Knowledge-Based Interpolation of Time-Oriented Clinical Data

Yuval Shahar

Section on Medical Informatics, Knowledge Systems Laboratory
Medical School Office Building (MSOB) X215, Stanford University, Stanford, CA 94305, USA
email: shahar@smi.stanford.edu Tel: 1-415-725-3393 Fax: 1-415-725-7944

Abstract

Temporal interpolation is the task of bridging gaps between time-oriented clinical data or abstracted concepts in a context-sensitive manner. It is one of the subtasks important for solving the **temporal-abstraction task**—abstraction of interval-based, higher-level concepts from time-stamped clinical data. We present a knowledge-based approach to the temporal-interpolation task. The temporal-interpolation mechanism we discuss relies, among other knowledge types, on a temporal-persistence model. This model employs local temporal-persistence functions that are temporally bidirectional (i.e., extend a belief measure in a predicate both into the future and into the past) and global, maximal-gap temporal-persistence functions that bridge gaps between interval-based predicates. We investigate the quantitative and qualitative properties implied by both types of persistence functions. Our goal is to formulate the knowledge required for solving the temporal-abstraction task, and in particular the temporal-interpolation subtask, so as to facilitate the acquisition of that knowledge, its maintenance, its reuse for the same task in different domains, and its sharing among different applications in the same domain. We have implemented our approach and evaluated it in several clinical domains.

1 Temporal-Abstraction and Temporal Interpolation

Time-stamped clinical data often need to be abstracted in a context-sensitive manner into more abstract, interval–based concepts, meaningful for a specific medical domain (e.g., oncology) and a particular task (e.g., monitoring of patients that are being treated by chemotherapy). We term this interpretation task the **temporal-abstraction (TA) task**. For instance, most clinical tasks require measurement and capture of numerous patient data. An automated, knowledge-based decision-support tool that assists physicians should provide short, informative, context-sensitive summaries, at various desirable levels of abstraction, of time-oriented clinical data stored on electronic media. Data abstraction assists both physicians and automated decision-support systems. A meaningful summary characterizes significant features over *periods* of time, such as "2 weeks of grade-II bone-marrow toxicity in the context of therapy for potential complications of a bone-marrow transplantation event" (Figure 1).

Solving the TA task involves the solution of several subtasks (see Section 2). One of these tasks is the **temporal-interpolation** task: bridging gaps between point- or interval-based temporal predicates of a similar-type that are temporally disjoint, to create longer intervals (see Figure 1). Temporal interpolation requires, among other knowledge types, some measure of **temporal persistence** of temporal predicates (denoting either raw data or abstract concepts). For instance, if we measured hemoglobin levels on Tuesday and on Friday, both being abstracted as LOW, was the patient's hemoglobin level on Thursday also LOW?" In fact, the very notion of an *episode* implies some form of bounded persistence of concepts over time, preventing the clumping together of similar, but distinct instances of the same concept. The concept of persistence was addressed previously; we discuss the relationship of such work to ours in Section 6.

2 Knowledge-Based Temporal Abstraction

The framework we employ for solving the TA task is the **knowledge-based temporal-abstraction (KBTA) method** [Shahar, 1997]. The KBTA method is a general *problem-solving method* [Eriksson *et al.*, 1995] for interpreting data in time-oriented domains, with clear semantics for both the *method* and its domain-specific *knowledge* requirements. The KBTA method comprises a knowledge-level representation of the TA task and of the knowledge required to solve that task. The KBTA method has a formal model of input and output entities, theirrelations, and properties associated with these entities—the **KBTA ontology.**

The KBTA method decomposes the TA task into five parallel *subtasks* :

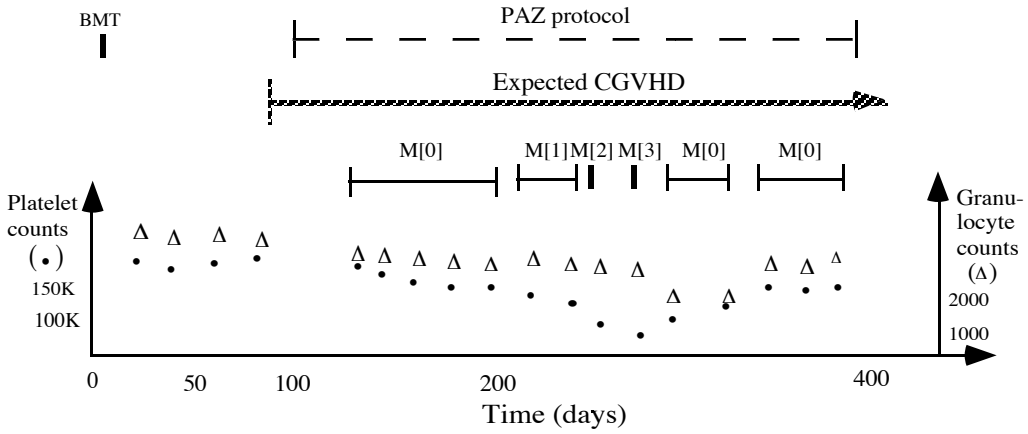

Figure 1: Typical inputs to and outputs of the temporal-abstraction task. The figure presents examples of abstractions of platelet and granulocyte values during administration of the PAZ clinical protocol for treating patients who have chronic graft-versus-host disease (CGVHD). The time line starts with a bone-marrow transplantation (BMT) event. ⊢ ⊣ = event; • = platelet counts; Δ = granulocyte counts; ⊢- - -→ = open context interval; ⊢⊣ = closed abstraction interval; M[n] = myelotoxicity (bone-marrow–toxicity) grade n.

(1) **temporal-context restriction**: creation of relevant contexts for interpretation of data (e.g., effect of a drug), crucial for focusing and limiting the scope of the inference

(2) **vertical temporal inference**: inference from values of contemporaneous input data or abstractions (e.g., results of several blood tests conducted during the same day) into values of higher-level concepts (e.g., classification into bone-marrow toxicity Grade II)

(3) **horizontal temporal inference**: inference from similar-type propositions that hold over different time intervals (e.g., joining different-value abstractions of the same parameter that hold over two meeting time intervals and computing the value of the new abstraction)

(4) **temporal interpolation**: bridging of gaps between similar-type but temporally disjoint point- or interval-based propositions to create longer intervals (e.g., joining two disjoint episodes of anemia, occurring during different days, into a longer episode)

(5) **temporal-pattern matching:** creation of intervals by matching patterns over disjoint intervals over which hold propositions of various types.

The five subtasks of the KBTA *method* are solved by five **temporal-abstraction mechanisms** (nondecomposable computational modules), which depend on four domain-specific **knowledge types**: *structural, classification* (functional), *temporal-semantic* (logical), and *temporal-dynamic* (probabilistic) knowledge. Values for the four knowledge types are specified as the domain's **temporal-abstraction ontology**. The KBTA method has been implemented in the **RÉSUMÉ** system and evaluated encouragingly in several medical and domains [Shahar and Musen, 1996] and even in an engineering domain [Shahar and Molina, 1996].

