Asbru), and some others are limited in the representation of time constraints (e.g. PROforma (Fox et al., 1998)). Whenever CAs comprise the concept of state as a description of a medical situation, they can be represented with State Transition Diagrams (STD's) (see Fig. 1) where each possible individual transition between two states is represented. As the levels of simplicity, intuitiveness, and time representation of the SDA\* formalism (Riaño, 2007) (see Fig. 2) were close to our requirements, this one has been the chosen formalism for representing macro- (Kamišalić et al., 2007) and micro-temporalities (Kamišalić et al., 2009) in CAs. We have also used Timed-transition diagrams (TTD's) (see Fig. 4, Fig. 5, Fig. 6) as representation formalism for macro-temporality constraints (Kamišalić et al., 2008).

The rest of the chapter is divided into 4 main sections. Section 2 provides an overview and examples of used representation formalisms - SDA\*s and TTDs. They are used for representation of time dimension in CAs. The process of generation of macro-temporalities from clinical data is described in the section 3. Section 4 describes a methodology for generation of micro-temporalities. Finally, section 5 gathers some conclusions.

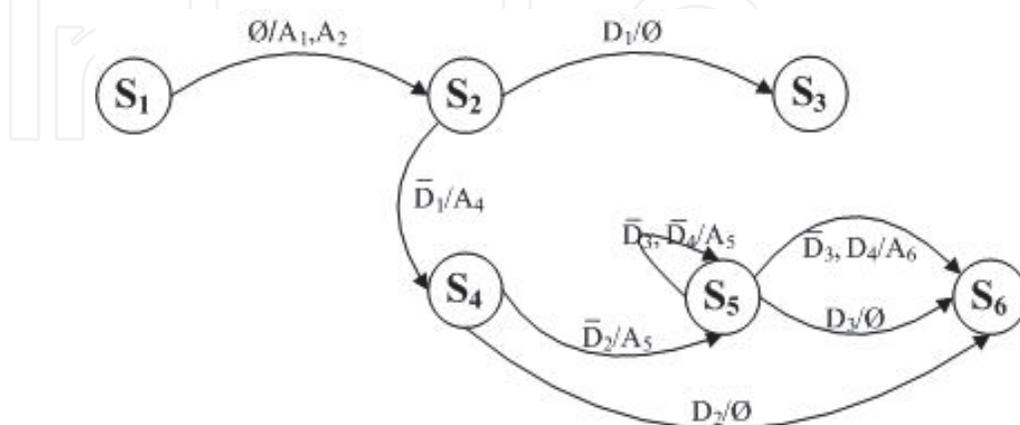


Fig. 1. Clinical algorithm for Hypertension in STD representation

## 2. Representation formalisms

As we have mentioned above, there are different formalisms which can be used to represent CAs. STDs can be used for representation of timeless CAs. As our goal is to introduce temporal constraints to CAs, the representation formalism have to permit representation of temporal restrictions for CAs. TTDs are used to represent CAs with generated macro-temporality constraints. These diagrams permit the integration of several sequences in a single structure, each sequence representing the individual treatment of one patient (Kamišalić et al., 2008). (Sections 3 and 4 introduce processes of macro- and micro-temporality generation, respectively.) SDA\* formalism is not convenient for representation of individual sequences. It can be used for representation of macro- and micro-temporality restrictions, but not for representation of all possible transitions from one to another state. When it is important to represent all possible transitions (sequences) from which generation of macro-temporality restriction is done, TTDs are used for CAs representation. Finally, obtained macro-temporality constraints as well as micro-temporality constraints can be introduced to SDA\* diagrams as it results more simple and intuitive for interpretation. To represent time dimension in CAs we are using TTD's and SDA\* formalism.

### 2.1 SDA\* formalism

SDA stands for state-decision-action notation and it is used to represent CAs as SDA\* diagrams, where each diagram is a directed graph that contains state, decision and action nodes (see Fig. 2). Each state node contains the set of terms which represent the signs and symptoms of a particular patient at the moment of making observations. In the diagram, state nodes are indicated as circles, which bring the feasible patient medical conditions. Decision nodes provide criteria to derive the treatment in one or other direction according to the patient's features and current condition. These branching points are represented as diamonds that deploy all the alternative clinical treatments at this point, each one attached to a condition that the patient must fully satisfy in order to follow this treatment. Action nodes represent the activities which should be taken as a result of an earlier made decision. It is represented as a square, which gives recommendations about medical orders considering medication and clinical procedures.

For each disease  $D$ ,  $T_D = \{t_1, t_2, \dots, t_n\}$  represents a set of terms related to  $D$ . These terms can be state, decision or action terms depending on whether they are able to be involved in the description of a SDA\* state, decision or action, respectively.  $SV \subset T_D$  is the set of state terms that are used to determine the condition of a patient considering some disease.  $DV \subset T_D$  is the set of decision terms that can be used to derive the treatment to some or some other medical actions.  $AV \subset T_D$  is the set of action terms representing the individual medical actions a physician may prescribe in the treatment of the disease the SDA\* diagram represents. In a SDA\* diagram the meaning of node connectors can vary according to the sort of nodes the connectors have as starting and ending elements. So,  $C_{SS}$ ,  $C_{SD}$ , and  $C_{SA}$  are the names given to connectors going from a state node to a state node, to a decision node or to an action node, respectively;  $C_{DS}$ ,  $C_{DD}$ , and  $C_{DA}$  are the names of the nodes going from decision nodes to other nodes, and  $C_{AS}$ ,  $C_{AD}$ ,  $C_{AA}$  the names of the connectors that start at an action node.

Macro-temporality constraints in the form of  $[t_{\min}, t_{\max}]$  intervals affect to some of the above sort of connectors:  $C_{SS}$ ,  $C_{AS}$ ,  $C_{AD}$ , and  $C_{AA}$ ; but not to the rest, for which they are defined as the constant  $[0, 0]$  (i.e. lack of delay or instantaneous transition).

