Hybrid Time Bayesian Networks

Manxia Liu[1(✉)], Arjen Hommersom[1,2], Maarten van der Heijden[1], and Peter J.F. Lucas[1]

[1] ICIS, Radboud University, Nijmegen, The Netherlands
{m.liu,arjenh,m.vanderheijden,peterl}@cs.ru.nl
[2] Open University, Heerlen, The Netherlands

Abstract. Capturing heterogeneous dynamic systems in a probabilistic model is a challenging problem. A single time granularity, such as employed by dynamic Bayesian networks, provides insufficient flexibility to capture the dynamics of many real-world processes. The alternative is to assume that time is continuous, giving rise to continuous time Bayesian networks. Here the problem is that the level of temporal detail is too precise to match available probabilistic knowledge. In this paper, we present a novel class of models, called hybrid time Bayesian networks, which combine discrete-time and continuous-time Bayesian networks. The new formalism allows us to more naturally model dynamic systems with regular and irregularly changing variables. Its usefulness is illustrated by means of a real-world medical problem.

Keywords: Continuous time Bayesian networks · Dynamic Bayesian networks · Dynamic systems

1 Introduction

Many real-world systems exhibit complex and rich dynamic behavior. As a consequence, capturing these dynamics is an integral part of developing models of physical-world systems. Time granularity is an important parameter in characterizing dynamics as it determines the level of temporal detail in the model. In cases where one time granularity is coarser than another, dealing with multiple time granularities becomes significantly important, e.g., in the context of mining frequent patterns and temporal relationship in data stream and databases [1].

Dynamic Bayesian networks (DBNs) are a general framework for modeling dynamic probabilistic systems. DBNs are an extension of standard Bayesian networks (BNs) assuming a discretization of time [2], and where the distribution of variables at a particular time point is conditional on the state of the system at the previous time point. A problem occurs if temporal processes of a system are best described using different rates of change, e.g., one temporal part of the process changes much faster than another. In that case, the whole system

ML is supported by China Scholarship Council. AH and MVDH are supported by the ITEA2 MoSHCA project (ITEA2-ip11027).

S. Destercke and T. Denoeux (Eds.): ECSQARU 2015, LNAI 9161, pp. 376–386, 2015.
DOI: 10.1007/978-3-319-20807-7_34

has to be represented using the finest time granularity, which is undesirable from a modeling and learning perspective. In particular, if a variable is observed irregularly, much data on discrete-time points will be missing and conditional probabilities will be hard to estimate.

As an alternative to DBNs, temporal processes can be modeled as continuous time Bayesian networks (CTBNs), where time acts as a continuous parameter [3]. In these models, the time granularity is infinitely small by modeling transition rates rather than conditional probabilities, thus multiple time granularities, i.e., slow and fast transition rates, can easily be captured. A limitation from a modeling perspective is that all probabilistic knowledge, for example derived from expert knowledge, has to be mapped to transition rates which are hard to interpret. Moreover, the transition rates assume that the time until a transition is exponentially distributed, which may not always be appropriate.

In this paper, we propose a new formalism, which we call hybrid time Bayesian networks (HTBNs), inspired by discrete-time and continuous-time Bayesian networks. They facilitate modeling the dynamics of both irregularly-timed random variables and random variables which are naturally described in a discrete way. As a result, the new formalism increases the modeling and analysis capabilities for dynamic systems.

2 Motivating Example

To illustrate the usefulness of the proposed theory, we consider the medical problem of heart failure and, in particular, one possible cause of heart failure: heart attack (myocardial infarction). This usually occurs as the result of coronary artery disease giving rise to reduced blood supply to the heart muscle (myocardium). One consequence is that part of the heart muscle will die, which is revealed later in a blood sample analysis in the lab by an increased level of particular heart muscle proteins, in particular troponine. Loss of heart muscle will inevitably have an impact on the contractability of the myocardium, and thus heart function will be negatively affected. This is known as *heart failure*. In particular, the heart fails with respect to its function as a pump. This will enforce an increase in the amount of extracellular fluid (the patient is flooded with water), which can be measured quite simply by means of the body weight. With regard to treatment, digitalis is considered as one of the drugs to improve contractability. This causal knowledge is formalized as a directed graph in Fig. 1.