In this paper, we analyze one of the key TA subtasks in clinical domains: context-specific temporal interpolation. First, we define briefly the KBTA ontology, and then discuss the temporal-interpolation mechanism which uses that ontology and analyze its theoretical foundations and the implications of the approach for acquisition and maintenance of temporal-dynamic knowledge.

3 The Knowledge-Based Temporal-Abstraction Ontology

The KBTA temporal model includes both time intervals and time points. *Time points* are the basic temporal primitives, but propositions can be interpreted only over *time intervals*. Therefore, all propositions are fluents [McCarthy and Hayes 1969] and in our model must be interpreted over a particular time period. The KBTA ontology [Shahar, 1997] contains the following entities:

1. *Time stamps*, $\tau_i \in T$, comprise the basic primitives of time. A time-standardization function, $f_s(\tau_i)$, can map a time stamp into an integer amount of any pre-defined temporal granularity unit $G_i \in \Gamma$ (e.g., hour). Time stamps are measured in G_i units with respect to a zero-point time stamp. A finite positive or negative amount of G_i units is a *time measure*.
2. A *time interval* is an ordered pair of time stamps that denote the endpoints, [*I*.start, *I*.end], of the interval. A zero length interval in which *I*.start = *I*.end is a time point.
3. An *interpretation context* $\xi \in \Xi$ is a proposition representing a relevant state of affairs (e.g., "the drug insulin exerts its effect during this interval"), within which certain parameters may be interpreted differently. IS-A and SUBCONTEXT relations are

defined over the set of interpretation contexts. *Basic* interpretation contexts are atomic propositions. *Composite interpretation contexts* are created by the temporal intersection of a basic or a composite interpretation context and one of its subcontexts, and enable a definition of increasingly specific interpretation contexts.

4. A *context interval* is a structure $<\xi, I>$ (i.e., interpretation context ξ holds during I).
5. An *event proposition* or event $e \in E$ is the occurrence of an external willful act or process, such as the administration of a drug. Events are instantiated event schemata; an event schema has a series a_i of event attributes (e.g., drug dose) that must be mapped to attribute values v_i. A PART-OF (or *subevent*) relation is defined over event schemata.
6. An *event interval* is a structure $<e, I>$ represents the occurrence of event e during I.
7. A parameter schema or *parameter* $\pi \in \Pi$ is a measurable or describable state of the patient. Parameters may represent raw input data (e.g., hemoglobin level) or abstractions from the raw data (e.g., state of hemoglobin). Parameter schemata have various properties, such as a domain V_π of possible symbolic or numeric values and measurement units. An *extended parameter* is a combination $<\pi, \xi>$ of a parameter π and an interpretation context ξ. An extended parameter can have a value $v \in V_\pi$, which is typically known only at runtime (i.e., parameter values require a context). A *parameter proposition* is the combination of a parameter, a parameter value, and an interpretation context, $<\pi, v, \xi>$ (e.g., "the state of hemoglobin is LOW in the context of chemotherapy"). Parameter propositions can have special properties, such as temporal persistence.
8. A *parameter interval* $<\pi, v, \xi, I>$ represents the fact that the value v of parameter π in a specific context ξ holds during interval I.
9. An *abstraction function* $\theta \in \Theta$ is a unary or multiple-argument function that takes one or more parameters as input and returns an abstract parameter. The abstract parameter may be one of three abstraction types: *state*, *gradient*, and *rate*. An additional abstraction type is *pattern* which defines a temporal pattern of several other parameters. An abstraction of a parameter (e.g., state(π)) is a parameter (e.g., hemoglobin value and the state of hemoglobin value are different parameters).
10. An *abstraction* is a parameter interval $<\pi, v, \xi, I>$ where π is an abstract parameter.
11. An *abstraction goal* $\psi \in \Psi$ is a proposition that indicates an intention relevant to the TA task (e.g., the intention to control a diabetes patient's blood-glucose values). Typically, it creates a context.
12. An *abstraction-goal interval* is a structure $<\psi, I>$, where ψ is an abstraction goal that is posted during the interval I.
13. Interpretation contexts are *induced* or inferred dynamically from event, parameter, or abstraction-goal propositions. The time intervals over which the inducing propositions hold impose temporal constraints on the interval in which the inferred context will be valid (e.g., the interpretation context of the effect of an AZT therapy event might begin 2 days following its start and end 2 weeks after its termination).

The **TA ontology** of a domain describes all potentially relevant (for the TA task) events, parameters, contexts, abstraction-goals, and relations (e.g., induction of contexts). The **TA task** is thus the following: Given a set of event, parameter, and goal intervals and the domain's TA ontology, produce an interpretation—a set of new abstractions that can answer any temporal query about all the abstractions derivable from the transitive closure of the input data and the domain's TA ontology. (A *temporal query* is a set of temporal and value constraints over the components of a set of parameter and context intervals.)

4 The Temporal-Interpolation mechanism

The temporal-interpolation subtask can be solved by a knowledge-based **temporal-interpolation mechanism.** The temporal-interpolation mechanism accepts as input two parameter points, two parameter intervals, or a parameter interval and a parameter point, and returns as output an abstraction, interpreted over a superinterval of the input's time points or intervals, interpolating over the gap between these time intervals. **Primary interpolation** accepts two parameter points and returns an abstraction interval. **Secondary interpolation** accepts two parameter intervals (or a parameter interval and a parameter point), and returns an abstraction (super)interval. Both interpolation types are relevant to primitive parameters and to all abstraction types (e.g., gradient)). Thus, **secondary gradient interpolation** infers, from two gradient-abstraction intervals of parameter π, a gradient-abstraction superinterval of π whose value is INCREASING, DECREASING, SAME, NONDECREASING, NONINCREASING, or NONMONOTONIC.

Temporal interpolation requires that the temporal distance between the two time points or intervals of the parameter propositions be less than a certain time *gap*. Within that time gap, a certain value of the parameter is then be assumed to hold.

The maximal allowed gap is a domain-, task-, and context-dependent function (e.g., the maximal allowed gap for LOW hemoglobin in the domain of oncology, the task of caring for patients using protocols, and the interpretation context of patients receiving X-ray therapy). The arguments of the maximal-gap function also include a measure of the rate of change of the parameter before and after the time gap;

as an approximation, we use the length of the intervals before and after the gap. A maximal-gap function Δ is a function $\Delta(\pi, \nu, L(I_1), L(I_2), \xi)$ of a parameter π (assuming that π includes its abstraction type) and lengths $L(I_1)$, $L(I_2)$ of the intervals I_1 and I_2, to be joined in the context ξ into an interval with an abstraction value ν. The Δ function returns the length of the maximal temporal gap that still allows interpolation between I_1 and I_2. For instance, in any context, joining two intervals where the hemoglobin-state abstraction was classified as LOW into a longer interval whose hemoglobin-state abstraction is classified as LOW depends on the time gap separating the two intervals, on the particular context, and on the length of time in which the LOW property was known both before and after the time gap. Primary interpolation is the initial *constructor* of abstraction intervals, since it joins two separate time points T_1 and T_2 into a new interval $[T_1, T_2]$, over which ν is true for π.