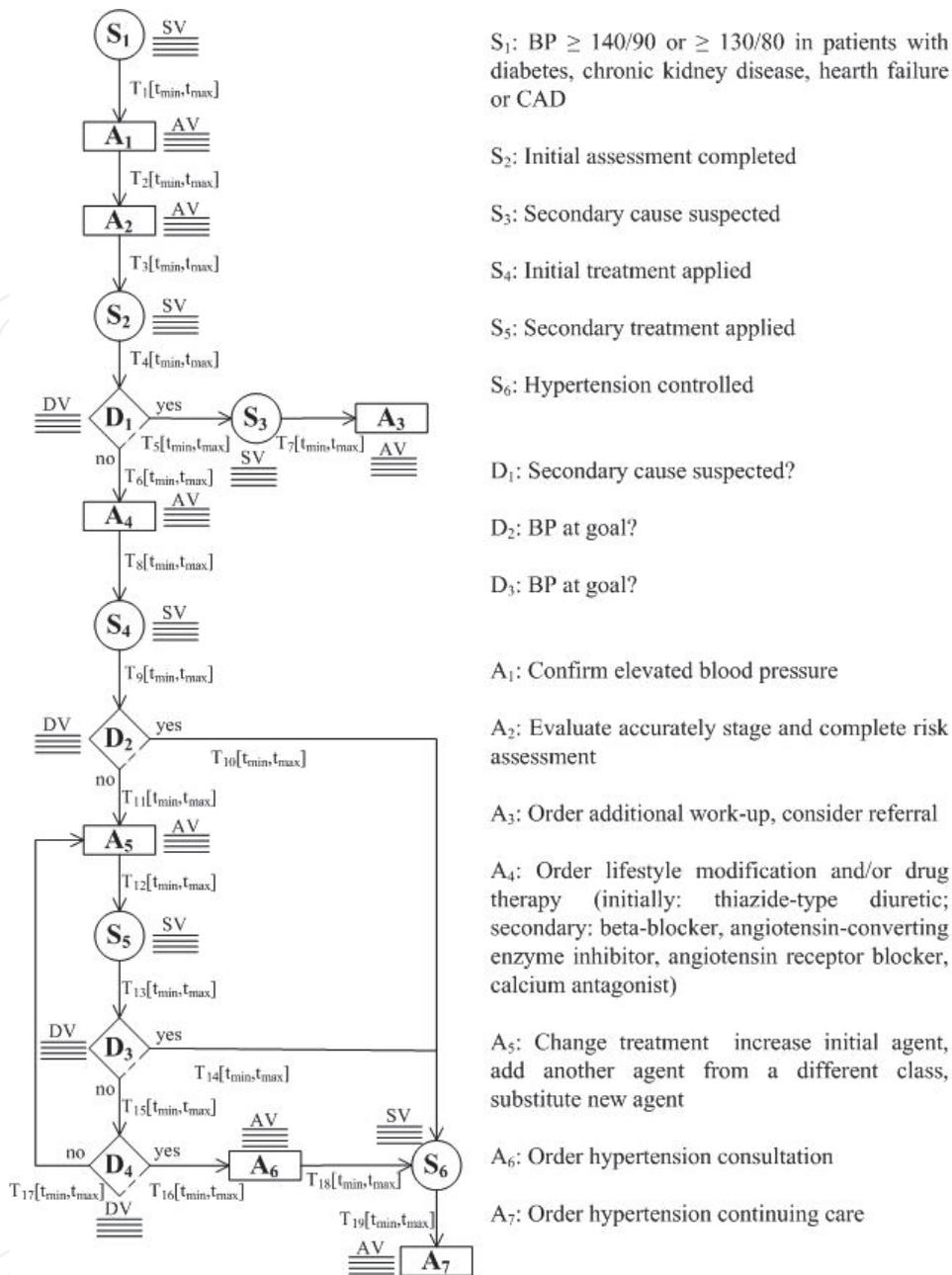


Fig. 2. Clinical algorithm for Hypertension in SDA\* representation

In  $C_{SS}$  connectors, macro-temporality is obtained from the times between consecutive encounters in which the patient has not received any treatment as a consequence of the first encounter. In  $C_{AS}$  connectors, macro-temporality is calculated from the times between consecutive encounters in which, during the first encounter the physician ordered the actions A, and in the second encounter the patient had evolved to state S. Finally in  $C_{AD}$  and  $C_{AA}$  connectors, macro-temporality is a combination of the times between consecutive encounters in which a treatment was proposed during the first visit and the state of the patient in the second visit is implicit, i.e. lacking of medical interest. The difference in the macro-temporalities of  $C_{AD}$  and  $C_{AA}$  connectors is in the fact that the first one is applied when not all the patients in encounters with the same implicit state receive the same state-caused

treatment. On the contrary, if all the patients that show sequences of state transitions in which the same implicit state (i.e. the state description of the patient during the encounter) is always followed by the same state-caused treatments, then a decision node will not be needed in the SDA\* diagram that will connect two consecutive action nodes.

Micro-temporality is assigned to the terms in  $T_D$  and it affects states, actions and decisions. Terms that represent the signs and symptoms of a particular patient at the moment of making an observation are called state terms. Decision terms in a CA are building decision criteria which lead the treatment in a specific direction considering the patient's current signs and symptoms, while action terms represent medical activities which should be performed as a result of an earlier analysis of the medical context (e.g., patient condition or hospital resources). Micro-temporality constraints is represented as a triplet  $[s_t, e_t, f_t]$  on the term  $t \in T_D$ ; where  $s_t$  stands for the start time of  $t$ ,  $e_t$  for the ending time, and  $f_t$  for the frequency. A term is atemporal if  $[-,-,-]$  micro-temporality is assigned, as presented in the Fig. 3. From temporal point of view it means that term  $t$  is present now. Considering temporal aspects, state and decision terms are past-to-present terms, while action terms are present-to-future terms. It means that state and decision terms have start time constraint  $s_t$  in the past, while end time constraint  $e_t$  can be in the past or in the present. Action terms have start time constraint  $s_t$  in the present or in the future, while the end time constraint  $e_t$  is in the future. Considering possible combinations of time constraints ( $s_t$ ,  $e_t$  and  $f_t$ ) existence in a micro-temporality, there are detected different modes of micro-temporalities, as presented in the Fig. 3. The first micro-temporality mode  $[s_t,-,-]$  includes start time constraint  $s_t$ , but excludes end time  $e_t$  and frequency  $f_t$  constraints, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present since start time  $s_t$  till now, without having information about the frequency of occurrence  $f_t$ . In the case of action terms it means that specific term will be present from start time  $s_t$ , but there is no information till when ( $e_t$ ) and with which frequency ( $f_t$ ) this term will be present. The second micro-temporality mode  $[-,e_t,-]$  includes end time constraint  $e_t$ , but excludes start time  $s_t$  and frequency  $f_t$  constraints, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present till end time  $e_t$ , without having information when it started ( $s_t$ ) and with which frequency  $f_t$  it occurred. In the case of action terms it means that specific term is present from now till end time  $e_t$ , without having information about the frequency constraint  $f_t$ . The third micro-temporality mode  $[-,-,f_t]$  includes frequency constraint  $f_t$ , but excludes start time  $s_t$  and end time  $e_t$  constraints, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present till now with frequency  $f_t$ , without having information when it started ( $s_t$ ). In the case of action terms it means that specific term is present from now with frequency  $f_t$ , without having information when it will end ( $e_t$ ).