Heart attacks usually happen repeatedly in patients, although after some interval of time, and this may negatively affect heart function. After administration of digitalis it will take some time, in terms of days, before the drug has a diminishing effect on heart failure. Thus, the course of heart failure will likely depend on various factors, and how they interact. Of particular importance here is the dynamic over time of the probability distributions.

In modeling processes such as heart failure, it is essential to notice the existence of different time granularities. There are *discrete, regular* variables which are observed regularly such as a routine checkup for body weight and a regular

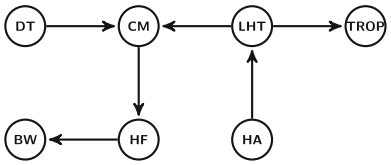


Fig. 1. Causal model for heart failure: CM = Contractility Myocardium, DT = Digitalis, LHT = Loss Heart Tissues, HA = Heart Attack, TROP = Troponine, HF = Heart Failure, BW = Body Weight.

intake of a drug. On the other hand, some variables are observed *irregularly*, such as the indicator troponine which is elevated after about half an hour after damage to the heart muscle is obtained; however its measurement is repeated with time intervals that increase after the patient's condition has been stabilized. Clearly, it is not possible to obtain a satisfactory representation of the clinical evolution of heart failure using only discrete time, regular or irregular, or continuous time. In the remainder of this paper we propose a method to deal with these heterogeneous time aspects.

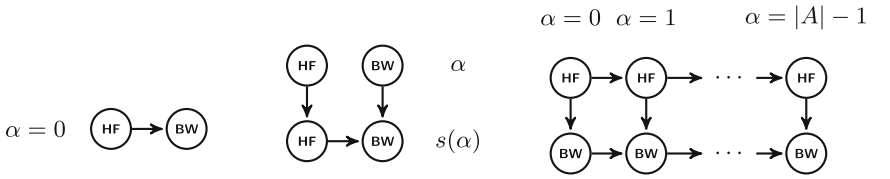
3 Preliminaries

We start by introducing Bayesian networks, dynamic Bayesian networks and continuous time Bayesian networks. In the following, upper-case letters, e.g. X, Y, or upper-case strings, e.g. HA, denote random variables. We denote the values of a variable by lower-case letters, e.g. x. We will also make use of a successor function s, which is defined on a countable, linearly ordered set of numbers Z in which every element $z_i \in Z$ with index i is mapped to element $s(z_i) = z_{i+1} \in Z$.

Bayesian Networks. A Bayesian network is a probabilistic graphical model which represents a joint probability distribution of a set of random variables. A *Bayesian network* $\mathcal{B}$ is defined as a pair $\mathcal{B} = (G, P)$, where G is an acyclic directed graph with $G = (V(G), E(G))$, where $V(G)$ is a set of nodes, and $E(G) \subseteq V(G) \times V(G)$ a set of directed edges or arcs. A joint probability distribution P is defined by a set of conditional probabilities of each random variable X given its immediate parents $\pi(X)$ in G, formally: $P(V(G)) = \prod_{X \in V(G)} P(X \mid \pi(X))$.

