

Engineering Time in Medical Knowledge-Based Systems through Time-Axes and Time-Objects

E.T. Keravnou

Department of Computer Science, University of Cyprus
Kallipoleos 75, P.O.Box 537, CY-1678 Nicosia, Cyprus
email: elpida@turing.cs.ucy.ac.cy

Abstract

Starting from the premise that time representation and temporal reasoning must constitute integral aspects of a competent, knowledge-based, medical system, the paper presents the relevant requirements and discusses their realisation in terms of a generic temporal kernel to be embedded in such a system. The kernel has a layered architecture where the bottom layer gives the ontological primitives and their associated axioms, and the higher layers implement the required temporal reasoning. The principal primitives of the ontology are the time-axis and the time-object.

1 Time representation requirements for medical problem solving

Time is intrinsically relevant to medical problem solving. Time representation requirements for medical tasks such as diagnosis, prognosis, monitoring, therapy planning, etc. demand more variety in expression and higher levels of abstraction than appears to be supported by well known, general theories of time proposed in the AI literature [1],[3],[12]. There are two basic issues here: how to model time per se and how to model time-varying situations or occurrences.

1.1 Model of time

Real time is infinite and dense. An abstraction of reality which models time as a single time-line (either in dense or discrete terms), a model often adopted in temporal databases and other applications, does not provide the appropriate abstraction for medical applications where a richer model providing a multidimensional structure to time, through a number of interrelated, *conceptual temporal contexts*, and *multiple granularities*, is often required. Examples of conceptual temporal contexts are the various developmental periods, eg fetal-period, infancy, early-childhood, etc whose

'natural' granularities are lunar months, calendar months, years, etc, respectively.

1.2 Model of occurrences

An appropriate model for occurrences must support the following:

- *Absolute and relative occurrences.* "No ossification of knee epiphyses at birth" expresses an occurrence in absolute terms, ie with respect to some fixed point, while "narrow thorax until kyphoscoliosis" expresses an occurrence, "narrow thorax", relative to another occurrence, "kyphoscoliosis".
- *Absolute and relative vagueness.* "Wide triradiate cartilage *upto about* the age of 11 years" expresses absolute vagueness while "nausea precedes or coincides with the headache" expresses relative vagueness.
- *Absolute and relative duration.* "Two days of headache" expresses the duration of headache in absolute terms while "headache during the nausea" does so in a relative way.
- *Incompleteness.* Patient information is often expressed in a temporally discrete and thus temporally incomplete fashion, eg the record of some patient could include statements like "mild scoliosis at the age of 2 years" and "severe scoliosis at the age of 7 years" without any mention of the status of scoliosis (absent or present and of what severity) at any other points in time.
- *Point and interval occurrences.* The same occurrence can be expressed with respect to different temporal granularities, thus giving it point or interval status (in a conceptual rather than real-life sense) depending on the temporal context of reference. For example "I had flu for most of January 1995" expresses an interval occurrence at the granularity of days but a point occurrence at the granularity of months. With respect to interval occurrences the issue of convexity, or nonconvexity [13], arises (convexity implies that the

unfolding of the occurrence entails some sort of activity throughout the particular period of time that defines its lifetime).

- *Compound occurrences.* Disease processes and therapeutic interventions define compound occurrences. Compound occurrences are categorised into: *periodic occurrences* (eg "headache every morning for about two hours over a period of one week which worsens each day", "administration of drug x every four hours until the pain stops but not for more than two days", etc); *temporal trends* that describe changes and their direction and rate of change, eg "low pressure increasing slowly", which in turn indicate whether a situation is normal or whether it is converging towards, or diverging away from, normality and at what rate; and general *temporal patterns*, eg a sequence of meeting trends, a set of relative occurrences, a set of causally related occurrences, etc. The modelling of compound occurrences requires mechanisms for abstraction and refinement.
- *Causality.* Changes are explained through causal relations and hence time is directly relevant to causality. The temporal principle underlying causality is that an effect cannot precede its cause.

2 Ontological primitives

The relative simplicity and crispness of the primitives and reasoning of most widely adopted general theories of time which in fact justify the wide interest in them, unfortunately do not render their expressive power sufficient enough for a number of real-life medical problems where happenings are not as simple and as orderly as someone walking on a street or getting a promotion. Such theories do not provide an adequate level of abstraction for knowledge engineering purposes in that their primitives, often the time-point, the time-interval, or the event, are too 'primitive' and not at a knowledge level from the perspective of modelling complicated, dynamic, domain concepts such as disease processes or therapeutic interventions. Another indirect limitation is based on the fact that still in many knowledge-based problem solvers temporal reasoning is seen as an accessory rather than as an integral aspect of the problem solver's reasoning. For example the temporal reasoner could be a background process demoted to the role of preprocessing data (from some temporal database) in order to select or derive the information to be fed to the problem solver; the problem solver, therefore, is not reasoning with time or change in any conscious way. An ontology of time that in a sense 'forces' time to be treated as an integral aspect of the particular problem solving and hence for temporal reasoning to constitute a very conscious activity on the part of the problem solver has significant advantages over an ontology that does not and which may even

unintentionally cultivate the treatment of time as an accessory. To achieve the above, an ontology must force time to be an integral aspect of the domain entities that constitute the processing elements of the problem solver. This the proposed ontology aims to achieve through its time-object primitive, which in addition results in the simple integration and uniform representation of structural, causal, and temporal knowledge. Furthermore, the ontology aims to provide the necessary temporal abstraction through multiple temporal granularities.