The fourth micro-temporality mode  $[s_t,e_t,-]$  includes start time  $s_t$  and end time  $e_t$  constraints but excludes frequency  $f_t$  constraint, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present since start time  $s_t$  till end time  $e_t$ , without having information about the frequency of occurrence  $f_t$ . In the case of action terms it means that specific term will be present from start time  $s_t$  till end time  $e_t$ , without information about frequency  $f_t$ . The fifth micro-temporality mode  $[s_t,-,f_t]$  includes start time  $s_t$  and frequency  $f_t$  constraints, but excludes end time  $e_t$  constraint, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present from start time  $s_t$  till now with the frequency  $f_t$ . In the case of action terms it means that specific term will be present from start time  $s_t$  with the frequency  $f_t$ , without having information about the end time  $e_t$ . The sixth micro-temporality mode  $[-,e_t,f_t]$  includes end time  $e_t$  and frequency  $f_t$  constraints, but

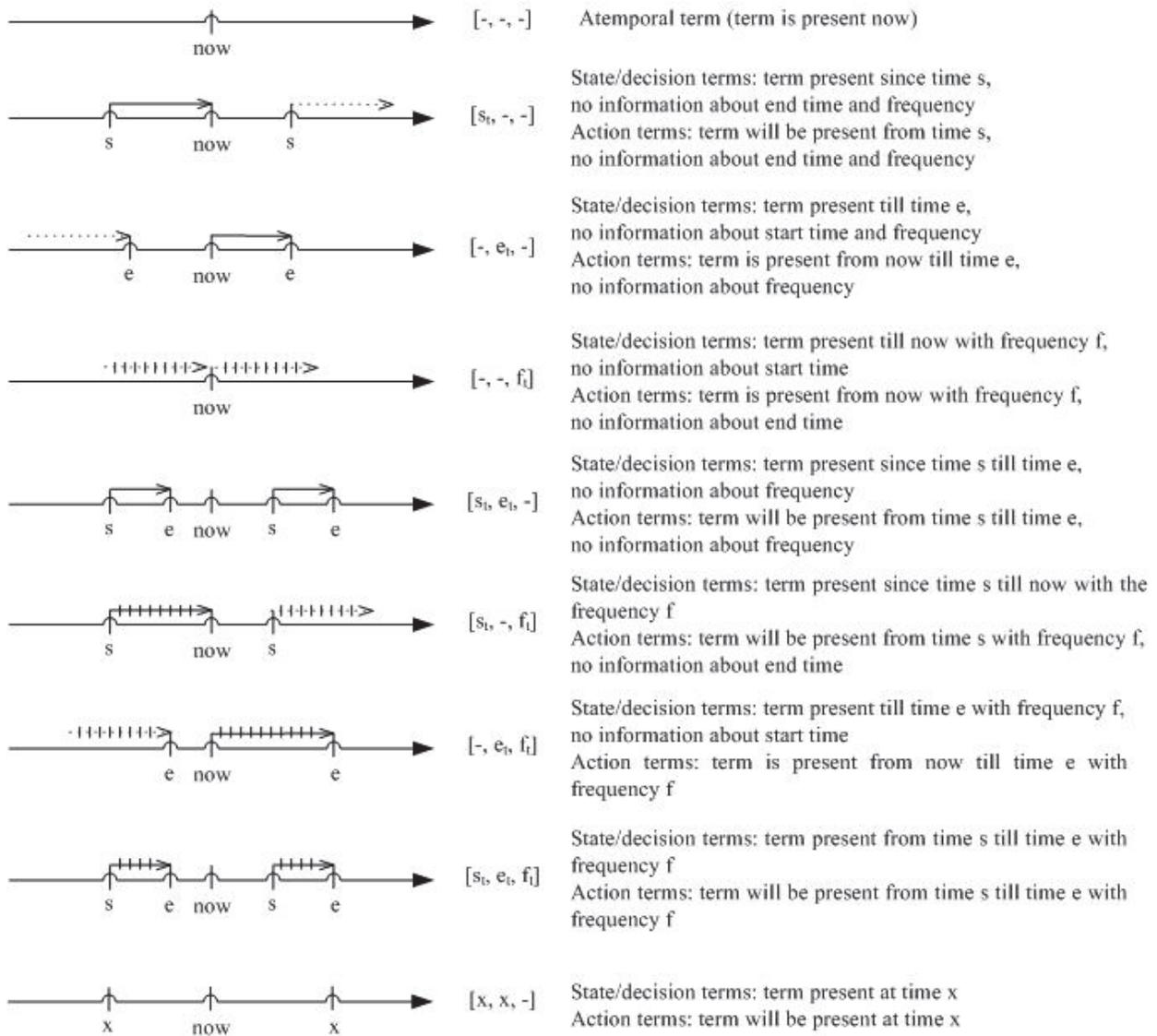


Fig. 3. Different micro-temporality modes

excludes start time  $s_t$  constraint, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present till end time  $e_t$  with frequency  $f_t$ , without having information when it started ( $s_t$ ). In the case of action terms it means that specific term is present from now till end time  $e_t$  with frequency  $f_t$ . The seventh micro-temporality mode  $[s_t, e_t, f_t]$  includes start time  $s_t$ , end time  $e_t$  and frequency  $f_t$  constraints, as shown in the Fig. 3. In the case of state or decision terms it means that specific term was present since start time  $s_t$  till end time  $e_t$  with the frequency of occurrence  $f_t$ . In the case of action terms it means that specific term will be present from start time  $s_t$  till end time  $e_t$  with the frequency  $f_t$ . The ninth micro-temporality mode  $[x, x, -]$  is specific as its start and end constraints are the same and there is no frequency constraint included, as shown in the Fig. 3. It means that term occurred at certain point in time. In the case of state or decision terms it means that term was present at time  $x$ . In the case of action terms it means that term will be present at time  $x$ .

**2.2 Timed-transition diagrams**

A Timed Transition System (TTS) (Herzinger et al., 1992) is a quintuple  $\langle V, \Sigma, T, l, u \rangle$  where  $V$  is a finite set of terms  $V = \{v_1, v_2, \dots, v_n\}$ ;  $\Sigma$  is a set of states  $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$  where every state  $\sigma_i \in \Sigma$  is a subset of terms (i.e.  $\Sigma \subseteq 2^V$ );  $T$  is a finite set of transitions where every transition  $t \in T$  is a binary relation on  $\Sigma$ ;  $l : T \rightarrow IN$  is the minimal delay function, and  $u : T \rightarrow IN \cup \{\infty\}$  is the maximal delay function such that for any  $t \in T, l(t) \leq u(t)$ . For every state  $\sigma_i \in \Sigma$ , a set of  $t$ -successors  $t(\sigma) \subseteq \Sigma$  is defined. TTSs use to be represented as timed-transition diagrams (TTDs)(see Fig. 4, Fig. 5, Fig. 6). In (Kamišalić et al., 2008) we have used TTDs to represent macro-temporality constraints in CAs.