Dynamic Bayesian Networks (DBNs). A DBN is defined as a pair $(\mathcal{B}_0, \mathcal{B}_{\rightarrow})$ over discrete-time variables $\mathbf{D}$, where $\mathcal{B}_0$ is taken as the initial Bayesian network model and $\mathcal{B}_{\rightarrow}$ is defined as a conditional distribution for a 2-time-slice Bayesian network (2-TBN). Given a set of discrete time points of interest $A \subseteq \mathbb{N}_0$ that includes 0, the joint distribution for a DBN with $|A|$ slices is defined by a product of the CPDs in the initial model and in the 2-TBN:

$$P(\mathbf{D}_A) = \prod_{D \in \mathbf{D}} P_{\mathcal{B}_0}(D_0 \mid \pi(D_0)) \prod_{D \in \mathbf{D}} \prod_{\alpha \in A \setminus \{\max A\}} P_{\mathcal{B}_{\rightarrow}}(D_{s(\alpha)} \mid \pi(D_{s(\alpha)}))$$



(a) Initial Model (b) Transition Model (c) A standard Bayesian network

Fig. 2. A DBN and its corresponding Bayesian network.

where $D_{s(\alpha)}$ is the random variable D at time $s(\alpha)$. Parent set $\pi(D_{s(\alpha)})$ may be from the same or the previous time slice. We can obtain a standard Bayesian network by unrolling the DBN over the time points of interest. In the remainder it is assumed that the intra-slice arcs of this BN are the same for every α.

Example 1. Consider a dynamic Bayesian network that has two random variables, HF and BW (see Fig. 1), with an initial model and a transition model as shown in Figs. 2a and 2b, respectively. Then the joint distribution for the DBN over time points of interest A with the corresponding Bayesian network as shown in Fig. 2c is: $P(\mathrm{HF}_A, \mathrm{BW}_A) = P(\mathrm{HF}_0)P(\mathrm{BW}_0 \mid \mathrm{HF}_0) \prod_{\alpha=0}^{|A|-2} P(\mathrm{BW}_{s(\alpha)} \mid \mathrm{BW}_\alpha, \mathrm{HF}_{s(\alpha)})P(\mathrm{HF}_{s(\alpha)} \mid \mathrm{HF}_\alpha)$.

Continuous Time Bayesian Networks (CTBNs). CTBNs [3] represent dynamic systems with continuous-time variables as a factorized homogeneous Markov process parameterized by *intensity matrices*. An entry (i, j) with $i \neq j$ in an intensity matrix gives the intensity of transitioning from state i to state j. Furthermore, the main diagonal makes each row sum to zero.

Example 2. Suppose we want to model the random process of body weight as the variable BW, which describes a patient's weight. Variable BW has three possible states, i.e., BW = {low, normal, high}, with a transition matrix as follows:

$$Q_{\mathrm{BW}} = \begin{pmatrix} -0.13 & 0.09 & 0.04 \\ 0.13 & -0.23 & 0.1 \\ 0.07 & 0.16 & -0.23 \end{pmatrix}$$

For example, the entry $(3, 2)$ means that the process will transition from high at time β to normal at time $\beta + \epsilon$ with a probability of 0.16/0.23=0.696 if a transition occurs at $\beta + \epsilon$.

The notion of a *conditional intensity matrix (CIM)* describes the dependence of a variable C on the current values of its parents $\pi(C)$. A *full amalgamation* product operator is defined over a set of CIMs to compute the joint intensity matrix, resulting in a single continuous-time Markov process for the entire system.

For a homogeneous Markov process over variables $\mathbf{C}$ with an intensity matrix $Q_{\mathbf{C}}$ and an initial distribution $P(\mathbf{C}_0)$, we can compute the distribution over the values of $\mathbf{C}$ at a particular time point or the joint distribution at different time points. The distribution at a point β is given by:

$$P(\mathbf{C}_\beta) = P(\mathbf{C}_0)\exp(Q_{\mathbf{C}}\beta)$$

The distribution over a finite set of time points of interest B is given by:

$$P(\mathbf{C}_B) = P(\mathbf{C}_0) \prod_{\beta \in B \setminus \{\max B\}} \exp(Q_{\mathbf{C}}(s(\beta) - \beta))$$

4 Hybrid Time Bayesian Networks

In this section, we define hybrid-time Bayesian networks, the semantics of these models in terms of its factorization, and finally, we show how such models can be interpreted as regular Bayesian networks. The latter is particularly important for practical purposes, as this implies that we may (dynamically) generate discrete-time versions of the model given time points for which we have observations, and in which we would like to compute marginals. After that, we can use existing methods for probabilistic inference in BNs.