The principal ontological classes are the *time-axis* and the *time-object* that respectively provide a model of time and a model of occurrences. Time-axes provide conceptual contexts for the definition of time-objects and hence they result in organising time-objects in meaningful (temporal) clusters. A time-axis is expressed at the granularity relevant to its semantics and multiple granularities are therefore appropriately modelled by confining their usage to relevant, conceptual, temporal contexts. The notion of an (anonymous) time-axis is not original per se. However, the notion of *multiple*, conceptual (ie named), interrelated, time-axes is. Multiplicity of time-axes is necessary for the natural representation of some problem. For example the ossification process for the first cervical vertebra of the spine begins at the second lunar month and terminates at the 25th year of age. The initiation of this process is expressed at the granularity of lunar months and its termination at the granularity of years. The time-values concerned refer to conceptual periods of time, namely fetal-period and maturity respectively. The lifetime of the particular ossification process extends over fetal-period, infancy, childhood, puberty and part of maturity. This process is temporally decomposed into a number of subprocesses. A process and its subprocesses are modelled as time-objects.

Time-objects are dynamic entities. The existence of some time-object could be expressed with respect to multiple temporal contexts (ie time-axes) and hence granularities, either in absolute or relative terms. In addition, the existence of a time-object, in some context, can be expressed with a degree of vagueness. Three important types of relations are defined for time-objects: structural, causal, and temporal relations. The ontological classes are:

- *Tunits*, the set of *time-units*.
- *Axes*, the set of discrete *time-axes*.
- *Times*(α), the sequence of literal *time-values* on time-axis α .
- *Tobjects*, the set of *time-objects*.
- *Tobjects*(α) (\subseteq Tobjects), the set of time-objects that have a valid existence on time-axis α .
- *Pobjects*(α) (\subseteq Tobjects(α)), the set of time-objects that have a point existence on time-axis α .
- *Iobjects*(α) (\subseteq Tobjects(α)), the set of time-objects that have an interval existence on time-axis α .

- *Props*, the set of *properties*.
- *Props*(μ) (\subseteq Props), the set of properties of relevance to time-unit μ .

2.1 Time-units and time-axes

Time-units define the possible granularities; *scale*(μ, μ', s) gives the scale relation between granularities μ and μ' . A time-axis, α , is expressed discretely at a specific granularity, μ , in terms of a sequence of time-values, *Times*(α) = $\{t_1, t_2, \dots, t_i, \dots, t_n\}$, given with respect to the *origin* of the time-axis. The origin of an abstract time-axis denotes a generic time-point; when an abstract time-axis is instantiated its origin gets bound to a real time-point. Hence a concrete time-axis is 'linked' to the universal, real-time, axis. The basic relation between time-axes is *t-link*(α, t, α', t') that links a time-value on one time-axis with a time-value on another time-axis. Other relations, eg concurrent, includes, intersects-with, etc can be expressed through t-link. Some of the time-axes are defined as *spanning* a chaining sequence of other time-axes. A spanning time-axis has a hybrid granularity inherited from its components. For example lifetime could be a spanning time-axis encompassing fetal-period, infancy, etc.

2.2 Time-objects

A time-object, τ , is a dynamic entity for which time constitutes an integral aspect. It is viewed as a tight coupling between a *property* and an *existence*, where its existence can be expressed with respect to different time-axes. Its existence with respect to the most appropriate time-axis for it, is called the time-object's *main existence*. Thus $\tau \equiv \langle \pi_\tau, \varepsilon_\tau \rangle$ where $\pi_\tau \in$ Props is the property of τ (π is a selector function, and the notation π_τ stands for function π applied to argument τ), and function $\varepsilon_\tau : \text{Axes} \rightarrow \text{Eexprs}$ is the existence function of τ . The domain of ε_τ is the set of time-axes and its range is the set of absolute existence expressions. An absolute existence expression gives the (earliest) initiation and (latest) termination of some existence. If time-object, τ , has a valid existence on some time-axis, α , then $\varepsilon_\tau(\alpha) = \langle t_s, t_f \rangle$; $t_s, t_f \in \text{Times}(\alpha)$; $t_s \leq t_f$ and the status $\varsigma \in \{\text{closed, open, open-from-left, open-from-right, moving}\}$. If there is openness in some valid existence of a time-object then its actual initiation and/or termination is not known but approximated through an admissibility margin. The initiation and termination admissibility margins are given by $\langle t_s, t_s \rangle$ and $\langle t_f, t_f \rangle$ respectively where $t_s = \text{le-fr}_\tau(\alpha)$ and $t_f = \text{ri-fr}_\tau(\alpha)$ (functions le-fr and ri-fr respectively stand for left-freedom and right-freedom). The existence of an open time-object, on a given time-axis, is therefore defined through an initial period of uncertainty, an in between period of certainty and a final period of uncertainty. If the earliest termination of an open time-object precedes or coincides with its latest initiation then there is no period of

certainty. Thus the duration of a non-closed existence of some time-object can only be shortened. If τ does not have a valid existence in the context of time-axis α , then $\varepsilon_\tau(\alpha) = \perp$.