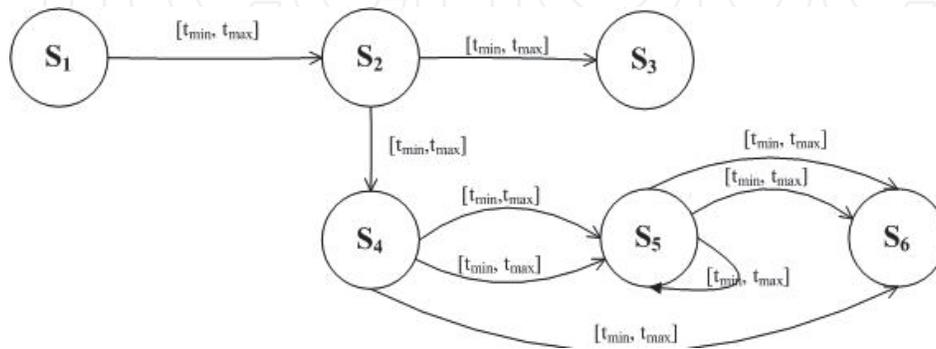


Fig. 4. Clinical algorithm for Hypertension (level 0 data) in TTD representation

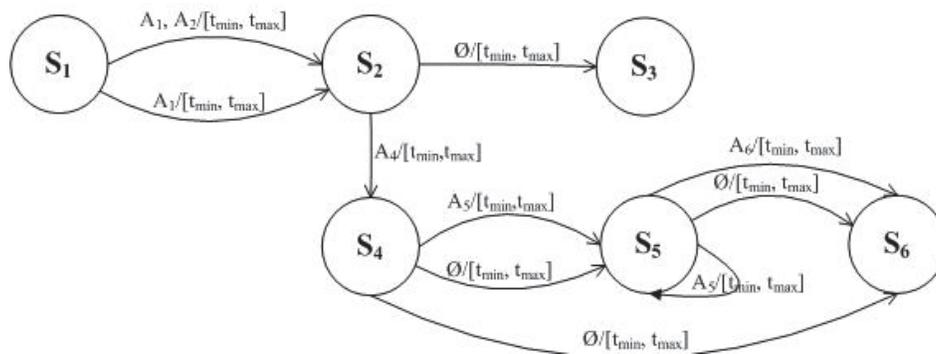


Fig. 5. Clinical algorithm for Hypertension (level 1 data) in TTD representation

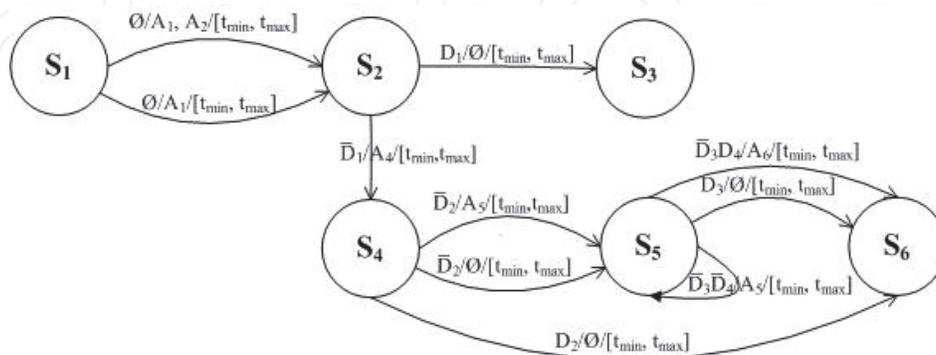


Fig. 6. Clinical algorithm for Hypertension (level 2 data) in TTD representation

### 3. Generation of macro-temporalities

#### 3.1 Data model

Input data for generation of macro-temporalities represent the evolution of patients through a medical treatment as sequences of state transitions. These data are based on the concept of encounter. An encounter is defined as a meeting between a medical professional and a patient in order to assess patient's condition and to determine the best medical course of action (Kamišalić et al., 2007). For each encounter some data about the patient condition (e.g. signs and symptoms) and the actions of the patient treatment (e.g. prescriptions or medical orders) are saved. Patient's evolution is seen as a sequence of state transitions through different consecutive encounters. In the  $i$ -th encounter, a patient  $P_k$  is in state  $S_i^k$  (described by the observed state terms) and, optionally, receives a state-caused treatment  $A_i$  (described by the action terms the physician orders). The time between this encounter and the next one is  $t_{i,i+1}$ . The sequences of state transitions of all the patients affected by and treated of a particular disease define a data model that describes the input data. For a particular disease, if the states in the data model are the same that the states in the CA describing the treatment of that disease, then each state-to-state transition of a sequence represents an instance of the sort of patient evolving between two consecutive states in the CA representation diagram (SDA\* or TTD).

There were identified different data levels in hospital databases. The description of the treatment of a particular patient is defined at level 0 when only the states the patient passes through and the times passing between consecutive states are provided. These structures are called level 0 sequences (Kamišalić et al., 2008)(see representation of level 0 sequences in Fig. 4). Level 1 data describe individual treatments of concrete patients as level 0 sequences of states together with the medical actions performed between each pair of consecutive states, and the time that each change of state takes. These are called level 1 sequences (Kamišalić et al., 2008)(see representation of level 1 sequences in Fig. 5). Level 2 data extends level 1 data with decisions representing the reasons that justify the actions. These are level 2 sequences (Kamišalić et al., 2008)(see representation of level 2 sequences in Fig. 6).

#### 3.2 Macro-temporality generalization

Considering different data levels introduced, there are also defined different procedures for generation of macro-temporalities. In some cases the input data can follow some of the known distributions, such as normal distribution, while in others not. For generation of macro-temporalities in both cases there were defined two different approaches considering distribution of input data. In the case of known distribution there can be used known statistical methods to generate macro-temporality. In this case it is obtained more adjusted macro-temporality interval and there can be reached a higher confidence in the obtained interval. In the case where the distribution function of the data is unknown, for generation of macro-temporality there are used quantiles, which give a first approximation about the interval with less confidence. Both approaches are described in the next subsections. In the case of level 0 data, for each pair of consecutive states  $S_i, S_j$ , it is applied process of macro-temporality constraint generation from all the  $t_{ij}$  times of all the sequences available. Each time a patient takes to evolve directly from  $S_i$  to  $S_j$  is taken in the calculation of macro-temporality interval, whether using one of the known statistical models or using quantiles. In the case of level 1 data, first there is applied process of actions classification, where all the actions of the sort  $A_{ij}$  are used to obtain a group of action classes in which each

action class contains  $A_{ij}$  actions that are mutually similar and dissimilar to the actions in other actions classes (see Fig. 9 and Section 4.2 for more details about similarity criterion) (see an example of 10 evolutions of level 1 sequence in Fig. 7; considering these evolutions there is made a classification of all the evolutions with the same actions, shown in Fig. 8).

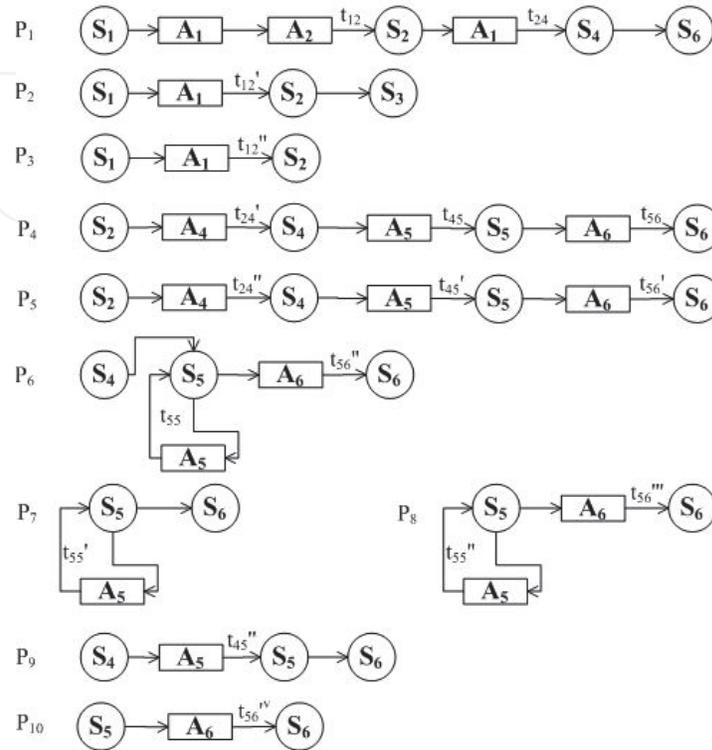


Fig. 7. Example of 10 evolutions of level 1 sequence for Hypertension

The  $t_{ij}$  times related to evolutions in which the actions applied belong to a concrete action class are used to calculate a macro-temporality constraint (see Fig. 8).