4.1 Model Definition

The formal definition of hybrid time Bayesian networks is as follows.

Definition 1 (Hybrid Time Bayesian Networks (HTBNs)). *A hybrid time Bayesian network is a triple $\mathcal{H} = (G, \Phi, \Lambda)$, where $G = (V(G), E^t(G), E^a(G))$ is a directed graph with each vertex in $V(G)$ either a continuous-time variable, collectively denoted by $\mathbf{C}$, or a discrete-time variable, collectively denoted by $\mathbf{D}$, $E^t(G)$ and $E^a(G)$ are temporal and atemporal arcs, respectively, such that $(V(G), E^a(G))$ is acyclic, Φ is a set of conditional probability distributions for variables $\mathbf{D}$, and Λ is a set of conditional intensity matrices and initial distributions for variables $\mathbf{C}$.*

Furthermore, graph G has the following properties:

(i) Arcs connecting continuous-time and discrete-time variables are atemporal;
(ii) Arcs connecting continuous-time variables are temporal;
(iii) A continuous-time variable has a temporal arc to itself.

Property (*iii*) indicates that a discrete-time variable does not necessarily have temporal dependences on itself. It is worthwhile to notice that the temporal cyclic property is inherited from discrete-time and continuous-time Bayesian networks. A temporal cycle is possible in two cases, either between continuous-time variables or between discrete-time variables. However, an atemporal cycle is not allowed, that is, there is no cycle in the graph involving both continuous-time and discrete-time variables.

Example 3. In the example discussed in Sect. 2, regular variables, i.e., BW, DT, HF and hidden variable CM can be represented in a discrete-time manner. The irregular variables, i.e., LHT, TROP, HA are modeled as continuous-time variables. The example is then represented in a hybrid time Bayesian network $\mathcal{H}$ as shown in Fig. 3.

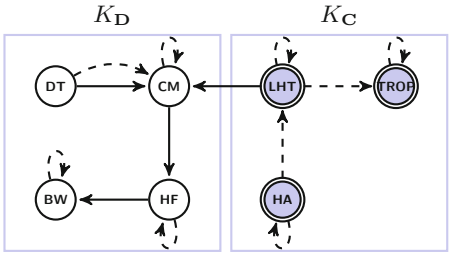


Fig. 3. An HTBN for the heart failure problem. Continuous-time variables are graphically represented by double-edged blue circles, atemporal arcs are solid, and temporal arcs are dashed.

4.2 Factorization

The joint probability distribution for hybrid time Bayesian networks is defined by multiplying the conditional joint probabilities for continuous-time and discrete-time Bayesian networks. To this end, we first need to introduce some new notions.

The *skeleton* $G^{\sim}$ of a directed graph G is obtained by changing the arcs in G by (undirected) edges. Every directed graph can be defined as the union of *connected components* by an equivalence relation $X - Y$, meaning that vertex Y can be reached by an undirected path from vertex X in its skeleton. Vertex X and Y are then members of the same equivalence class $[X]$ and the corresponding graph is a connected component. A graph $G' = (V(G'), E(G'))$ is said to be an *induced subgraph* of G if $E(G') = (V(G') \times V(G')) \cap E^t(G)$ and $V(G') = \mathbf{C}$, called a *continuous-time induced subgraph*, denoted as $G_{\mathbf{C}}$, or $E(G') = (V(G') \times V(G')) \cap (E^t(G) \cup E^a(G))$ and $V(G') = \mathbf{D}$, when it is called a *discrete-time induced subgraph*, denoted as $G_{\mathbf{D}}$.