Hence a time-object can exist as a *point-object* on some time-axis but as an *interval-object* on another time-axis. In the former case the extent of the time-object is less than the time-unit of the particular time-axis. A special moving time-object is *now* which exists as a point-object on any relevant time-axis and functions to partition time-objects into past, future, or ongoing. A time-object whose existence is expressed with respect to an abstract/concrete time-axis is an abstract (generic)/concrete time-object.

Regarding temporal relations between time-objects, Allen's set [1] has been adapted and extended to fit the discrete, multidimensional, and multigranular, model of time. More specifically a time-axis constitutes an argument of these relations, instances of which may be derived from absolute existence expressions; furthermore new relations based on temporal distance are added and useful disjunctions are directly defined. The structural relations between time-objects are *isa-component-of*, and its inverse *contains*, and *variant-component*, and its inverse *variant-contains*; the latter two express conditional containment.

contains(τ_i, τ_j)

$\Leftarrow \text{variant-contains}(\tau_i, \tau_j, \text{cs}) \wedge \text{conds-hold}(\text{cs})$

A variant component can only be assumed, in some situation (eg a specific patient), if the specified conditions are satisfied. A compound time-object has a valid existence under any context in which at least one of its components has a valid existence, and a time-object exists within the one that contains it. Temporal views of a compound time-object, from the perspective of specific temporal contexts, can thus be defined. Finally, relation *causes* between a pair of (abstract) time-objects τ_i and τ_j specifies various constraints (temporal and other) that need to be satisfied in order for a *causality-link* to be established to hold between a pair of concrete instances of τ_i and τ_j . A generic constraint, that always needs to be satisfied, is that a potential effect cannot precede its potential cause.

causality-link($\tau_i, \tau_j, \text{cf}$)

$\Leftarrow \text{causes}(\tau_i, \tau_j, \text{cs}, \text{cf}) \wedge \text{conds-hold}(\text{cs}) \wedge$

$\neg \text{starts-before}(\tau_j, \tau_i)$

The fourth argument of relation *causes* denotes a certainty factor, ie even if all specified conditions are satisfied, in some situation, still it may not be definite that the particular *causality-link* actually exists in that situation. The uncertainty is due to knowledge incompleteness.

Trends and periodic occurrences are modelled as compound time-objects, subsuming a number of other time-objects at a lower data and/or temporal level. A generic periodic time-object is defined through a

repetition element, a repetition pattern, and a progression pattern.

2.3 Properties

Properties (ontology class Props), that constitute the other half of time-objects, are atomic or compound (negations, disjunctions, or conjunctions), passive or active, and some are time-invariant. Examples of properties are "sex male", "sore throat", "severe coughing", etc. A property is associated with relevant granularities, eg "headache present" is associated with hours and days, but not months or years. This way the time-axes meaningful to a property or the subset of properties that can be instantiated in the context of some time-axis, α , AxisP(α), can be defined.

A property, ρ , either has an infinite persistence, $inper(\rho)$, or a finite persistence, $finper(\rho)$. In the latter case the following are additionally specified: whether the property can *reoccur* (multiple instantiations of the given property in the same context are possible); maximum and minimum durations (*max-dur*, *min-dur*) under any of the relevant granularities, where the default is any duration; and a default admissibility margin for the initiation of any instantiation of the property, under a specific relevant temporal context (*earliest-init*, *latest-init*), which if not specified is assumed to be the entire extent of the particular time-axis. In addition, the proposed ontology adopts the semantic attributes of properties specified by Shoham, eg downward hereditary, upward hereditary, solid, gestalt, etc [19].

Causality, with explicit temporal constraints, is specified at the level of properties as well, through relation *cause-spec* which is a 6-place relation where the first two arguments are properties, the third is a granularity, the fourth and fifth are sets of relative and absolute temporal constraints respectively, and the last argument is a certainty factor. This relation also enables the derivation of the existence of a *causality-link* between a pair of (concrete) time-objects through the following axiom.

$$\begin{aligned} & \text{causality-link}(\tau_i, \tau_j, cf) \\ & \Leftarrow \text{cause-spec}(\rho_i, \rho_j, \mu, TRel, C\text{ss}, cf) \wedge \\ & \quad \pi(\tau_i) = \rho_i \wedge \pi(\tau_j) = \rho_j \wedge \\ & \quad \text{satisfied}(\tau_i, \tau_j, \mu, TRel) \wedge \text{satisfied}(\tau_i, \tau_j, \mu, C\text{ss}) \wedge \\ & \quad \neg \text{starts-before}(\tau_j, \tau_i) \end{aligned}$$

Other property relations include exclusion, necessitation, etc.

In summary the ontological primitives provide the necessary conceptual abstraction for the adequate modelling of medical concepts (disorders, patient data, therapeutic-actions) [6]. Medical knowledge-bases and patient databases are viewed as collections of abstract and concrete time-objects respectively (and their relevant time-axes) and hence time becomes an integral aspect of the particular applications. All the ontological classes and their associated axioms are discussed in detail in [7].