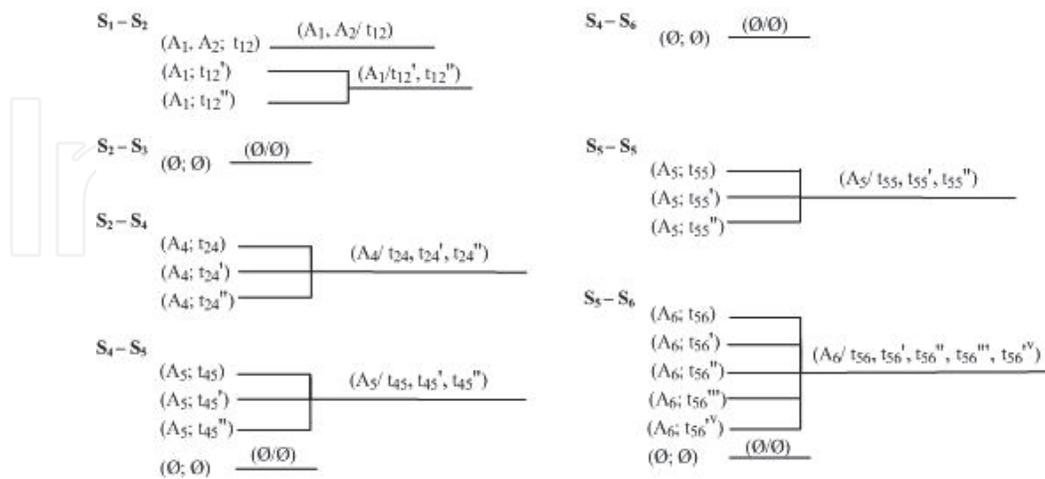


Fig. 8. Actions classification for Fig. 7 example

Finally for level 2 data, there is applied not only the process of actions classification (as in the case of level 1 data) but also the process of decisions classification (the same procedure as in the case of actions classification for level 1 data) (see Fig. 9 and Section 4.2 for more

details about similarity criterion). The next step consists of calculating the time  $t_{ij}$  between states from all of the times  $t_{ij}$  of all the classified transitions from  $S_i$  to  $S_j$  taken from all the sequences available (Kamišalić et al., 2008).

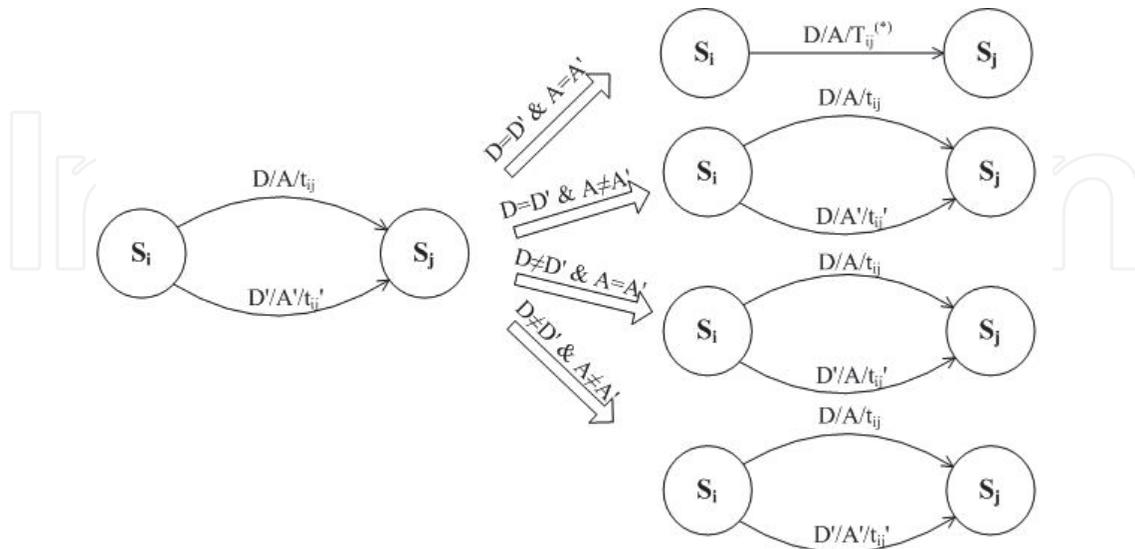


Fig. 9. Macro-temporality generalization: considering decisions and actions similarity

### 3.2.1 The statistical model

If we assume that for each pair of consecutive states  $S_i$ ,  $S_j$  and time between them  $t_{ij}$ , the sample of  $t_{ij}$  times of all the transitions from  $S_i$  to  $S_j$  taken from patient sequences approach one of the known distributions then we can use statistical methods for macro-temporality constraints generation. Here it is presented the statistical model for the sample approaching a  $t$ -student distribution. In this case,  $\bar{t}_{ij}$  represents the mean of all  $t_{ij}$  values in the sample of patients evolving from  $S_i$  to  $S_j$ .  $S_{t_{ij}}$  is the standard deviation of that same sample. Equation 1 is used to calculate the macro-temporality, where  $t_n$  is the  $z$ -value of  $t$ -student distribution (Kamišalić et al., 2007).

$$t = \bar{t}_{ij} \pm t_n \cdot S_{t_{ij}} \quad (1)$$

The calculated macro-temporalities between consecutive states  $S_i$  and  $S_j$  can be assigned to the SDA\* connectors ending in  $S_j$ . The connectors could be of the sort  $C_{SS}$ ,  $C_{AS}$  or  $C_{DS}$ .