Thus, a necessary requirement for primary interpolation is that $L([T_1, T_2]) \leq \Delta(\pi, \nu, 0, 0, \xi)$, where $L(I)$ is the length of I.

A prerequisite to an interpolation operation is that the value ν of the parameter π is has the value TRUE for the *concatenable* inferential property [Shoham, 1987] in the context ξ (i.e., the parameter propositions involved can indeed be joined). This prerequisite involves *temporal-semantic knowledge*. We summarize the *temporal-semantic knowledge* for a domain in an **inference-properties table** [Shahar *et al.*, 1992], a relation in which every tuple $(\pi, \nu, \phi, \omega, \xi)$ represents the knowledge that the temporal-semantic property $\phi \in \Phi$, for value ν, of parameter π, in the context ξ, has the truth value ω ($\omega \in$ {TRUE, FALSE}) (π is assumed here to include its abstraction type).

Similarly, deciding *what* is the value of the resulting abstraction when joining two abstraction intervals with different values, ν_1 and ν_2, of the same parameter π requires using *horizontal classification knowledge*. A **horizontal-inference table** [Shahar *et al.*, 1992] is a relation that includes tuples of the form $(\pi, \nu_1, \nu_2, \nu_3, \xi)$, meaning that, for parameter π (assuming that π includes its abstraction type), when an abstraction interval with parameter value ν_1 meets an abstraction interval with parameter value ν_2, in the context ξ, the value of the parameter of the joined abstraction interval should be ν_3. That is, $\nu_1 \oplus \nu_2 = \nu_3$. In a horizontal-inference table, it is assumed that concatenated abstractions are of the same type—for instance, state (e.g., HIGH or LOW) or gradient (e.g., INCREASING $\oplus$ SAME = NONDECREASING). The $\oplus$ operator is the **horizontal-join** operator. In the case of joining different values, both the temporal-semantic knowledge (inferential property) and the temporal-dynamic knowledge (Δ function) that are used for interpolation are those specific to the value ν_3.

Secondary state, gradient, and rate interpolation require additional conditions to preserve consistency, apart from an upper bound on the temporal gap between intervals. An **interpolation-inference table** defines the interpolation operation for every relevant parameter (e.g., hemoglobin-state) and value combination (e.g., INCREASING and SAME). An interpolation-inference table represents horizontal-classification knowledge, persistence knowledge, and the special temporal conditions that should hold between the temporal elements of the involved abstractions for successful interpolation.

For example, we need to check that, when we use secondary temporal interpolation to join two INCREASING abstractions for π that are true over two intervals I_1, I_2, into a INCREASING abstraction for π over a superinterval I_j, the value of π has indeed increased, or at least has not decreased below a certain predefined threshold during the time gap [I_1.end, I_2.start] (see Figure 2). In other words, we have to check that $I_1.\text{end}.\pi \leq I_2.\text{start}.\pi + C_\pi$, where C_π represents a measurement variation for π—the maximal decrement in parameter π, below which a change in π will not be considered as a decrease. C_π can be interpreted as a measurement error of π, or as a natural random variation of π over time, or as a significant change of π, for a particular task, depending on the context. In general, C_π is a function of π, $f_c(\pi)$, that is defined either by the domain expert or through analysis of the distribution of π. In principle, $f_c(\pi)$ might also use a context argument ξ and the initial value of π, $I_1.\text{end}.\pi$ (e.g., what is considered as a significant variation in the value of the hemoglobin-value parameter might have a different value within the interpretation context BONE-MARROW DEPRESSION, and furthermore, when the last hemoglobin value known is abstracted as VERY LOW).

Primary temporal interpolation for the INCREASING gradient abstraction, requires that $T_2.\pi - T_1.\pi \geq C_\pi$. Primary temporal interpolation for the DECREASING gradient abstraction requires that $T_1.\pi - T_2.\pi \geq C_\pi$. Primary temporal interpolation for the SAME gradient abstraction requires that $|T_2.\pi - T_1.\pi| \leq C_\pi$.

Using the C_π property, we can ignore minor absolute changes in the value of π that are less than a certain threshold when we wish to identify general qualitative trends.

5 Local and Global Persistence Functions

The maximal-gap Δ functions, which allow interpolation between point and interval primitive and abstract parameters, can be interpreted as creating a default abstraction during the maximal-gap interval. Like all conclusions inferred by the temporal-abstraction mechanisms, the inference that creates such default abstractions is nonmonotonic and can be overridden by additional data or by other inferences. The maximal-gap functions represent domain- and task-dependent knowledge regarding the rate of change of a parameter proposition $<\pi, \nu, \xi>$ over time, or the **persistence** of the truth of that proposition over a temporal gap. In general, however, we distinguish two types of persistence functions: **Local** (ρ) **persistence functions** and **global** (Δ) **functions**. For the purpose of the following

discussion, we assume that the context ξ and the value of π, unless mentioned explicitly, are known.

5.1 Local Persistence Functions

Local (ρ) persistence functions represent the local persistence of the truth of a parameter proposition, given a single parameter point or interval: $\rho(\pi, L(I), t)$, where $L(I)$ is the length of the interval I during which the parameter proposition is known to hold, and t is the time since an endpoint of I. The ρ function returns a degree of belief—a probability distribution—in the proposition $<\pi, \nu>$ being true at time $t_0 + t$, given that $<\pi, \nu>$ was true at endpoint t_0. The ρ function extends a proposition temporally in *both* directions: to the *future* and to the *past*. Assuming that time t_0 is a random (first) time in which the proposition was measured, there is no particular reason to assume that a parameter proposition was not true *before* time t_0. Thus, t might actually have a *negative* value. We need this extension if we are to include an approximation of the past value of a parameter, for purposes of interpretation, as opposed to forecasting a future value of the parameter. Thus, our model includes both **forward decay** and **backward decay** in belief. The function describing this decay is equivalent to a statistical **survival function**.

In practice, the important question for performing an interpolation using a local persistence function is how long t can be before the belief in the parameter proposition $\varphi \in P$ (i.e., its probability) drops below a certain context-specific threshold φ_{th} (Figure 2).

5.2 Global Persistence Functions

Global (Δ) maximal-gap functions bridge the gap between two propositions. Δ functions are an extension of ρ functions, and, in special cases, as we show in this section, they can be constructed from the latter functions. The Δ function returns the maximal time gap that still allows us to join the propositions into an abstraction that is believed to be true, with a sufficient, task-specific, predefined degree of belief in the proposition, during the gap (and thus over a superinterval of the input propositions, given that both were true for some time before and after the gap). Thus, the Δ functions are a global extension of the local (ρ) persistence functions, since they assume both forward and backward decay of the propositions involved.