Some of the axioms are treated as deductive rules (eg the causality axioms) and others as integrity constraints (eg the containment axioms), a distinction adopted from [15]. A detailed comparison between the proposed ontology and three widely adopted time ontologies, namely Allen's interval-based temporal logic [1], Kowalski and Sergot's event calculus [12], and Dean and McDermott's time token maps [3] is given in [7].

3 Medical problem solver with integral temporal reasoning

The ontological primitives and their associated axioms form the ground layer of the embedded temporal kernel. This is the layer that provides the interface with the (medical) system's knowledge-base and (patient) database. The other layers of the temporal kernel provide various temporal reasoning functionalities.

3.1 Temporal reasoning requirements

The required temporal reasoning functionalities are classified, in ascending order of level, into mapping and clipping, derivation, and consistency and querying, functionalities. These are listed below:

- Mapping and clipping:
 - *Determining bounds for absolute existences of occurrences*, ie determining admissibility margins for initiations and terminations of occurrences.
 - *Mapping occurrences across temporal contexts*.
 - *Detecting direct conflicts and clipping persistences*.
- Derivation: There are two types of derivations; deriving new occurrences, or deriving new relations between occurrences.
 - *Deriving new occurrences through merging* (see section 3.3).
 - *Deriving new occurrences through decomposition*. The potential components of compound occurrences are inferred.
 - *Deriving new occurrences through temporal data abstraction*:
 - a. *Persistence derivation* (see section 3.3).
 - b. *Deriving temporal trends*.
 - c. *Deriving periodic occurrences*.
 - d. *Deriving potential clusters of related occurrences*. This is the opposite of the decomposition derivation.
 - *Deriving causal antecedents/consequents of some occurrence*.
 - *Deriving a relation between a pair of occurrences*. The types of relations are temporal, structural, and causality links.
- Consistency and querying:
 - *Establishing the overall consistency of some world of occurrences*.

- *Querying occurrences.* A (hypothesised) occurrence, of any degree of complexity is queried against some world.
- *Deriving the state of some world at a particular time.* In an explorative mode, the problem solver may need to be informed about what is believed to be true, at some specific time, in some specific context. The query may be expressed relative to another point in time which defaults to now, eg at time point t , what was/is/will be believed to be true during some specified period p ?

Below we give algorithms for some of the mapping and clipping, and derivation functionalities.

3.2 Mapping and clipping functionalities

First we discuss the mapping of time-objects across time-axes. In order to map a time-object, τ , from time-axis, α , to time-axis, α' , the following conditions must be satisfied:

- The property of τ , π_τ , can be instantiated in the context of time-axis α' , ie $\pi_\tau \in \text{AxisP}(\alpha')$.
- The scale relation between μ and μ' , the respective time-units for α and α' , is known.
- A linkage between the two time-axes is specified. This means that an instance of relation $t\text{-link}$ involving the relevant pair of time-axes is designated as defining the (basic) linkage between the two.

If any of these conditions is not satisfied then $\varepsilon_\tau(\alpha') = \perp$. A basic function of the mapping operation, *map-val*, is to map a time-value, $t \in \text{Times}(\alpha)$, from α to α' [7]. If a particular time-value cannot be mapped then *map-val* returns \perp . The mapping operation essentially involves mapping the base (t_b) and limit (t_l) time-values of τ onto α' . Further, if the existence of τ on α entails uncertainty, then time-values $le\text{-fr}_\tau(\alpha)$ and $ri\text{-fr}_\tau(\alpha)$, if applicable, are also mapped. Overall, there are four cases of mapping: mapping an interval-object to a coarser/finer granularity time-axis; and mapping a point-object to a coarser/finer granularity time-axis. Information is lost when the existence of a time-object is mapped onto a coarser granularity time-axis. It is therefore quite possible that $\text{map}(\text{map}(\varepsilon_\tau(\alpha), \alpha'), \alpha) \neq \varepsilon_\tau(\alpha)$. Since every time-object has a main existence, to avoid the above problem, where possible, mappings should be from the main existence.

Next we discuss direct conflict detection and clipping of persistence. Two time-objects are implicated in a *direct* conflict, if their existences are not disjoint and their properties are mutually exclusive; they are implicated in an *indirect* conflict if there is a direct conflict between one of these and a derivative of the other time-object. A conflict is denoted by a triple (τ, τ', α) that gives the pair of time-objects and the time-axis in whose context the conflict is established. Once a conflict is established, the next step is to try and resolve it.