### 3.2.2 Quantiles

Quantiles are more useful than other statistical methods if the distribution function of the data we are analysing is unknown. For each pair of consecutive states ( $S_i$ ,  $S_j$ ) there is an assigned time  $t_{ij}$  between them. The  $t_{ij}$  times of all transitions from  $S_i$  to  $S_j$  taken over all the patient sequences available defines a sample that is used to generate a macro-temporality constraint on  $t_{ij}$  times. Quantiles of  $q$ -quantiles are the data values that divide an ordered list of data into  $q$  essentially equal-sized data subsets. Quantiles are data values which mark the boundaries between consecutive subsets. For example, percentiles are 100-quantiles. Using the precentile is a way of providing estimation of proportions of the data that should fall above and below a given value. The 1st percentile cuts off the lowest 1% of data and the 99th percentile cuts off the highest 1% of data. Considering the confidence we aim to reach for the generated

interval we can decide which lowest and highest percentiles we will cut off. For example, if we decide to keep 90% of confidence in generated interval, we would cut off till 5th and from 95th percentile. In this case, macro-temporality interval would consist of  $t_{min}$  as a value of 5th percentile and  $t_{max}$  as a value of 95th percentile (Kamišalić et al., 2008). If the input data follow some of the known distributions, such as normal distribution, then we could use other statistical methods to generate macro-temporality which would give us a more adjusted interval and a higher confidence in the obtained interval.

### 3.3 Incorporation of macro-temporalities in representation diagrams

Transitions (connectors) in temporal CAs can be extended with macro-temporalities with the help of the instances of treatment in the medical databases under the data model described in section 3.1. SDA\* diagrams or TTDs are chosen as representation formalism for temporal CAs (representation of macro-temporality constraints). Each encounter in the database represents the medical actions taken for a patient in a particular state. Identifying the state of the patient  $S_E$  among all the states in the SDA\* is made with the operation of similarity  $S_E \approx S_{SDA}$  defined in subsection 4.2. Identifying the medical action of the encounter  $A_E$  in the SDA is made as well with the operation of similarity  $A_E \approx A$ , where  $A$  is each one of the actions in  $Actions(S_{SDA}, S_E)$ , and  $Actions(S_{SDA}, S_E)$  the set of alternative actions in the SDA that patients in condition  $S_E$  may follow after the SDA state  $S_{SDA}$ .

After this identification of the encounter in the states, decisions and actions of the CA, the macro-temporalities of the encounter are attached to the corresponding connectors between identified states, decisions and actions in the SDA\* or TTD. This process is repeated for all the encounters in the database. At the end, all the macro-temporalities attached to each one of the connectors in the SDA\* or TTD are combined with the procedure of macro-temporality generalization (using statistical model or quantiles), described in subsection 3.2. The final SDA\* or TTD has several macro-temporalities attached to each one of its connectors. If all the micro-temporalities of the terms are close (i.e., temporalities in similar cases are similar) then the connectors in the SDA\* or in TTD are related a single macro-temporality (see example of attached macro-temporalities in Fig. 2 for SDA\* representation or Fig. 5 for TTD representation).

### 3.4 Case study

For the connector  $C_{AS}$  between the action  $A_4$  and the state  $S_4$ , shown in Fig. 2, we consider times  $t_{ij}$  between consecutive encounters in which, during the first encounter the physician ordered the action  $A_4$ , and in the second encounter the patient had evolved to state  $S_4$ . Assume we have different patients who evolved through the same transition with different times (2d, 14h, 4d, 3d, 5d, 20h, 1w, 14h, 1d, 6d, 18h, 2d, 20h, 1w, 4d, 10h, 2d, 4d, 14h, 3d). We assume that the distribution function of the given data is unknown, so we calculate quantiles of the times in the list, to obtain macro-temporality  $[t_{min}, t_{max}]$  for given transition. First step consists of deviding the times in an ordered list. In this example we will use deciles (10-quantiles). It means that the list of times will be divided into 10 equal-sized subsets. Depending on the confidence we aim to reach for generated interval, we decide which lowest and highest deciles we cut of. Say that we want to keep 60% of confidence in the interval, then we cut the first 2 and the last 2 deciles. As a result we obtain the macro-temporality interval  $T_8=[18h, 4d]$ .

## 4. Generation of micro-temporalities

### 4.1 Data model

The data model is based on the concept of encounter. For each encounter all the relevant data about the patient condition and the treatment actions are saved. These data use state, decision and action terms to describe both the condition in which the professional finds the patient at the particular moment of the encounter, and the decision which that professional took for the patient. That is to say, the actions which were proposed in the encounter as part of the treatment.

In the data model, all state, decision and action terms are attached a micro-temporality constraint that is not necessarily complete (as presented in Section 2.1, Fig. 3).

### 4.2 Micro-temporality generalization

In order to be able to generate micro-temporality from data, we have to know how to compare states, decisions and actions. Two states, two decisions or two actions can be considered the same, if they are similar enough to be accepted as the same. That is to say, they are atemporarily the same (represented as  $\approx$ ) if each term of the states, decisions or actions corresponds to some of the terms of the other state, decision or action; or if they differ in some terms that have no major signification (i.e., their effects do not change the essence of a particular state, decision or action). Our method uses the similarity criterion (Kamišalić et al., 2009):

$$S \approx S' \Leftrightarrow \frac{(S \cap S')}{(S \cup S')} > \alpha$$

for some predefined parameter  $\alpha \in [0, 1]$ , where 0 determines total difference (no single term matching in compared states, decisions or actions) and 1 determines total equality (all terms matching) of compared states, decisions or actions.

If we come to the conclusion that several states, several decisions or several actions are atemporarily the same, then we also consider the micro-temporalities of the corresponding terms in the states, decisions or actions as it is depicted in Fig. 10 in order to evaluate if they are also the same at the level of time.

For all the micro-temporalities  $[s_t, e_t, f_t]$  corresponding to the same term, it is necessary to measure the distance between these temporal constraints. The distance is measured for all constraints  $s_t$ ,  $e_t$  and  $f_t$ , separately. This approach is called S-time-distance (Sicherl, 2005). The distance between temporal constraints  $s'_t$  and  $s''_t$  is calculated with Equation 2, between  $e'_t$  and  $e''_t$  with Equation 3, and between  $f'_t$  and  $f''_t$  with Equation 4.

$$d_s(s'_t, s''_t) = \Delta s_t = |s'_t - s''_t| \quad (2)$$

$$d_t(e'_t, e''_t) = \Delta e_t = |e'_t - e''_t| \quad (3)$$

$$d_f(f'_t, f''_t) = \Delta f_t = |f'_t - f''_t| \quad (4)$$

The result of calculating the distances of temporal constraints are three values which have to be compared to some predefined parameter  $\beta$ . When all three distances are lower than or equal to a predefined parameter  $\beta$ , we say the terms with such micro-temporalities are the