Figure 2 presents a graphic view of the Δ function as an interpretation of a decay in the belief in the truth of a proposition. For instance, in the case that the abstractions' parameter values are identical—that is, the propositions are the same before and after the gap interval—and the forward and decay times are relatively independent, we are interested in whether, at all points inside the gap interval, *either* of the values, approximated by the forward belief decay in proposition φ, $BEL_{forward}(\varphi)$, or by the backward belief decay, $BEL_{backward}(\varphi)$, is true with a probability $p \geq \varphi_{th}$. As the time gap Δt between the two abstractions increases, the belief that either the backward- or forward- decay value is true will eventually fall below the predefined threshold value φ_{th} (see Figure 2).

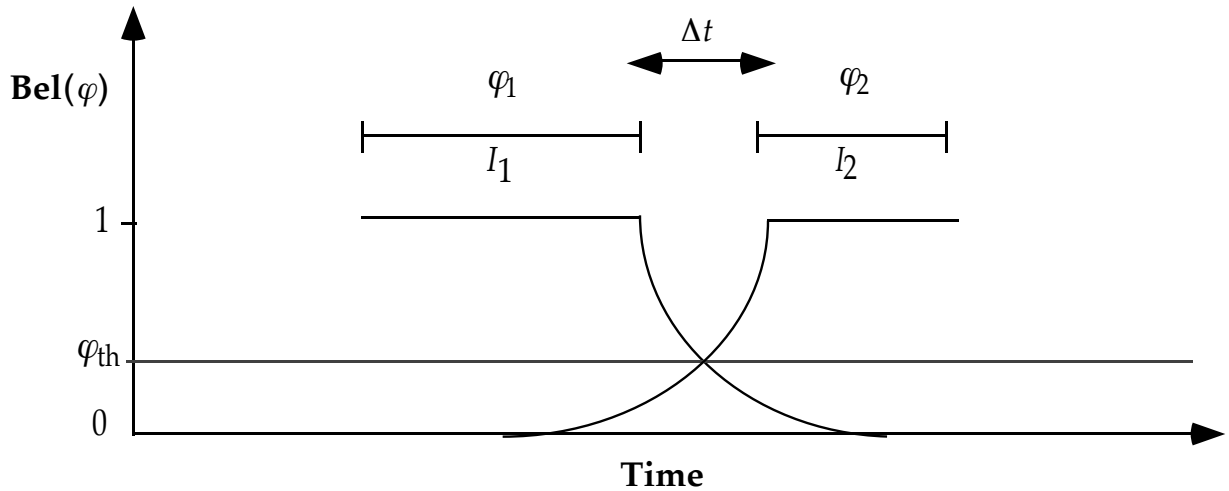

Figure 2: Local and global persistence functions. The maximal time gap Δt returned by a global Δ function is used to decide whether the parameter propositions φ_1 and φ_2, attached to intervals I_1 and I_2, can be joined (possibly, if they do not denote the same value of the relevant parameter, into a new proposition $\varphi_3 = \varphi_1 \oplus \varphi_2$). The time gap Δt can be interpreted—in the case that $\varphi_1 \equiv \varphi_2$, and that the truth values of the propositions are relatively independent—as the maximal time gap in which the belief produced by either the local forward or backward decay (represented by a local persistence ρ function) stays above the predefined confidence threshold φ_{th}. Bel(φ) = degree of belief in φ; φ_{th} = the task- and context-specific belief threshold value.

If the local (ρ) persistence function is an exponential-decay survivor function and the backward- and forward-decay rates are independent, we can compute the Δ function. Assume that the probability $p(t)$ of the parameter proposition φ being true is $e^{-\lambda t}$, a function of the time t since the reference time in which P was true, regardless of the length of the time interval I during which φ was true. Let the forward decay rate be λ_1 and the backward decay ratebe λ_2. Then, we need to know the maximal gap Δt such that, in the point of minimal belief, $p(t)$ is at or above the threshold φ_{th}. Note that the minimum point of $BEL_{forward}(\varphi)$ or $BEL_{backward}(\varphi)$ is when the values of the forward- and backward-decay functions are equal (see Figure 2).

Thus, at the minimal $p(t)$,

$$BEL_{forward}(\varphi) = BEL_{backward}(\varphi),$$

that is,

$$e^{-\lambda_1 t} = e^{-\lambda_2(\Delta t - t)},$$

so, when $p(t)$ is minimal,

$$t = [\lambda_2/(\lambda_1 + \lambda_2)]\ \Delta t;$$

but $p(t) \geq \varphi_{th}$ implies, after substituting for t in $BEL_{forward}(\varphi)$, that

$$e^{-[(\lambda_1 * \lambda_2)/(\lambda_1 + \lambda_2)]\Delta t} \geq \varphi_{th} = e^{-K},$$

and thus

$$\Delta t \leq [(\lambda_1 + \lambda_2)/(\lambda_1 * \lambda_2)]\ K, \quad K = -\ln\varphi_{th}.$$

In other words, the Δ function for two parameter points, $\Delta(\pi, 0, 0)$, or for two parameter intervals when the duration of the intervals has no effect on the persistence of the propositions, is a constant determined by the forward- and backward-decay rates and the desired level of confidence.

We can generalize this analysis. Assume that the longer φ is known to be true in the past or in future, the longer we are likely to keep believing it or to believe that it already existed in the past, before we measured it (this assumption will be discussed in Section 5.3). One (not necessarily the only) way to represent that assumption would be to modify the decay rate λ by assuming that it is inversely proportional to the length of the relevant intervals, $L(I_i)$, which we denote simply as L_i. Let

$$BEL\ (P) = e^{[-\lambda_i/L_i]t}, \quad i = 1,2.$$

So, if $p(t)$ is minimal, and as before, $BEL_{forward}(\varphi) = BEL_{backward}(\varphi)$,

$$e^{[-\lambda_1/L_1]t} = e^{[-\lambda_2/L_2](\Delta t - t)};$$

that is, when $p(t)$ is minimal,

$$t = [(L_1\lambda_2)/(\lambda_1 L_2 + \lambda_2 L_1)]\Delta t.$$

Substitute for t in $BEL_{forward}(\varphi)$, and assume $p(t) \geq \varphi_{th}$:

$$\Delta t \leq [(\lambda_2 L_1^2 + \lambda_1 L_1 L_2)/\lambda_1\lambda_2 L_1]K, \quad K = -\ln\varphi_{th}.$$

For instance, if $\lambda_1 = \lambda_2 = \lambda$ and $L(I_1) = L(I_2) = L$, then

$$\Delta t \leq [(\lambda L^2 + \lambda L^2)/\lambda^2 L]K;$$

that is,

$$\Delta t \leq [2L/\lambda]K, \quad K = -\ln\varphi_{th}.$$

In other words, if exponential decay rates decrease (equally) linearly forward and backward as a function of the duration of the proposition, then the maximal time gap allowing us to join equal-length abstractions would be proportional to a linear function of the length of either interval, with the rest of the factors kept constant. The duration of the gap would be inversely proportional to the uniform decay rate.