Algorithm for detecting direct conflicts: The algorithm accepts a list of time-object, time-axis pairs, (τ, α) , where

α refers to the main existence of τ , and returns a list of direct conflicts.

```

detect-conflicts obsaxs
  let conflicts be initialised to the empty list
  for each  $(\tau, \alpha) \in \text{obsaxs}$  do
    for each  $(\tau', \alpha')$  that follows  $(\tau, \alpha)$  in obsaxs do
      if  $\pi_\tau \text{ excludes } \pi_{\tau'}$ 
        then let  $\alpha''$  be the finest granularity
           time-axis between  $\alpha$  and  $\alpha'$ 
        if  $\varepsilon_\tau(\alpha'') \neq \perp \wedge \varepsilon_{\tau'}(\alpha'') \neq \perp$ 
          then if  $\neg \text{disjoint}(\alpha'', \tau, \tau')$ 
            then enter  $(\tau, \tau', \alpha'')$  into conflicts
        else let  $\alpha''$  be the finest granularity
           time-axis amongst the remaining
           relevant time-axes on which both
            $\tau$  and  $\tau'$  have a valid existence
        if there is such a time-axis and
           $\neg \text{disjoint}(\alpha'', \tau, \tau')$ 
        then enter  $(\tau, \tau', \alpha'')$  into conflicts
    end for
  end for
  return conflicts
end detect-conflicts

```

Algorithm for resolving conflicts by clipping persistences:

The algorithm uses auxiliary functions *apparent-overlap*, *terminating-relation*, and *clip-persistence*. Predicate function *apparent-overlap* accepts a conflict and returns true if the relevant overlap could be apparent. It is based on the heuristic that two time-objects which are disjoint under some granularity may appear as coinciding (eg if both of them are mapped onto the same point existence) or chaining, under a coarser granularity.

Function *terminating-relation* accepts a conflict, the relevant domain of time-objects, and an operation mode, *reactive* (meaning that (therapeutic) actions are only instigated in order to combat some established abnormality, and hence such actions could not have been started prior to what they aimed to terminate), or *proactive* (where therapeutic actions could also be instigated in order to prevent an anticipated, future, undesired happening; such actions would normally commence prior to what they aim to prevent). The function returns a triple of values, where either all three values are equal to nil or they denote three time-objects (τ_i, τ_j, τ') , meaning that τ' causes τ_j in order to terminate τ_i (time-object τ' therefore denotes a terminating action for τ_i); time-objects τ_i and τ_j are the two specified in the parameter of the function, but not necessarily in the given order.

```

terminating-relation  $(\tau, \tau', \alpha)$  tobs mode
  if the mode is proactive or
    the initiations of  $\tau$  and  $\tau'$  on  $\alpha$  coincide
  then the following reasoning is repeated for
    either combination of the two time-objects,
    ie  $(\tau_i, \tau_j)$  is set both to  $(\tau, \tau')$  and to  $(\tau', \tau)$ 
  elseif  $\tau$  begins before  $\tau'$  on  $\alpha$ 
    then  $\tau_i$  is set to  $\tau$  and  $\tau_j$  is set to  $\tau'$ 

```

else τ_i is set to τ' and τ_j is set to τ
 end if
 let $\text{relact} \subseteq \text{tobs}$ such that the property for each of
 the selected time-objects denotes a potential
 terminating action for the property of τ_i
 repeat for each $\tau^* \in \text{relact}$
 if a causality-link can be established from τ^* to τ_i
 and τ^* does not begin before τ_i on α
 (if the reasoning is done in a reactive mode),
 then τ^* is the relevant action
 until a relevant action is found
 if a relevant action, τ^* , is found
 then return (τ_i, τ_j, τ^*) else return $(\text{nil}, \text{nil}, \text{nil})$
 end terminating-relation

Finally function *clip-persistence* accepts a pair of
 time-objects (τ, τ') and modifies the main existence of τ to
 the minimum necessary so that this is disjoint from the
 existence of τ' . If in order to do so the whole existence of
 τ is revoked or its duration drops below a specified
 minimum for the relevant granularity then the function
 returns false else it returns true.

The main algorithm accepts a list of conflicts, the
 relevant domain of time-objects, and an operation mode
 (reactive or proactive). It returns a triple (appoverlaps,
 unresolvedc, revokedobs) where appoverlaps gives the
 conflicts that it has not investigated because the relevant
 overlap may well be apparent, unresolvedc gives the
 conflicts that it has not managed to resolve, and
 revokedobs gives the time-objects whose existence is
 revoked in order to resolve the conflict.

3.3 Derivation functionalities

Derivation operations aim to derive, from the
 specified world of occurrences, new occurrences or new
 relations between occurrences. A derivation is
nondirected, if only the type is specified, or *directed* if a
 derivation pattern or template is specified which is
 matched against the particular world.

First we discuss the case of a semi-directed *merging*
derivation where the property concerned, ρ , and the
 time-axis, α , are specified. The property ρ is assumed to
 be *concatenable* [19], and it may well be a compound
 property such as a disjunction of related properties, eg ρ
 $\equiv \text{mild-coughing} \vee \text{moderate-coughing} \vee \text{severe-coughing}$.
 The algorithm is outlined below:

merging-deriv (ρ, α)
 do the following:
 1. select the relevant subset of time-objects from
 the particular world, $\text{relobs}(\rho, \alpha)$
 2. partition relobs into clusters that constitute
 maximal, conjoined, chains
 3. for each cluster, derive the time-object that
 subsumes it
 end
 return the list of associations (τ, cl) where τ is a
 newly derived time-object ($\pi_\tau = \rho$)

and cl is the cluster of existing time-objects
 subsumed by it
 end merging-deriv