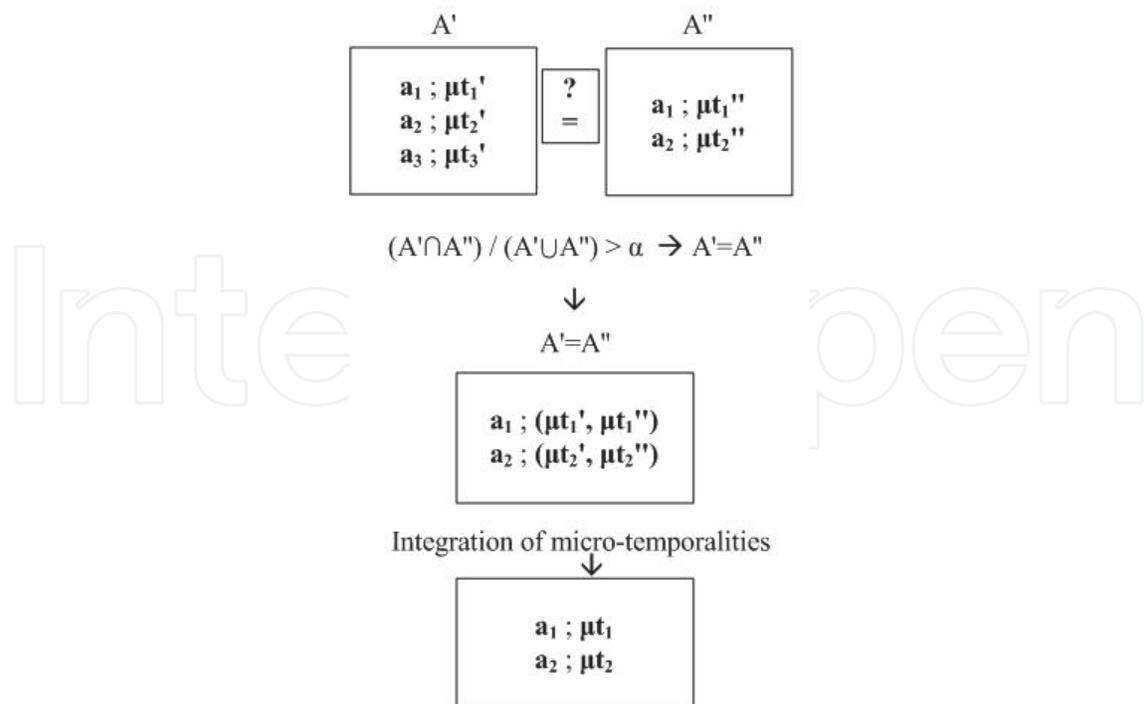


Fig. 10. Comparing action terms and providing integration of micro-temporalities

same as far as time is concerned. Formally speaking, we define two micro-temporalities  $\mu t'$  and  $\mu t''$  are the same (represented as  $\mu t' \approx \mu t''$ ) if:

$$\mu t' \approx \mu t'' \Leftrightarrow d_s(s'_t, s''_t) \leq \beta, \text{ and} \\ d_e(e'_t, e''_t) \leq \beta, \text{ and} \\ d_f(f'_t, f''_t) \leq \beta.$$

When there are several micro-temporalities related to a same term, these micro-temporalities may be combined with a clustering algorithm that employs the distance functions  $d_s$ ,  $d_e$ , and  $d_f$  defined above. Hierarchical taxonomy clustering algorithms (Sneath & Sokal, 1973) are numerical approaches to the construction of dendrograms on a set of numerical data. Given a set of micro-temporalities  $[s_i, e_i, f_i], i = 1, \dots, n$  that we want to combine, our method constructs a dendrogram for each one of the square matrices  $\{d_s(s_i, s_j)\}_{i,j=1,\dots,n}$ ,  $\{d_e(e_i, e_j)\}_{i,j=1,\dots,n}$ , and  $\{d_f(f_i, f_j)\}_{i,j=1,\dots,n}$ . Once the heights of these dendrograms are scaled to the range  $[0,1]$ , a predefined parameter  $\beta$  is used to horizontally cut the dendrograms. Parameter  $\beta$  is chosen considering whether we can accept the increasing of error and therefore decreasing the number of micro-temporalities, or we prefer getting more adjusted temporal constraints representatives (increasing the number of micro-temporalities) and therefore decreasing the number of micro-temporalities. These  $\beta$ -cuts provide three sets of clusters. The median of the values in each cluster defines the centroid of this cluster. Constraints which are the closest to centroids, are chosen as cluster representatives. These cluster representatives of different dendrograms are combined in the final micro-temporalities.

#### 4.3 Incorporation of micro-temporalities in representation diagrams

Terms in temporal CAs can be extended with micro-temporalities with the help of the instances of treatment in the medical databases under the data model described in section 4.1.

As representation formalism for temporal CAs (representation of micro-temporality constraints) are chosen SDA\* diagrams. Each encounter in the database represents the medical actions taken for a patient in a particular state. The terms in the encounter come related to micro-temporalities. Identifying these terms in the atemporal SDA is the first step of the process. Identifying the state of the patient  $S_E$  among all the states in the SDA is made with the operation of similarity  $S_E \approx S_{SDA}$  defined in subsection 4.2. Identifying the medical action of the encounter  $A_E$  in the SDA is made with the operation  $A_E \approx A$ , where  $A$  is each one of the actions in  $Actions(S_{SDA}, S_E)$ , and  $Actions(S_{SDA}, S_E)$  the set of alternative actions in the SDA that patients in condition  $S_E$  may follow after the SDA state  $S_{SDA}$ . For example in Fig. 2,  $Actions(S_5, \{noD_3\}) = \{A_5, A_6\}$  because a patient in state  $S_5$ , which is not  $D_3$ , may be derived to action  $A_5$  if he is not  $D_4$  or to  $A_6$  if he is  $D_4$ .

After this identification of the encounter in the states, decisions and actions of the SDA, the micro-temporality of the terms of the encounter are attached to the corresponding terms in the SDA states, decisions, and actions. This process is repeated for all the encounters in the database. At the end, all the micro-temporalities attached to each one of the terms in the SDA are combined with the procedure of clustering algorithm, described in subsection 4.2. The final SDA has several micro-temporalities attached to each one of its terms. If all the micro-temporalities of the terms are close (i.e., temporalities in similar cases are similar) then the terms in the SDA are related a single micro-temporality.

#### 4.4 Case study

For two action terms (take thiazide-type diuretic and take beta-blocker) of the CA shown in Fig. 2 as a part of the action A4, we consider three sorts of constraints,  $s_t$ ,  $e_t$  and  $f_t$ . We have to find clusters (and centroids) for each one of them. Micro-temporality  $[s_t, e_t, f_t]$  of the applied action term is given for each patient. Assume we have different patients with the same action term but different micro-temporalities ( $[-, 2w, 6h]$ ,  $[-, 3w, 4h]$ ,  $[-, 2w, 6h]$ ,  $[-, 1w, 8h]$ ,  $[-, 6d, 6h]$ ,  $[-, 10d, 8h]$ ,  $[-, 4w, 6h]$ ,  $[-, 15d, 6h]$ ,  $[-, 4d, 8h]$ ,  $[-, 20d, 4h]$ ). We apply the clustering algorithm separately for all  $s_t$ ,  $e_t$  and  $f_t$ . Finding the clusters for groups of  $s_t$ ,  $e_t$  and  $f_t$  gives us the centroids of those clusters. The closest constraint to each centroid constructs the micro-temporality of a term. For example, if all  $e_t$  are divided into more than one cluster, the result is more than one centroid. The closest constraint to the centroid is taken as a time constraint on the end time in different micro-temporalities of the same action term (for above introduced example of patients' micro-temporalities, the result of the clustering are four clusters and therefore four constraints closest to the centroids for  $e_t$  ( $6d, 2w, 3w, 4w$ ) and two clusters and therefore two constraints closest to the centroids for  $f_t$  ( $4h, 6h$ ); which give as the result four micro-temporalities  $[-, 6d, 6h]$ ,  $[-, 2w, 6h]$ ,  $[-, 3w, 4h]$  and  $[-, 4w, 6h]$  for the action term 'take thiazide-type diuretic'). In this particular case, we would choose  $[-, 2w, 6h]$  micro-temporality from the four possible, because this one is the representative of most micro-temporalities. However, choosing one representative is resulting in increasing the error and decreasing the confidence.