Both $G_{\mathbf{C}}$ and $G_{\mathbf{D}}$ can be decomposed into connected components; each individual connected component is indicated by $K_{\mathbf{C}}$ and $K_{\mathbf{D}}$, respectively. Clearly connected components are disjoint as they represent equivalence classes and together the connected components form partitions of the continuous-time and discrete-time subgraphs, respectively. A subset $\mathbf{X} \subseteq V(G_{\mathbf{D}})$ is said to constitute the parents of $V(K_{\mathbf{C}})$, denoted as $\pi(V(K_{\mathbf{C}}))$, if and only if there exists an arc (D, C) in G, $C \in V(K_{\mathbf{C}})$, for every $D \in \mathbf{X}$. Parents $\pi(V(K_{\mathbf{D}}))$ are defined analogously. In the example shown in Fig. 3, there is only one continuous-time connected component with $V(K_{\mathbf{C}}) = \{\text{LHT}, \text{TROP}, \text{HA}\}$ and one discrete-time connected component with $V(K_{\mathbf{D}}) = \{\text{DT}, \text{CM}, \text{HF}, \text{BW}\}$.

We are now in the position to define a conditional distribution of connected components given their parents.

Definition 2 (Conditional Joint Distribution for Component $K_{\mathbf{D}}$). *Given a discrete-time component $K_{\mathbf{D}}$, the conditional joint distribution for $K_{\mathbf{D}}$ over time points of interest A is defined as:*

$$P(V(K_{\mathbf{D}})_A \mid \pi(V(K_{\mathbf{D}}))_A) = \prod_{D \in V(K_{\mathbf{D}})} (P(D_0 \mid \pi^a(D)_0) \prod_{\alpha \in A \setminus \{0\}} P(D_\alpha \mid \pi^a(D)_\alpha, \pi^t(D)_{\alpha-1}))$$

where $\pi^t(D)$ are the temporal and $\pi^a(D)$ are the atemporal parents of D.

Definition 3 (Conditional Joint Distribution for Component $K_{\mathbf{C}}$). *Given a continuous-time component $K_{\mathbf{C}}$ over variables $V(K_{\mathbf{C}})$ with an initial distribution $P(V(K_{\mathbf{C}})_0)$ and corresponding parents $\pi(V(K_{\mathbf{C}}))$ over time points of interest A. The conditional joint distribution for $K_{\mathbf{C}}$ over a finite set of time points of interest B, $\{0\} \subset A \subseteq B \subset \mathbb{R}^+$, is defined as:*

$$P(V(K_{\mathbf{C}})_B \mid \pi(V(K_{\mathbf{C}}))_A) = P(V(K_{\mathbf{C}})_0) \prod_{\beta \in B \setminus \{\max B\}} \exp(Q_{V(K_{\mathbf{C}})|\pi(V(K_{\mathbf{C}}))_a}(s(\beta) - \beta))$$

$$a = \max\{\alpha \mid \alpha \leq \beta, \alpha \in A\}$$

where $Q_{V(K_{\mathbf{C}})|\pi(V(K_{\mathbf{C}}))_a}$ is the conditional intensity matrix for variables $V(K_{\mathbf{C}})$ given the values of parents $\pi(V(K_{\mathbf{C}}))$ at time a.

Now we can define the full joint probability distribution of a hybrid-time BN given sets of time points of interest.

Definition 4 (Joint Probability Distribution). *Given a hybrid time Bayesian network $\mathcal{H}$ and sets of components $\boldsymbol{K}_{\mathbf{D}}$, $\boldsymbol{K}_{\mathbf{C}}$ with associated time points of interest A, B. The joint distribution for $\mathcal{H}$ over B is defined as:*

$$P(V(G)_B) = \prod_{K_{\mathbf{C}} \in \boldsymbol{K}_{\mathbf{C}}} P(V(K_{\mathbf{C}})_B \mid \pi(V(K_{\mathbf{C}}))_A) \prod_{K_{\mathbf{D}} \in \boldsymbol{K}_{\mathbf{D}}} P(V(K_{\mathbf{D}})_A \mid \pi(V(K_{\mathbf{D}}))_A)$$

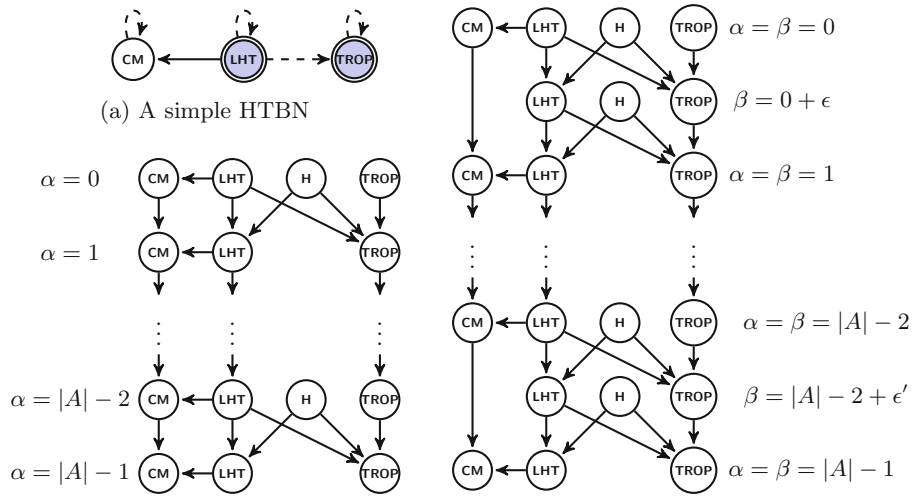
The following propositions establish that HTBNs are proper generalizations of both DBNs and CTBNs.

Proposition 1. *A DBN $(\mathcal{B}_0, \mathcal{B}_\rightarrow)$ with random variables $\mathbf{D}$, and an HTBN $(G, \Phi, \varnothing)$ define the same joint probability distribution for any set of time points of interest A, if $V(G) = \mathbf{D}$; $E^a(G)$, $E^t(G)$ correspond to the temporal and atemporal arcs of $\mathcal{B}_\rightarrow$, and Φ are the parameters of the DBN.*

Proposition 2. *A CTBN with graph G and parameters Λ and an HTBN $(G, \varnothing, \Lambda)$ define the same probability distribution for any set of time points of interest B.*

4.3 Discrete-Time Characterization

A natural question is whether the joint distribution defined on a HTBN, given the fixed time points of interest, can also be graphically represented as a regular (discrete-time) Bayesian network. The benefit is that the parameters of the resulting Bayesian network are conditional probabilities, which are easier to understand for domain experts. Furthermore, this construction is convenient as it enables the use of standard software for inference in HTBNs.



(a) A simple HTBN

(b) Same time points of interest $A = B$

(c) Different time points of interest: $A \neq B$

Fig. 4. Discretization of an HTBN.

Below we show that there is a construction to a regular Bayesian network possibly at the expense of introducing additional hidden variables that model the dependence structure of continuous-time variables. The reason for these hidden variables is as follows. Consider a simple structure $X \rightarrow Y \rightarrow Z$. In a regular Bayesian network, it holds that Z is independent of X given its parent Y. Interpreting this graph as a continuous-time component (where arcs are temporal), a continuous-time variable is conditionally independent of its non-descendants given the full trajectories of its parents. In the structure given, we thus can only conclude that Z at time β is independent of X given the full trajectory for Y from time 0 to time β, otherwise X and Z are dependent. In order to represent this, we introduce additional dependences between X and Z at each time point of interest using auxiliary hidden variables. We illustrate the process in Fig. 4.

Proposition 3 (Discretization). *Given a hybrid time Bayesian network $\mathcal{H}$ described by a graph G with associated probability distribution P and time points of interest, there exists a Bayesian network $\mathcal{B} = (G_{\mathcal{B}}, P_{\mathcal{B}})$, $P_{\mathcal{B}}(V(G)) = P(V(G))$, which represents all independences of $\mathcal{H}$.*

Proof (Sketch). We only show the construction of this Bayesian network $\mathcal{B}$.

Let $G_{\mathcal{B}} = (V(G_{\mathcal{B}}), E(G_{\mathcal{B}}))$. Set $V(G_{\mathcal{B}})$ are variables mapped from variables V. Set $V(G_{\mathcal{B}})$ is composed of three parts, i.e., $V(G_{\mathcal{B}}) = \Delta \cup \Omega \cup \Theta$, where: 1) Δ are variables $\mathbf{D}$ induced by time points A, 2) Ω are variables $\mathbf{C}$ induced by time points B, 3) Θ are hidden variables induced by temporal dependence between continuous-time variables and time points of interest B, $\Theta = \{H_{\beta}^{ij} \mid (C^i, C^j) \in E^t, \beta \in B\}$, where H_{β}^{ij} models the dependence between variable C^i and C^j.

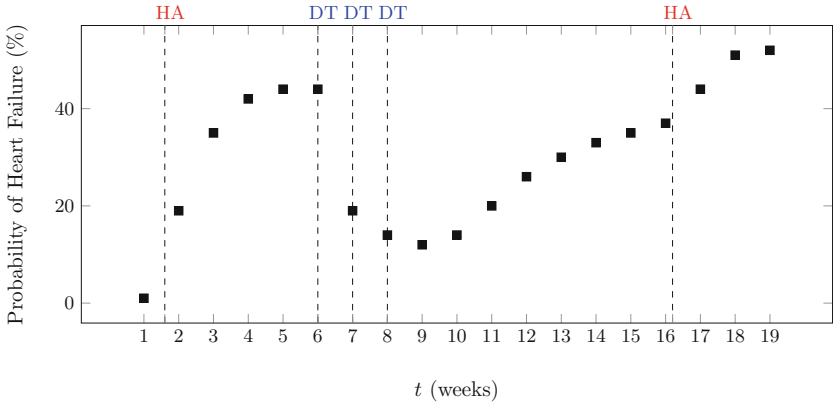


Fig. 5. Effects of heart attack and digitalis on heart failure. 'DT' indicates that digitalis was administered at that moment in time. 'HA' indicates that a heart attack was observed. Note that observations for HA are continuous-time, so observed at an arbitrary point in time; digitalis is observed once a week at most.

Set $E(G_{\mathcal{B}})$ are arcs mapped from E^a and E^t, $E(G_{\mathcal{B}}) \subseteq V(G_{\mathcal{B}}) \times V(G_{\mathcal{B}})$. Basically, the dependence mapping can be categorized by the type of dependences and variables, denoted as $E(G_{\mathcal{B}}) = \Xi \cup \Pi \cup \Upsilon \cup \Gamma$, where: (1) Ξ models the dependence for discrete-time child while its parents could be continuous-time or discrete-time, specified as $\Xi = \{(X_\alpha, D_\alpha) \mid X \in \mathbf{C} \cup \mathbf{D}, (X, D) \in E^a, \alpha \in A\}$, (2) The atemporal dependence for continuous-time variables conditioned on discrete-time parents is specified as $\Pi = \{(D_a, C_\beta) \mid (D, C) \in E^a, \beta \in B\}$, where $a = \max\{\alpha \mid \alpha < \beta, \alpha \in A\}$, (3) Temporal dependences for variables $\mathbf{C}$ and $\mathbf{D}$ are denoted as Υ: $\Upsilon = \{(C_\beta, C_{s(\beta)}) \mid \beta \in B \setminus \{\max B\}\} \cup \{(D_\alpha, D_{s(\alpha)}) \mid (D, D) \in E^t, \alpha \in A \setminus \{\max A\}\}$, 4) Γ are additional dependences for continuous-time variables, $\Gamma = \{(H^{ij}_\beta, C^i_{s(\beta)}), (H^{ij}_\beta, C^j_{s(\beta)}) \mid H^{ij}_\beta \in \Theta, \beta \in B \setminus \{\max B\}\} \cup \{(C^i_\beta, C^j_{s(\beta)}) \mid \beta \in B \setminus \{\max B\}\}$. Thus we have a graph $G_{\mathcal{B}} = (V(G_{\mathcal{B}}), E(G_{\mathcal{B}}))$.

It can be shown that G and $G_{\mathcal{B}}$ represent the same independences on $V(G)$ on the points of interest. Also, the parameters for $\mathcal{B}$ can be derived from $\mathcal{H}$. □

5 Experiments

The power of HTBNs is illustrated in the domain of myocardial contractability in relationship to heart attack, heart failure and its medical treatment, introduced in Sect. 2 and summarized in Fig. 3. In particular, of interest is the question of how the dynamics of the occurrence of heart failure is affected by heart attacks and the administration of digitalis. As discussed in Sect. 2, a single DBN and CTBN can not provide a satisfactory representation of the evolution of variables with different rates: changes in the occurrence of heart failure happen often, in contrast to the more sparse and irregular occurrence of heart attacks.

We parameterized the model using medical expert knowledge given that discrete-time transitions occur weekly. Then, we computed the probability distribution of heart failure for a period of 19 weeks given the observed (regular or irregular) evidence. Results of this experiment are plotted in Fig. 5. The plot shows the negative effects of a heart attack (see the jumps at time $t = 2$, $t = 3$ and $t = 17$) and the positive effect of digitalis on heart failure (see the rapid fall at time $t = 7$). The model also implies that the condition of the heart stabilizes after administering the drug through an increase in the contractility. However, a damaged heart does not fully recover, not even with the help of digitalis.

6 Discussion

We have described hybrid time Bayesian networks for modeling dynamic systems with different types of time granularities: the proposed models provide a generalization of continuous-time and discrete-time Bayesian networks. As an inherited property from CTBNs, the joint distribution is propagated over time even when evidence is spaced irregularly. In addition, we established a mapping of hybrid-time networks into a standard BN given time points of interest.

The formalism is related to non-stationary dynamic Bayesian networks, where the structures and parameters are determined by time points of interest [4,5]. These are related in the sense that non-stationary Bayesian networks allow for different time granularities of the (complete) temporal process. The key difference here is that we consider the case where different random variables evolve at different kinds of rates.

A limitation of HTBN is that so far the granularities of discrete-time variables are assumed to be fixed, as the focus of this paper has been on the combination of continuous and discrete-time models. As future work, we will also combine different discrete-time granularities within the hybrid-time framework as proposed in irregular-time Bayesian networks (ITBNs) [6] and also discussed by van der Heijden and Lucas [7]. Furthermore, as a final piece of future work, we would like to extend the formalism to also allow random variables that are completely atemporal. For example, in classification, one might want to predict a single outcome indicator based on time-series. This would complete the full spectrum of temporal models of random variables.

References

1. Bettini, C., Jajodia, S., Wang, S.: Time Granularities in Databases, Data Mining, and Temporal Reasoning. Springer, Heidelberg (2000)
2. Murphy, K.P.: Dynamic Bayesian networks: representation, inference and learning. Ph.D. thesis, University of California, Berkeley (2002)
3. Nodelman, U., Shelton, C.R., Koller, D.: Continuous time Bayesian networks. In: Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence, pp. 378–387. Morgan Kaufmann Publishers Inc. (2002)
4. Grzegorczyk, M., Husmeier, D.: Non-stationary continuous dynamic Bayesian networks. In: Advances in Neural Information Processing Systems, pp. 682–690 (2009)

5. Robinson, J.W., Hartemink, A.J.: Non-stationary dynamic Bayesian networks. In: Advances in Neural Information Processing Systems, pp. 1369–1376 (2009)
6. Ramati, M., Shahar, Y.: Irregular-time Bayesian networks. In: UAI 2010: Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, pp. 484–491 (2010)
7. van der Heijden, M., Lucas, P.J.: Probabilistic reasoning with temporal indeterminacy. In: PGM 2012: Proceedings of the 6th European Workshop on Probabilistic Graphical Models (2012)