The relevant subset of time-objects is defined as:
 $\text{relobs}(\rho, \alpha) = \{ \tau \mid \tau \in \text{Tobjects}(\alpha); \pi_\tau \Rightarrow \rho \}$. Thus the set
 of relevant periods of time on α , $I(\rho, \alpha)$, is given by:
 $I(\rho, \alpha) = \{ \epsilon_\tau(\alpha) \mid \tau \in \text{relobs}(\rho, \alpha) \}$. Let $\text{int} \subseteq I(\rho, \alpha)$; int
 constitutes a *conjoined chain* iff its elements collectively
 cover all time-values included in the minimal period of
 time that spans all elements of int :

conjoined-chain(int)
 $\Leftrightarrow \forall t \in \text{extent}(\text{int})$
 $(\exists \tau \in \text{int} \text{ such that } t \text{ is included in } \tau)$
 where $\text{extent}(\text{int})$ is the sequence of time-values

$\langle t_b, \dots, t_e \rangle$
 $t_b = \min_\mu \{ t_s \mid \langle t_s, t_f, \zeta \rangle \in \text{int} \}$
 $t_e = \max_\mu \{ t_f \mid \langle t_s, t_f, \zeta \rangle \in \text{int} \}$
 μ is the granularity of α

Let $\text{int} \subseteq \text{Int} \subseteq I(\rho, \alpha)$; int constitutes a *maximal*
conjoined chain in Int , iff it is a conjoined chain and no
 other subset of Int constitutes a conjoined chain that
 strictly includes int 's chain.

max-conj-chain(int, Int)
 $\Leftrightarrow (\text{conjoined-chain}(\text{int}) \wedge$
 $\neg (\exists \text{int}' \subseteq \text{Int} \text{ such that}$
 $(\text{conjoined-chain}(\text{int}') \wedge \text{extent}(\text{int}) \subset \text{extent}(\text{int}'))))$

Let $\text{int} \subseteq I(\rho, \alpha)$, where $\text{conjoined-chain}(\text{int})$. The
 minimal existence that spans all elements of int is given
 by function *chain*:

chain(int) = $\langle t_s, t_f, \zeta \rangle$
 where $t_s = \min_\mu \{ t_s \mid \langle t_s, t_f, \zeta \rangle \in \text{int} \}$
 $t_f = \max_\mu \{ t_f \mid \langle t_s, t_f, \zeta \rangle \in \text{int} \}$
 $c_1 = (\langle t_s, t_f, \text{closed} \rangle \in \text{int})$
 $\vee (\langle t_s, t_f, \text{open-from-right} \rangle \in \text{int})$
 $c_2 = (\langle t_s, t_f, \text{closed} \rangle \in \text{int})$
 $\vee (\langle t_s, t_f, \text{open-from-left} \rangle \in \text{int})$
 $\zeta = \text{closed}$, if $c_1 \wedge c_2$
 $\zeta = \text{open-from-right}$, if $c_1 \wedge \neg c_2$
 $\zeta = \text{open-from-left}$, if $\neg c_1 \wedge c_2$
 $\zeta = \text{open}$, if $\neg c_1 \wedge \neg c_2$

Let $\text{int} \subseteq I(\rho, \alpha)$. The minimum set of maximal
 existences that collectively cover all elements of int is
 given by function *merge*:

merge(int) = $\{ \text{chain}(i) \mid i \in \text{power}(\text{int});$
 $\text{max-conj-chain}(i, \text{int}) \}$

The refined algorithm is now given.

Algorithm for deriving time-objects through merging

merging-deriv (ρ, α)
 let result be initialised to the empty list
 if ρ is a concatenable property then do:
 let $\text{relobs} = \{ \tau \mid \tau \in \text{Tobjects}(\alpha); \pi_\tau \Rightarrow \rho \}$
 $\text{int} = \{ \epsilon_\tau(\alpha) \mid \tau \in \text{relobs} \}$
 $\text{maxexist} = \text{merge}(\text{int})$
 repeat for each $\text{maxe} \in \text{maxexist}$
 let $\text{ts} = \{ \tau \mid \tau \in \text{relobs};$

```

    extent { $\varepsilon_\tau(\alpha)$ }  $\subseteq$  extent {maxe}
    if  $\exists$  (maxe',ts)  $\in$  result
    then replace this entry of result with
        the new entry (chain {maxe, maxe'}, ts)
    else enter the association (maxe,ts) into result
    end repeat
    return result
end merging-deriv

```

As a side point we define strong and weak chains of time-objects. Let $ts \subseteq \text{relobs}(\rho, \alpha)$, $\text{int} = \{\varepsilon_\tau(\alpha) \mid \tau \in ts\} \subseteq I(\rho, \alpha)$, such that $\text{conjoined-chain}(\text{int})$. A conjoined chain formed by the existences of a set of time-objects is considered *strong* if the chain is not likely to break due to potential uncertainties regarding its elements:

```

strong-chain(ts)  $\Leftarrow$  conjoined-chain(int')
where int' = {< le-fr $_\tau(\alpha)$ , ri-fr $_\tau(\alpha)$ , closed>
              |  $\tau \in ts$ ; le-fr $_\tau(\alpha) \leq$  ri-fr $_\tau(\alpha)$ }

```

Thus, weak-chain(ts) $\equiv \neg$ strong-chain(ts)

Next we discuss one form of temporal data abstraction, persistence derivation. The problem of persistence derivation is explained through an example. Suppose there is some patient whose record includes two radiographs of the spine, one taken at the age of 2 years and the other at the age of 4 years. Both radiographs show the presence of abnormality "kyphoscoliosis". These observations correspond to two point-objects at the granularity of years. The problem is how to see beyond these two discrete sightings, both forwards and backwards in time, eg what can be inferred about the existence of this condition with regard to the particular patient, before the age of 2 years, between the ages of 2 and 4 years, and after the age of 4 years? Disorder expectations usually refer to interval occurrences and hence persistence derivation is a necessary functionality for the proper matching of a disorder profile against a patient profile.

The algorithm given below operates in a nondirected fashion; it accepts a time-axis, α , and a collection of time-objects, which have the same property, and exist as point-objects under α , and returns the interval-objects that represent distinct, maximal occurrences of the particular property, under the given temporal context.

Algorithm for persistence derivation

```

derive-persistence  $\alpha$  tobs
let  $\mu$  be the time-unit for  $\alpha$ 
 $\rho$  be the property shared by the elements of tobs
 $\tau_1, \tau_2, \dots, \tau_i, \dots, \tau_n$  be the elements of tobs
    in ascending temporal order
 $t_1, t_2, \dots, t_i, \dots, t_n \in \text{Times}(\alpha)$  be the
    respective positions of  $\tau_1, \tau_2, \dots, \tau_i, \dots, \tau_n$  on  $\alpha$ 
iobs be initialised to the empty set
if  $\rho$  is an infinitely persistent property ;; case 1
then enter into iobs a single interval-object with
     $t_f = \text{upper-time}(\alpha)$  ;; termination
     $t_s = \text{earliest-init}(\rho, \alpha)$  ;; earliest initiation
     $t_s = \min_\mu \{\text{latest-init}(\rho, \alpha), t_1\}$  ;; latest initiation

```

```

    if  $t_s > t_s$  then a conflict is noted
elseif  $\rho$  is a finitely persistent, but not a
    reoccurring, property ;; case 2
then enter into iobs a single interval-object with
     $t_s = \max\{\text{earliest-init}(\rho, \alpha), (t_n - \mu \text{max-dur}(\rho, \alpha))\}$ 
     $t_s = \min\{t_1, t_n - \mu \text{min-dur}(\rho, \alpha), \text{latest-init}(\rho, \alpha)\}$ 
     $t_f = \max\{(t_s + \mu \text{min-dur}(\rho, \alpha)), t_n\}$ 
     $t_f = (t_s + \mu \text{max-dur}(\rho, \alpha))$  ;; earliest termination
    if  $t_s > t_s$  or  $t_f > t_f$  then a conflict is noted
else ;;  $\rho$  is a finitely persisting, reoccurring,
    ;; property (case 3)
    let uncovered-obs be initialised to the
        temporal sequence  $\tau_1, \tau_2, \dots, \tau_i, \dots, \tau_n$ 
    repeat
        1. select the longest initial segment of
            uncovered-obs that constitutes a
            non-conflicting instance of case 2
        2. enter the interval-object that corresponds
            to the selected segment into iobs
        3. remove selected segment from
            uncovered-obs
    until uncovered-obs is the empty sequence
    return iobs
end derive-persistence

```

The clustering of the point-objects performed under case 3 determines significant temporal distances (on the basis of the property concerned) between neighbouring objects which signify separate incidents. The derivation of temporal trends is based on a modification of the above algorithm [7].

4 Validation

The original motivation for the ontology came mainly through the SDD system, a diagnostic expert system for the domain of skeletal dysplasias [11], although many of the temporal requirements identified for medical problem solving were in parallel detected in the context of other research projects whose scope was the modelling of biochemical and other industrial processes [4],[6]. In all these applications the relevance of temporal notions such as periodicity, trends, delays, prematurity, etc, was quite evident and hence of the notions of temporal distance and absoluteness, as well as the need for different granularities and conceptual temporal contexts. In addition, through such practical work, it was appreciated that the ontological primitives had to be at a high level of abstraction (knowledge level) to adequately support the knowledge engineering of complex dynamic processes. The original ideas, which were demonstrated through SDD's temporal reasoner [10], have been considerably enhanced and consolidated in the context of the CEC AIM Project GAMES II whose objective was to develop a theoretical framework and practical tools for the development of medical knowledge-based systems. The proposed ontology and temporal reasoning mechanisms constitute part of the overall GAMES II product [20].

Furthermore, the ontology has been applied in building a simple prototype system modelling the normal ossification processes for parts of the human skeleton, and has been utilised in formulating a temporal model for medical diagnostic reasoning [9]. The plan for the immediate future is to fully implement the ontology and associated temporal reasoning algorithms in the form of an efficient and reusable temporal kernel for knowledge-based problem solvers, primarily for medical applications, but for other applications as well.

5 Conclusions

Time is essential in problem solving. With regard to medical problem solving time is intrinsically relevant to all medical concepts [6] (disorders, patient data, therapeutic interventions) and hence temporal reasoning ought to constitute an integral aspect of the overall problem solving. Many of the general theories of time proposed in the AI literature [1],[3],[12] focus on natural language processing, or a specific class of problem solving tasks, usually planning or monitoring, or the management of temporal databases. Some of these general theories have been effectively adapted in the context of specific medical applications [2].

By and large the work done in temporal reasoning for medical applications focuses on temporal data abstraction and temporal causality [5],[14],[16],[18]. The overall aim of the research presented in this paper is to develop an ontology of time, and associated temporal reasoning, based on a global analysis of time representation requirements for different medical tasks (diagnosis, monitoring, therapy planning) [8] and to effectively implement these through a generic temporal kernel. The ontology must provide the required level of abstraction for knowledge modelling purposes and its usage must result in treating time as an integral aspect of the application. The chosen ontological primitives, the time-axis and the time-object, satisfy these design objectives and in addition the notion of a time-object integrates in a uniform and natural way temporal, structural, and causal knowledge. However, the high power of expression provided by the proposed ontology is associated with relevant computational overheads.

Acknowledgements

The research reported here has been partly supported by the University of Cyprus, the Leverhulme Trust (UK) (under projects SDD and MSD), and the CEC (under AIM Project A2034 GAMES II). The development of the ontology has benefit from discussions with other participants on these projects, in particular John Washbrook, Hauke Kindler and Mark Leaning. In addition Kazem Sadegh-Zadeh, John Mylopoulos, and the three unknown referees have offered useful suggestions.

References

- [1] J.F. Allen, "Towards a General Theory of Action and Time", *Artificial Intelligence*, Vol.23, pp.123-154, 1984.
- [2] L. Chittaro, M. Del Rosso, and M. Dojat M., "Modeling Medical Reasoning with the Event Calculus: an application to the management of mechanical ventilation", *Proc AIME 95*, Lecture Notes in Artificial Intelligence, Vol. 935, Springer, pp.79-90, 1995.
- [3] T. Dean and D. McDermott, "Temporal Data Base Management", *Artificial Intelligence*, Vol.32, pp.1-55, 1987.
- [4] G.A. Dervakos, J.M. Woodley, E.T. Keravnou, J. Washbrook and M.D. Lilly, "Development of a KBS for Biotransformation Process Design, *Proc. I.CHEME.E. Symposium Series No.114*, pp.283-291, 1989.
- [5] I.J. Haimowitz and I.S. Kohane, "Managing Temporal Worlds for Medical Trend Diagnosis", to appear in [8].
- [6] E.T. Keravnou, "Modelling Medical Concepts as Time-Objects", *Proc AIME 95*, Lecture Notes in Artificial Intelligence, Vol. 935, Springer, pp.67-78, 1995.
- [7] E.T. Keravnou, "An Ontology of Time Using Time-Axes and Time-Objects as Primitives", Technical Report, Department of Computer Science, University of Cyprus, 1995.
- [8] E.T. Keravnou (ed.), *Temporal Reasoning in Medicine*, special issue of *Artificial Intelligence in Medicine*, Vol.8, No.3, June 1996.
- [9] E.T. Keravnou E.T., "Temporal Diagnostic Reasoning Based on Time-Objects", to appear in [8]
- [10] E.T. Keravnou, J. Washbrook, "A Temporal Reasoning Framework Used in the Diagnosis of Skeletal Dysplasias", *Artificial Intelligence in Medicine*, Vol.2, No.5, pp.239-265, 1990.
- [11] E.T. Keravnou, F. Dams, J. Washbrook, C.M. Hall, R.M. Dawood and D. Shaw, "Modelling Diagnostic Skills in the Domain of Skeletal Dysplasias", *Computer Methods and Programs in Biomedicine*, Vol.45, pp.239-260, 1994.
- [12] R. Kowalski and M. Sergot, "A Logic-Based Calculus of Events", *New Generation Comput.*, Vol.4, pp.67-95, 1986.
- [13] P. Ladkin, "Primitives and Units for Time Specification", *Proc AAAI-86*, pp.354-359, 1986.
- [14] W. Long, "Temporal Reasoning for Diagnosis in a Causal Probabilistic Knowledge Base", to appear in [8].
- [15] J. Mylopoulos, A. Borgida, M. Jarke and M. Koubarakis, "Telos: A Language for Representing Knowledge About Information Systems (Revised)", *Technical Report KRR-TR-89-1*, Department of Computer Science, University of Toronto, 1990.
- [16] A. Riva and R. Bellazzi R., "Learning Temporal Probabilistic Causal Models from Longitudinal Data", to appear in [8].
- [17] B. Rosser, J. Washbrook, J. Campbell, E.T. Keravnou and D. Long, "A Framework for Time Dependent Reasoning Systems", Product P111-1, Esprit Project P2409 (EQUATOR), 1989.
- [18] Y. Shoham and M.A. Musen, "Knowledge-Based Temporal Abstraction in Clinical Domains", to appear in [8].
- [19] Y. Shoham, "Temporal Logics in AI: Semantical and Ontological Considerations", *Artificial Intelligence*, Vol.33, pp.89-104, 1987.
- [20] G. van Heijst, M. Ramoni, G. Schreiber and M. Stefanelli (eds.), CEC AIM Project A2034 GAMES-II Final Report, 1995.