These simplified examples serve to show that even though the decay rates λ_i are in general unknown, and the decay function is perhaps difficult to compute, the resulting global Δ function (using a belief threshold) might be a simple constant or polynomial, and thus can be more easily described, computed, or acquired, than the underlying local-persistence function.

Furthermore, if there is evidence for a particular type of decay function (e.g., logarithmic), we can compute the latter's coefficients by acquiring from the domain expert a few maximal-gap values—that is, several examples of Δt. We might even check the expert's consistency (or the adequacy of the decay function) by repeating the calculation for several other examples. Alternatively, we can simply acquire a table of typical Δt values for various common $L(I_1)$ and $L(I_2)$ values, and can interpolate between these values, or extrapolate from them, when necessary.

Due to the dependence between the forward decay of a parameter proposition attached to one time point and the backward decay of that proposition at a later time point, and, therefore, an implied *joint distribution* of the forward and backward belief values, we usually need the actual global (Δ) function, in addition to (or instead of) the local (ρ) persistence function. (In the example above, we in fact computed a *lower bound* for the Δ function.) In practice, the domain expert often knows several Δ function values (such as what is the maximal time gap allowed in order to join two parameter points for several parameter values in each context), even if she cannot define any particular, precise, local-decay function ρ (except, possibly, for specifying the forward and backward local decay times Δt corresponding to reaching the local threshold value φ_{th}). Knowing only the global Δ function still enables interpolation between two point-based or interval-based parameter propositions. In view of the preceding discussion, in many domains, knowing only the values needed to maintain $Bel(\varphi)$ above the threshold value φ_{th}—that is, the (simpler) Δ function—would be a common state of affairs.

5.3 A Typology of Persistence Functions

Global (Δ) persistence functions can have four **qualitative types**, depending on whether the Δ function is either (1) positive monotonic or (2) negative monotonic, with respect to (a) the length of the first parameter interval $L(I_1)$ or (b) the length of the second parameter interval $L(I_2)$ (see Figure 2). (For example, the maximal allowed gap might be longer, the longer the interval before the gap.) Theoretically, there are **positive-positive** (**PP**), **positive-negative** (**PN**), **negative-positive** (**NP**), and **negative-negative** (**NN**) monotonic Δ functions. We refer to these categories as **qualitative persistence types**.

Formally, PP Δ functions are functions such that

$$L(I') > L(I) \Rightarrow \forall i\ [\Delta(I', i) \geq \Delta(I, i) \wedge \Delta(i, I') \geq \Delta(i, I)]$$

NN Δ functions are functions such that

$$L(I') > L(I) \Rightarrow \forall i\ [\Delta(I', i) \leq \Delta(I, i) \wedge \Delta(i, I') \leq \Delta(i, I)]$$

where $L(I)$ is the length of interval I and $\Delta(I, i)$ stands for $\Delta(L(I), L(i))$.

In the case of local (ρ) persistence functions, whether representing backward or forward local persistence, we can categorize functions qualitatively into **positive** (**P**) and **negative** (**N**) categories with similar meaning (i.e., whether the longer I, the longer or shorter the relevant validity interval, before or after I).

Most Δ functions, in practice, seem to be of the PP type. In other words, the longer we know that a parameter proposition was true either before or after a time gap, the longer we would allow that gap to be while maintaining our belief that the parameter proposition stayed true throughout that gap (i.e., its probability was always above a certain threshold). (For instance, the proposition denoting the MODERATE-ANEMIA value of the hemoglobin-state parameter usually would be associated with a PP Δ function, as would be the proposition denoting the DEEP-COMA value of the consciousness parameter).

Negative-monotonic Δ functions occur when a longer duration of either I_1 or of I_2 lowers the probability that the abstraction was true during the gap, and the longer the lengths, the shorter the allowed Δt. For instance, knowing about a longer I_1 interval of an almost-fatal cardiac arrhythmia (say, ventricular fibrillation) actually *lowers* the probability that the (following) gap interval had the same characterization, given the same I_2 interval and assuming that the patient is alive. Most of the negative-monotonic functions emerge from a total-length constraint on the time allowed for the abstraction (or an analogous probabilistic distribution on the expected total time), or from a total cardinality constraint on the number of events allowed.

We often can limit ourselves, as a first approximation, to the common PP Δ functions. Note that the exponential-decay local (ρ) functions that were given as an example in Section 5.2 for decay functions dependent on the length of either of the two intervals implied, with the independence assumption, a PP-type Δ function. However, there is also an important computational advantage in adhering to PP Δ functions.

lemma 1: PP Δ functions are associative. (The order of joining intervals and points cannot change the resulting set of abstractions.)

Proof: Assume a situation where parameter points T_1, T_2, and T_3 exist in that temporal order. If we can form both the parameter interval $[T_1, T_2]$ and the parameter interval $[T_2, T_3]$, then, if we can eventually form the interval $[T_1, T_3]$, we can do so by forming initially either subinterval, since the Δ function is PP. That is, if we can join one point to another, we can certainly join that point—forwards or backwards, as necessary—to an interval starting or ending, respectively, with the other point. For instance,

$$L([T_1, T_2]) \leq \Delta(0,0) \Rightarrow L([T_1, T_2]) \leq \Delta(0, L([T_2, T_3])),$$

since the Δ function is PP, and therefore $\Delta(0,0) \leq \Delta(0, L([T_2, T_3]))$.

A similar argument holds for any four consecutive points.

Thus, the claim is true for any sequence of primary or secondary interpolations, since Δ functions are applied only when there are no intervening points between the two intervals or points to be joined. ❑

The associativity property is important for data-driven systems, in which the order of the parameter intervals the system reasons with might be arbitrary. This property is necessary also to guarantee that the final abstractions do not depend on the order of *arrival* of the input data.

lemma 2: NN Δ functions are not associative.

Proof: It is easy to construct a case for consecutive parameter points T_1, T_2, and T_3, where, if we create the interval $[T_1, T_2]$, we no longer can join it to T_3, and if we create the interval $[T_2, T_3]$, the Δ function value will prevent our joining it to T_1 (e.g., a total-sum constraint does not allow creating the interval $[T_1, T_3]$ with high enough probability). ❑

NP and PN functions cannot be associative for similar reasons. Whether such functions can even exist is doubtful, and we leave it as an open research question. It would seem that appropriate semantic restrictions on the nature of Δ functions might preclude the existence of PN and NP functions.

In the case of ρ (local) persistence functions, we can categorize functions into P and N categories with similar meaning (i.e., whether the longer I, the longer or shorter the validity interval before or after I).

The dynamic knowledge about the domain does not necessarily need to include complete, closed, definitions of Δ functions—partial tables may suffice, or the actual functions might be approximated. But knowing whether a maximal-gap function is positive (PP) or negative (NN) is important for estimating the value of that function from a few

examples or for interpolating that value from several discrete entries in a table. This qualitative-persistence type is easy to acquire, since domain experts usually have an excellent intuition about whether, qualitatively, a longer duration of a parameter proposition before or after a gap increases or decreases the probability of the proposition being true during a longer gap, even if the probabilities involved are in fact unknown.

6 Related Work

Several temporal logics include some form of a persistence axiom for facts, that states that a proposition stays true until known to be otherwise. The ρ local–persistence function can be viewed as an extension of McDermott's persistence assumption [McDermott, 1982; Dean and McDermott, 1987] and of McCarthy's inertia principle [McCarthy, 1986]. Both, however, include infinite persistence as a default. McDermott [1982] suggested that a fact does not cease to be true unless we explicitly hear that it no longer is true. Since this assumption is not always realistic, McDermott introduced the idea of a typical *lifetime* of a fact. Thus, an event causes *persistence* of a fact. Our ρ function belief threshold creates a value- and context-specific validity time for a parameter proposition, but ρ functions extend temporally in *both* directions.

Tawfik and Neufeld [1996] have computed the relevance of time-stamped knowledge in a temporal Bayesian framework, modeling relevance as a Markov process and looking only at a single predicate and a forward projection. Their analysis can be viewed as providing bounds on relevance due to a local persistence function, with certain independence assumptions.

Dean and Kanazawa [1988] proposed a model of probabilistic temporal reasoning about propositions that *decay* over time. They modeled explicitly the probability of a proposition P being true at time t, P(<P, t>), given the probability of <P, t-Δ>. The assumption is that there are events of type E_p that can cause proposition p to be true, and events of type $E_{\neg p}$ that can cause it to be false. Thus, one can define a *survivor function* for P(<P, t>) given <P, t-Δ>, such as an exponential decay function. Our ρ function model is somewhat similar. However, Dean and Kanazawa's main intention was not to solve an *interpretation* task (such as the TA task) but to solve a *projection* task, in particular in the context of the planning task. Thus, unlike in our model, persistence is only considered forwards in time. In a later work, Kanazawa [1991] presented a logic of time and probability, L_{cp}. Propositions asserted in L_{cp} were stored in a *time network*, which maintained probabilistic dependencies among various facts, such as the time of arrival of a person at a place, or the range of time over which it is true that the person stayed in one place, and was used to answer queries about probabilities of facts and events over time.

In medical domains, two approaches tat are somewhat similar to the one used by Dean and Kanazawa are de Zegher-Geets' **time-oriented probabilistic functions** (**TOPFs**) in the **IDEFIX** system [de Zegher-Geets *et al.*, 1988] for summarization of medical records, and Blum's [1982] time-dependent database access functions and proxy variables to handle missing data in the context of the **Rx** project for automated discovery in clinical databases. The goals of these systems were also closer in nature to the TA task—that is, *interpretation* of time-stamped data. When de Zegher-Geets' TOPFs represent the probability of a state or disease given a previous identical state, they simulate a forward ρ function; in addition, states in IDEFIX can have an *expected length* attribute.

Russ [1995] has analyzed the computational cost of limited temporal persistence, considering medical domains in particular, and has shown the improvements enabled by data abstraction. Since the KBTA method operates at multiple levels of abstraction, it often capitalizes automatically on such improvements.

7 Implementation and an Example From the Diabetes-Monitoring Domain

The KBTA method had been implemented by the **RÉSUMÉ** system [Shahar and Musen, 1993] and was evaluated in various areas of clinical medicine [Shahar and Musen, 1996], with highly encouraging results. The results emphasized not only the validity of the methodology, but the advantages of explicit representation of temporal-abstraction knowledge for acquiring, maintaining, and reusing that knowledge. A graphical tool for acquiring temporal-abstraction knowledge from expert physicians was constructed [Stein *et al.*, 1996], using the **PROTÉGÉ-II** framework's set of tools [Tu *et al.*, 1995]. The RÉSUMÉ system is currently integrated within the **EON** component-based architecture for guidline-based care [Musen *et al.*, 1996].

An example of using the RÉSUMÉ system in a medical domain is an evaluation that we performed in the domain of monitoring patients who have insulin-dependent diabetes [Shahar and Musen, 1996]. We collaborated with two endocrinologists, acquiring within several meetings a TA ontology from one of the experts. We created a parameter-properties ontology for the domain of insulin-dependent diabetes (Figure 3), an event ontology (e.g., insulin therapy, meals, physical exercise), and an interpretation-context ontology (e.g., preprandial [measured at fasting time, before a meal] and postprandial [after a meal] contexts and subcontexts, and postexercise contexts).

Glucose_state_state values (i.e., values of the state(state(glucose)) abstract parameter) that are measured within *different phases* (e.g., prelunch and presupper), but within the *same day*, can be joined by interpolation within the *nonconvex context* [Shahar, 1996] version of the PREPRANDIAL *generalized interpretation context* [Shahar, 1996], thus creating an abstraction comprising several

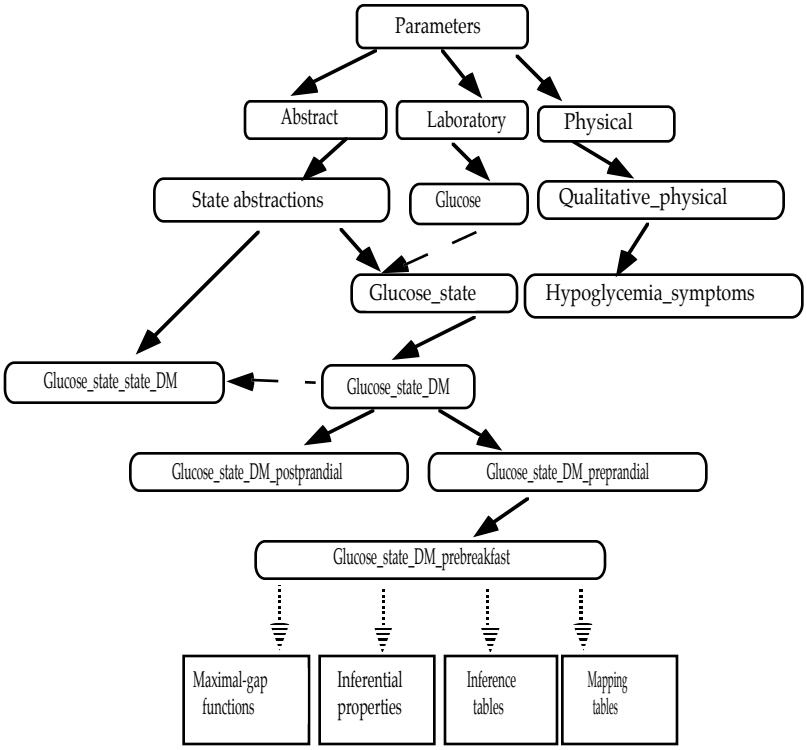

Figure 3: Part of the diabetes parameter-properties ontology. The Glucose parameter is abstracted into the Glucose_state parameter. This abstract parameter has a specialized subclass in the DM context, and is abstracted in that context into the Glucose_state_state parameter. The Glucose_state_DM class is further specialized in the preprandial and postprandial contexts, each of which has several subclasses corresponding to the different relevant premeal contexts. = class; = property; = IS-A relation; = ABSTRACTED-INTO relation; = PROPERTY-OF relation; DM = diabetes mellitus.

preprandial abstractions within the same day, up to 6 to 8 hours apart. The maximal gap is defined by a interphase Δ function. Diurnal state abstractions that are measured in the *same phase* but over *different* (usually consecutive) *days*, such as several values of the Glucuse_state_DM_prebreakfast parameter, can be joined by interpolation within the same interpretation context (e.g., a nonconvex PREBREAKFAST context interval, that comprises all breakfasts within a given interval), up 24 to 28 hours apart, using another interphase Δ function.

In the study, the two experts formed (independently) temporal abstractions from more than 800 points of data, representing two weeks of glucose and insulin data from each of eight patients. The RÉSUMÉ system created 132 (80.4%) of the 164 temporal abstractions noted by both experts [Shahar and Musen, 1996]. An example of the output is shown in figure 4. Examination of the output for the first three cases by one of the experts showed that the expert agreed with almost all (97%) of the produced abstractions—a result similar to the one we found in a previous study in the domain of growth monitoring. We expected this high predictive value, since the domain's TA ontology directly reflected that expert's knowledge about these low- and intermediate-level abstractions.

8 Discussion and Conclusions

The knowledge requirements for the temporal-interpolation mechanism include (1) *structural knowledge*: the qualitative-dependency aspect of the ABSTRACTED-INTO relation; domain time units; (2) *classification knowledge*: classification of domain-specific gradient and, in particular, rate abstraction values (e.g., SLOW, FAST) as changes per time unit; horizontal-classification knowledge, that is, the horizontal-inference table; (3) *temporal-dynamic knowledge*: maximal-gap (Δ) functions and local (ρ) persistence functions, both specific to each parameter proposition (which includes an explicit context); significant change values C_π or functions $f_c(\pi)$ for the relevant parameters in various contexts; additional temporal constraints for completing the interpolation-inference table; and (4) *temporal-semantic knowledge*: truth values for the *concatenable* property [Shoham, 1987] for input and inferred parameters.

Temporal-dynamic knowledge about a domain does not necessarily need to include complete definitions of Δ functions—partial functions may suffice, and knowing whether a maximal-gap function is PP or NN might complete the picture. The qualitative type of a persistence

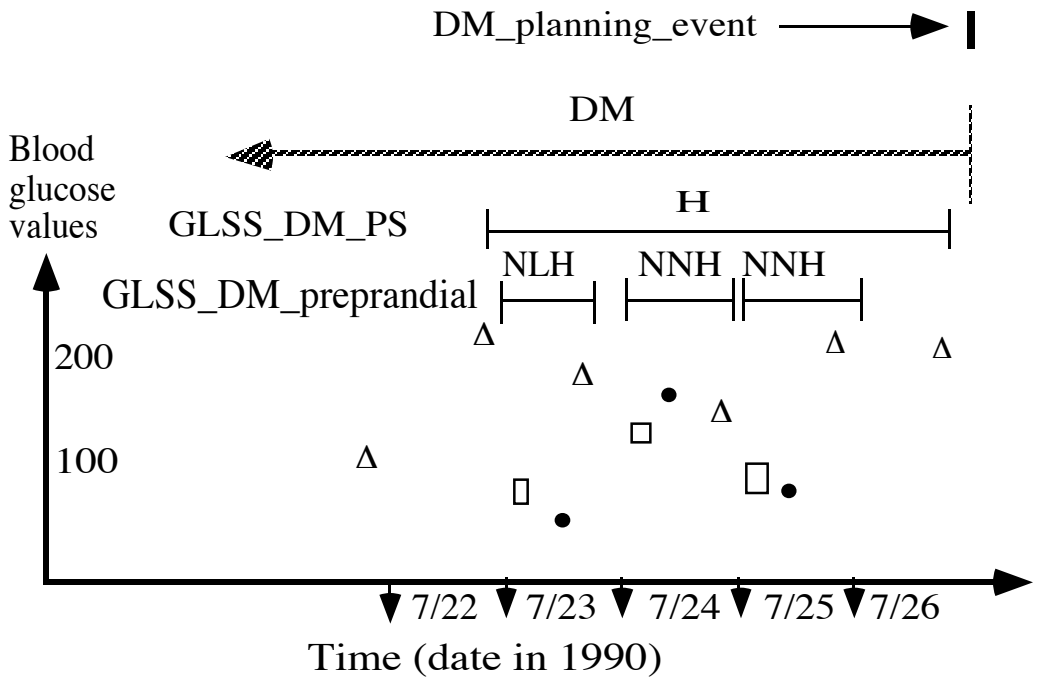

Figure 4: Abstraction of data by the RÉSUMÉ system in the domain of monitoring diabetes patients. ⊢→ = (open) context interval; ⊢⊣ = abstraction interval; ❒ = prebreakfast glucose; • = prelunch glucose; Δ = presupper glucose; DM = diabetes mellitus therapy context; GLSS_DM_PS = Glucose_state_state abstraction in the DM and presupper context; GLSS_DM_PREPRANDIAL = Glucose_state_state abstraction in the DM and preprandial context.

function is easy to acquire from expert physicians in any particular clinical area.

Furthermore, one of the insights underlying our model is that higher-level abstractions are often more persistent. Since temporal interpolation operates simultaneously at all abstraction levels, the more stable abstract conclusions often mask faster changes (and uncertainties) in lower-level abstractions and raw data.

The bidirectional temporal persistence model we present is relevant when data is abstracted and interpreted retrospectively, as is the goal of the TA task. Furthermore, both ρ and Δ functions are context sensitive and are thus represented explicitly. Finally, as shown in Sections 5.2, the use of global (Δ) persistence functions facilitates acquisition of temporal-dynamic knowledge.

The current knowledge-based temporal-interpolation model has three major limitations. From the *soundness* aspect, the threshold cutoff assumed by the model is convenient in practice, but might potentially lead to unsound conclusions (from the clinical domain's point of view) of higher-level abstractions that use the result of the interpolation (which is assumed to hold with certainty once its probability is higher than a domain-specific threshold). Thus, a confidence value should still be attached to the conclusion. From the *completeness* point of view, the model cannot conclude values of the parameter during the gap in the specific case when the values before and after the gap are different and also are not part of a horizontal-join relation. Finally, from the *knowledge acquisition* point of view, even when using the results of the analysis in Section 5.2, considerable amounts of knowledge might still need to be acquired from expert physicians. (Currently, we are using a graphic knowledge-acquisition tool that uses three-dimensional tables to represent Δ functions [Stein et al., 1996] and that is generated automatically, given the KBTA ontology, by tools from the PROTÉGÉ-II project [Tu et al., 1995].)

Thus, our future plans are to (1) construct a Bayesian-semantics framework for the interpolation operation, (2) attempt to learn local and global interpolation functions from large temporal databases (given some domain knowledge, such as the abstraction hierarchy and classification functions, and the temporal-semantic properties of relevant parameters), and (3) test the automatically acquired functions using methodologies that have been proven valuable in similar cases in clinical domains. One example we are considering is the Stanford-based ARAMIS project, in which records of patients who have rheumatoid arthritis and related chronic diseases have been collected for more than 30 years. In one experiment in that project, various interpolation functions for missing raw data have been tested and compared by attempting to "guess"

values of data that were temporarily made invisible, thus simulating the case of missing data [Albridge *et al.*, 1984]. Such methodologies might be applicable also for higher-level abstractions, once the data had been abstracted partially (e.g., only vertically).

Acknowledgments

This work has been supported by grants LM05708 and LM06245 from the National Library of Medicine and IRI-9528444 from the National Science Foundation. Computing resources were provided by the CAMIS project, funded under grant No. LM05305 from the National Library of Medicine.

References

[Albridge *et al.*, 1984] K.M. Albridge, J. Standish, and J.F. Fries. Hierarchical time-oriented approaches to missing data inference. *Computers and Biomedical Research*, 21: 349–-366, 1984.

[Allen, 1984] J.F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–-154, 1984.

[Blum, 1982] R.L. Blum. Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project. In D.A. Lindberg and P.L. Reichartz, eds., *Lecture Notes in Medical Informatics*, vol. 19 (Springer-Verlag, New York, 1982).

[De Zegher-Geets *et al.*, 1988] I.M. De Zegher-Geets, A.G. Freeman, M.G. Walker, R.L. Blum, and Wiederhold,G. Summarization and display of on-line medical records. *M.D. Computing*, 5:38–-46, 1988.

[Dean and Kanazawa, 1986] T. Dean and K. Kanazawa. Probabilistic temporal reasoning. In *Proceedings of the Eight National Conference on Artificial Intelligence*, pages 524--528, 1986. Minneapolis, MN.

[Eriksson *et al.*, 1995] H. Eriksson, Y. Shahar, S.W. Tu, A.R. Puerta, and M.A. Musen. Task modeling with reusable problem-solving methods. *Artificial Intelligence*, 79(2): 293--326, 1995.

[Forbus, 1984] K.D. Forbus. Qualitative process theory. *Artificial Intelligence*, 24(1–3):85--168, 1984.

[Kanazawa , 1991] K. Kanazawa. A logic and time nets for probabilistic inference. In *Proceedings, Ninth National Conference on Artificial Intelligence*, pages 360--365, Los Angeles, CA. MIT Press, Cambridge, MA, 1991.

[Ladkin, 1986] P. Ladkin. Time representation: A taxonomy of interval relations. In *Proceedings of the Sixth National Conference on Artificial Intelligence*, pages 360-366, Philadelphia, PA, 1986.

[McCarthy and Hayes, 1969] J. McCarthy and P. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence*, University Press, Edinburgh, UK, 1969.

[McCarthy, 1986] J. McCarthy. Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, 28(1):89--116, 1986.

[McDermott, 1982] D.V. McDermott. A temporal logic for reasoning about processes and plans. *Cognitive Science*, 6(2):101–-155, 1982.

[Musen *et al.*, 1996] M.A. Musen, S.W. Tu, A.K. Das, and Y. Shahar. EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Association*, 3(6):367–-388, 1996.

[Russ, 1989] T.A. Russ. Using hindsight in medical decision making. In L. C. Kingsland, ed., *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*, pages 38--44. IEEE Computing Society Press, Washington, 1989.

[Russ, 1995] T.A. Russ. Use of data abstraction methods to simplify monitoring. *Artificial Intelligence in Medicine*, 7 (6):497-–514, 1995.

[Shahar and Molina, 1996] Y. Shahar and M. Molina. Knowledge-based spatiotemporal abstraction. In *Proceedings of the AAAI-96 Workshop on Spatial and Temporal reasoning*, pages 21--30, Portland, Oregon, 1996.

[Shahar *et al.*, 1993] Y. Shahar, S.W. Tu, and M.A. Musen. Knowledge acquisition for temporal-abstraction mechanisms. *Knowledge Acquisition*, 4:217–-236, 1993.

[Shahar and Musen, 1993] Y. Shahar and M.A. Musen. RÉSUMÉ: A temporal-abstraction system for patient monitoring. *Computers and Biomedical Research*, 26: 255-–273, 1993. Reprinted in van Bemmel, J.H., and McRay, A.T. (eds), *Yearbook of Medical Informatics 1994*,

pages 443-–461. Stuttgart: F.K. Schattauer and The International Medical Informatics Association, 1994.

[Shahar, 1996] Y. Shahar. Dynamic temporal interpretation contexts for temporal abstraction. In *Proceedings of the 1996 Third International Workshop on Temporal Representation and Reasoning*, pages 64--71, Key West, Florida, 1996.

[Shahar and Musen, 1996] Y. Shahar and M.A. Musen. Knowledge-based temporal abstraction in clinical domains. *Artificial Intelligence in* Medicine, 8(3):267--298, 1996.

[Shahar, 1997] Y. Shahar. A framework for knowledge-based temporal abstraction. *Artificial Intelligence*, 90(1–2):79-–133, 1997.

[Shoham, 1987] Y. Shoham. Temporal logics in AI: Semantical and ontological considerations. *Artificial Intelligence*, 33(1):89–104, 1987.

[Stein *et al*., 1996] A. Stein, M.A. Musen, and Y. Shahar. Knowledge acquisition for temporal abstraction. In *Proceedings of the 1996 AMIA Annual Fall Symposium (formerly the Symposium on Computer Applications in Medical Care)*, pages 204--208, Washington D.C., Hanley and Belfus, Philadelphia, PA, 1996.

[Tawfik and Neufeld, 1996] A. Y. Tawfik and E. M. Neufeld. Irrelevance in uncertain temporal reasoning. In *Proceedings of the Third International Workshop on Temporal Representation and Reasoning (TIME '96)*, pages 196--202, Key West, Florida, 1996.

[Tu *et al*., 1995] S.W. Tu, H. Eriksson, J. Gennari, Y. Shahar, and M.A. Musen. Ontology-based configuration of problem-solving methods and generation of knowledge-acquisition tools: Application of PROTÉGÉ-II to protocol-based decision support. *Artificial Intelligence in Medicine*, 7(3):257–289, 1995.