## 5. Conclusion

CAs derived from CPGs make explicit knowledge necessary to assist physicians in order to make appropriate decisions. They should give procedural and temporal indications. The first one are indications on what to do, and the second one are indications on what are the time restrictions. Still the most CAs are atemporal, offering indications on what to

do but not on time restrictions. To deal with this problem, we have defined micro- and macro-temporality constraints and offered the methodologies for generalization of micro- and macro-temporalities. We are introducing temporal constraints from existing patients' temporal data. We have presented two possible approaches to generation of macro-temporality constraints, using statistical methods or quantiles. Quantiles are less susceptible to long tailed distributions and outliers. We came to the conclusion that if the data we are analyzing are not distributed according to some assumed distribution or if we have other potential sources for outliers that are far removed from the mean, then quantiles may be more useful than other statistical models. Obtained macro-temporalities are attached to connectors of the CAs. We have also offered the methodology for micro-temporality constraints generation using a clustering algorithm. The obtained micro-temporalities are attached to a state, decision or action term of the CAs. We have presented different representation formalisms which can be used for representation of CAs, and have explained different situations in which STDs, TTDs and/or SDA\*s are used. As CAs now include temporal constraints (time dimension), physicians are able to provide also indications on time restrictions considering medical procedures.

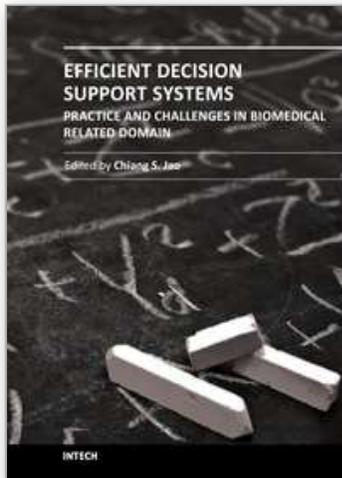
## 6. References

- Combi, C.; Keravnou-Papailiou, E. & Shahar, Y. (2010). *Temporal Information Systems in Medicine*. Springer, 978-1-4419-6543-1, New York, 2010
- Fox, J.; Johns, N. & Rahmzadeh, A. (1998). Disseminating Medical Knowledge-The PROforma Approach. *Artificial Intelligence in Medicine*, 14(1-2): 157-182, 1998
- Grabmeier, J. & Rudolph, A. (2002). Techniques of Cluster Algorithms in Data Mining. *Data Mining and Knowledge Discovery*, 6(4): 303-360, October 2002
- Henzinger, T. A.; Manna, Z. & Pnueli A. (1992). Timed Transition Systems. *Real Time: Theory in Practice*, LNCS 600: 226-251, 3-540-55564-1, Springer Verlag 1992
- Institute for Clinical Systems Improvement(2006). Health Care Guideline: Hypertension Diagnosis and Treatment, October 2006, [http://www.icsi.org/hypertension\\_4/hypertension\\_diagnosis\\_and\\_treatment\\_4.html](http://www.icsi.org/hypertension_4/hypertension_diagnosis_and_treatment_4.html)
- Kamišalić, A.; Riaño D.; Real F. & Welzer T. (2007). Temporal Constraints Approximation from Data about Medical Procedures. *Machine Learning and Management of Health-Care Procedural Knowledge, 20th IEEE Int. Workshop on CBMS2007*, Maribor, Slovenia
- Kamišalić, A.; Riaño D. & Welzer T. (2008). Generating Macro-Temporality in Timed Transition Diagrams. 'From Knowledge to Global Care' AIME2007 workshop K4CARE, LNCS 4924: 62-74, 978-3-540-78623-8, Springer Verlag 2008
- Kamišalić, A.; Riaño D. & Welzer T. (2009). First Approach to Micro-temporality Generation for Clinical Algorithms. *European-Japanese Conference EJC2009*, Maribor, Slovenia
- Musen, M. A.; Tu, S.W.; Das, A.K. & Shahar, Y. (1996). EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Information Association*, 3(6): 367-388, 1996
- Riaño, D. (2007). The SDA Model: A Set Theory Approach. *Machine Learning and Management of Health-Care Procedural Knowledge, 20th IEEE Int. Workshop on CBMS 2007*, Maribor, Slovenia
- Riaño, D.; Lopez-Vallverdu, J. A. & Tu, S. (2007). Mining Hospital Data to Learn SDA\* Clinical Algorithms. 'From Knowledge to Global Care' AIME2007 workshop, LNCS 4924: 46-61, 978-3-540-78623-8, Springer Verlag 2008

- Seyfang, A.; Miksch, S. & Marcos M. (2002). Combining Diagnosis and Treatment using Asbru. *International Journal of Medical Informatics*, 68 (1-3): 49-57, December 2002
- Sicherl, P. (2005). Time distance - a new generic approach for analysis and visualization of time related data, *Complex objects visualization - COV2005*, Ljubljana, Slovenia
- Sneath, P. H. A. & Sokal, R. R. (1973). Numerical Taxonomy, *Freeman*, San Francisco, 1973

IntechOpen

IntechOpen



## **Efficient Decision Support Systems - Practice and Challenges in Biomedical Related Domain**

Edited by Prof. Chiang Jao

ISBN 978-953-307-258-6

Hard cover, 328 pages

**Publisher** InTech

**Published online** 06, September, 2011

**Published in print edition** September, 2011

This series is directed to diverse managerial professionals who are leading the transformation of individual domains by using expert information and domain knowledge to drive decision support systems (DSSs). The series offers a broad range of subjects addressed in specific areas such as health care, business management, banking, agriculture, environmental improvement, natural resource and spatial management, aviation administration, and hybrid applications of information technology aimed to interdisciplinary issues. This book series is composed of three volumes: Volume 1 consists of general concepts and methodology of DSSs; Volume 2 consists of applications of DSSs in the biomedical domain; Volume 3 consists of hybrid applications of DSSs in multidisciplinary domains. The book is shaped decision support strategies in the new infrastructure that assists the readers in full use of the creative technology to manipulate input data and to transform information into useful decisions for decision makers.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Aida Kamišalić, David Riaño and Tatjana Welzer (2011). Temporal Knowledge Generation for Medical Procedures, Efficient Decision Support Systems - Practice and Challenges in Biomedical Related Domain, Prof. Chiang Jao (Ed.), ISBN: 978-953-307-258-6, InTech, Available from:  
<http://www.intechopen.com/books/efficient-decision-support-systems-practice-and-challenges-in-biomedical-related-domain/temporal-knowledge-generation-for-medical-procedures>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen