# On the Modularisation of Independence in Dynamic Bayesian Networks

Ildikó Flesch[a], Peter Lucas[a] and Stefan Visscher[b]

[a]Department of Information and Knowledge Systems, ICIS,
Radboud University Nijmegen, The Netherlands
[b]Department of Infectious Diseases
University Medical Center Utrecht, The Netherlands
Email: {ildiko,peterl}@cs.kun.nl,S.Visscher@umcutrecht.nl

## Abstract

Dynamic Bayesian networks are Bayesian networks which explicitly incorporating the dimension of time. They are distinguished into repetitive and non-repetitive networks. Repetitive networks have the same set of random (statistical) variables and independence relations at each time step, whereas in non-repetitive networks the set of random variables and the independence relations between these random variables may vary in time. Due to their structural symmetry, repetitive networks are easier to use and are, therefore, often considered as the standard dynamic Bayesian networks. However, repetitiveness is a very strong assumption, which usually does not hold, because dependences and independences that only hold at certain time steps may be lost.

In this paper, we propose a new framework for the modularisation of non-repetitive dynamic Bayesian networks, which offers a practical approach to coping with the computational and structural difficulties associated with unrestricted dynamic Bayesian networks. This framework is based on separating temporal and atemporal independence relations in the model. We investigate properties of the modularisation and show to be compositive.

## 1 Introduction

Probabilistic graphical models are increasingly adopted as tools for the modelling of domains involving uncertainty. For the development of practical applications, especially Bayesian networks have gained much popularity. When considering these application domains, it appears that so far limited attention has been given to the modelling of uncertain time-related phenomena, which occur in many of these domains. Bayesian networks in which some notion of time is explicitly dealt with are usually called *dynamic* Bayesian networks (DBNs) [4]. In some domains involving time, such as speech recognition, the use of DBNs has been extensively explored (e.g. [2]), and technical issues such as concerning reasoning (e.g. [6]) and learning (e.g. [3]) in DBNs have been investigated.

DBNs are distinguished into two main classes: repetitive and non-repetitive networks. Repetitive networks have the same set of random variables and independence relations at each time step, whereas in non-repetitive networks the set of random variables and the independence relations between these random variables may vary in time. The simpler structure of repetitive networks provides significant advantages in terms of computational complexity

and ease of modelling. Therefore, they are often seen as the standard model of DBNs and they have been extensively explored (See [5] for an overview). However, repetitiveness is a very strong assumption that normally will not hold.

In recent literature, the practical use of non-repetitive models has already been explored [8]. Separating temporal and atemporal information in DBNs may be useful, as it (i) helps experts gaining insight into the relations in the network, (ii) may help overcome computational limitations and (iii) may provide an opportunity for learning procedures to obtain more accurate models. However, so far no research has been carried out into how the temporal and atemporal independence relations can be characterised.

In this paper, it is studied what happens with the represented Markov properties, and therefore also with the associated conditional independence assumptions, when we make an explicit distinction between temporal and atemporal structures. It is shown that this distinction allows decomposing the Markov properties into parts, such that the properties of the individual parts can be investigated separately. As we will see, these individual parts cannot be joined together to define the entire set of relations in the DBN without the consideration of some significant properties. Therefore, the way temporal and atemporal parts of the Markov properties interact, is also studied, defining an operator that joins atemporal and temporal relations together in a *correct* way. As considering the temporal and atemporal parts, and their interaction, of a DBN involves studying particular fragments of an independence relation, our research is related to work on multiply sectioned Bayesian networks [7] and multinetwork models [1]. Yet, the constraints imposed by the modelling of time in DBNs give rise to results which are nevertheless distinct.

This paper is organised as follows. In Section 2, we introduce a real-world medical example that illustrates the need of non-repetitive DBNs. In Section 3, some necessary concepts from graph theory as well as the basic principles of Bayesian networks are briefly reviewed. In Section 4, the definition of DBNs is introduced, which includes definitions for the ability to distinguish temporal and atemporal independence relations. Next, in Section 5, the join operator and its properties are introduced. Subsequently, in Section 6, the join operator will be applied to join atemporal and temporal relations to obtain the entire set of independence relations. Finally, in Section 7, we summarise the results that have been achieved.

## 2 Motivating example: the disease course of VAP

A real-world non-repetitive DBN of the disease course of a form of pneumonia is used as motivating example of this paper. We briefly describe this disease and then discuss the construction of a DBN for pneumonia.

Pneumonia frequently develops in ICU patients, as these patients are critically ill and often do they need respiratory support by a mechanical ventilator. After admission to a hospital, all patients become colonised by bacteria. In particular, mechanically ventilated patients run the risk of subsequently developing pneumonia caused by these bacteria; this type of pneumonia is known as *ventilator-associated pneumonia*, or VAP for short. Typical signs and symptoms of VAP include: high *body temperature*, decreased lung function (measured by the $PaO_2/FiO_2$ ratio) and evidence of pneumonia on the chest *X-ray*. By carrying out a dependency analysis on a retrospective, temporal dataset, with data of ICU patients collected during a period of three years, we were able to study how independence information changed in the course of time. Taking the duration of mechanical ventilation as the parameter defining the time steps,
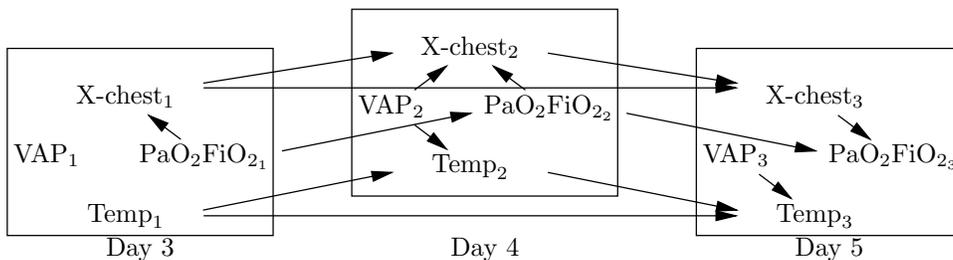
Figure 1: The non-repetitive dynamic Bayesian network for VAP.

taking into account knowledge of two infectious disease experts, we have focused on modelling the course of the development of VAP at 3, 4 and 5 days after admission. The resulting DBN is shown in Figure 1.

The dependence analysis for data on day 3 suggests that there is at that time no dependence between VAP and the signs and symptoms. However, a dependence between $PaO_2/FiO_2$ and chest X-ray was found, which seems logical: a pneumonia-affected lung, as demonstrated on the chest X-ray, will give rise to decreased lung function, measured by the $PaO_2/FiO_2$ ratio. These results are consistent with clinical evidence: after 3 days, VAP signs and symptoms become manifest. The dependence analysis for data on day 4 suggested, in addition, that as VAP develops (which is an infectious disease), body temperature is increased due to fever. Analysis of the data for day 5 shows again the relation between lung function ($PaO_2/FiO_2$) and chest X-ray. Apparently this is an important and strong relation. For day 5 the relation between VAP and X-ray of the chest has disappeared. This is explained by noting that signs of the chest X-ray improve after start of treatment (a variable not modelled). The temporal arcs were subsequently learnt from the data taking the dependences just mentioned as a starting point. The result is a non-repetitive DBN.

## 3 Basic notions

We will be concerned in this paper with acyclic directed graphs (ADGs), denoted as a pair $G = (V, A)$, where $V$ is a set of vertices and $A \subseteq V \times V$ is a set of *arcs*. A *path* is a sequence of vertices $v_1, v_2, \ldots, v_m$, with $(v_k, v_{k+1}) \in A$, also denoted by $v_k \to v_{k+1}$, where $v_1, v_2, \ldots v_{m-1}$ are required to be distinct. A path where for each $k$ it holds that $v_k \to v_{k+1}$ is called a *directed path*. A *directed cycle* is a directed path with $v_1 = v_m$. A *trail* in a graph is a sequence of unique vertices $v_1, v_2, \ldots, v_m$, where we have for each $k$ that $v_k \to v_{k+1}$ or $v_{k+1} \to v_k$; each arc occurs only once. A *subtrail* of a trail $v_1, v_2, \ldots, v_m$ is a sequence $v_i, v_{i+1}, \ldots, v_j$. A trail $\tau$ connecting vertices $u$ and $v$ is also written as $u \mapsto v$. The set of all trails of a direct graph $G$ is denoted by $\Theta$. A graph $G'_{|\Theta'} = (V', A')$ is called a *reduced subgraph* of graph $G = (V, A)$ with associated set of trails $\Theta$ if $V' \subseteq V$ and $A'$ consists of all arcs of the set of trails $\Theta'$ with $\Theta' \subseteq \Theta$.

Let $X$ be a set of *discrete* random variables and let $V$ act as its index set, i.e., $X_v$ with $v \in V$ denotes a random variable and $X_W$ with $W \subseteq V$ denotes a set of random variables. Furthermore, let $P$ denote a joint probability distribution(JPD) of $X$. Let $U, W, Z \subseteq V$ be disjoint sets of indices, then $X_U$ is said to be *conditionally independent* of $X_W$ *given* $X_Z$, if

$$P(X_U \mid X_W, X_Z) = P(X_U \mid X_Z), \tag{1}$$

3

denoted by $U \perp\!\!\!\perp_P W \mid Z$. The entire set of independence relations of $P$ is denoted by $\perp\!\!\!\perp_P$.

The independence relation $\perp\!\!\!\perp_P$ can also be represented although not always perfectly (see below) by means of an ADG $G$, in which the entire set of independence relations is denoted by $\perp\!\!\!\perp_G$. In the graph arcs represent dependences, and absence of arcs represents (conditional) independences. These independences can be read off by the d-separation criterion, defined as follows. A trail $\tau$ in an ADG $G$ is said to be *blocked* by a set $Z$ if one of the following conditions is satisfied: (i) $v \in Z$ and $v$ appears on the trail $\tau$, and either *no* or *only one* of the arcs of $\tau$ meeting at $v$ is directed to $v$; (ii) $v \notin Z$, $\delta(v) \cap Z = \varnothing$, where $\delta(v)$ are the descendants of $v$, and *both* the arcs meeting at $v$ on $\tau$ is directed to $v$. It is said that the sets $U$ and $W$ are *d-separated* by $Z$ if any trail between a vertex in $U$ and a vertex in $W$ is blocked by the set $Z$; formally: $U \perp\!\!\!\perp_G W \mid Z$ (*global directed Markov property*). Otherwise, $U$ and $W$ are *d-connected* by $Z$, denoted by $U \not\perp\!\!\!\perp_G W \mid Z$.

Unfortunately, not every independence encoded in a JPD can be represented graphically by means of d-separation in an associated ADG. An ADG $G$ is said to be a directed *I-map* if each independence in $G$ is also valid in $P$. If it is impossible to omit any of the arcs in an ADG $G$ without losing the property that it is an I-map of $P$, $G$ is said to be a *minimal* I-map of $P$.

Let $G = (V, A)$ be an ADG and let $X_V$ be a set of random variables corresponding to the vertex set $V$ and let $P$ denote the joint probability distribution of $X_V$, then a *Bayesian network* is a pair $\mathcal{B} = (G, P)$, with $G$ being an I-map of $P$.

# 4   Dynamic Bayesian networks

In this section, we begin by defining DBNs, which are an extension of ordinary Bayesian networks and allow modelling the uncertainty involved in processes regarding the dimension of time. Subsequently, two independence reading-off methods are defined for the ability to distinguish between atemporal and temporal independences. In this paper, *time* is denoted by $T$ and is assumed to be a subset of the set of the natural numbers; a *time point* $t$ is then a member of $T$. From now on, let $T$ stand for the time axis with associated total order $< \subseteq T \times T$.

## 4.1   Basic elements

Independence relationships between random variables with the *same* time point are represented by means of an acyclic directed graph, called a *timeslice*. Between timeslices, vertices corresponding to random variables may be linked to each other by means of so-called temporal arcs. Thus, a DBN consists of two parts: (i) a time-independent atemporal part (the timeslices), and (ii) a time-dependent temporal part. First, we consider the atemporal part.

**Definition 1** (*timeslice and atemporal arcs*) *An ADG $G_t = (V_t, A_t^a)$, with the set of vertices $V_t$ and the set of* atemporal arcs $A_t^a \subseteq V_t \times V_t$, $t \in T$, *is called a* timeslice *at time point $t$.*

The set of all timeslices $G$ of a DBN is taken as:

$$G = \{G_t \mid t \in T\} = \{(V_t, A_t^a) \mid t \in T\} = (V_T, A_T^a) \ . \tag{2}$$

4

An arc $(u_t, v_{t'})$ with $t < t'$ is called a *temporal arc*. The set of temporal arcs of set $G$ is denoted by $A^t$. Thus temporal arcs connect timeslices with strict direction from the past to the future giving rise to the following structure.

**Definition 2** *(temporal network)* *A temporal network $N$ is defined as a pair $N = (V_T, A)$, where $G = (V_T, A_T^a)$ and $A = A_T^a \cup A^t$.*

Clearly, a temporal network $N$ is also an ADG. A DBN is defined as a pair $\mathcal{DBN} = (N, P)$, where $P$ is the JPD on $X_{V_T}$. In the remaining part of this paper, when the symbol $t$ is used as a superscript it indicates a temporal property; if it is used as a subscript it acts as a time index.

Both temporal and atemporal relations in the network can be represented by means of trails. An *atemporal trail* contains no temporal arcs and is denoted by $\tau^a$. A *temporal trail* consists of *at least one* temporal arc and is denoted by $\tau^t$. The sets of atemporal and temporal trails are denoted by $\Theta^a$ and $\Theta^t$, respectively. Considering the temporal relationships we only need to consider temporal trails resulting into a reduced temporal network.

**Definition 3** *(reduced temporal network)* *Let $N = (V_T, A)$ be a temporal network. Then $N_{|\Theta^t} = (V_T, A_{\Theta^t})$ with the set of vertices $V_T$ and the set of all arcs that are included in any temporal trail $A_{\Theta^t} \subseteq A$, is called a reduced temporal network.*

Observe that the reduced temporal network is based on the set of temporal *trails*, which may consist of both atemporal and temporal trails. A further partitioning of the reduced temporal network is based on its set of *arcs*. This partitioning is obtained by decomposing a reduced temporal network into two parts, where one part consists of only atemporal and the another part of only temporal arcs. The *atemporal part of the reduced temporal network* is denoted by $N_{|\Theta^t, A_T^a} = (V_T, A_{\Theta^t, A_T^a})$, where $V_T$ is the set of vertices and $A_{\Theta^t, A_T^a} \subseteq A_T^a$ consists of all atemporal arcs in the reduced temporal network. The *temporal part of the reduced temporal network* is denoted by $N_{|\Theta^t, A^t} = (V_T, A_{\Theta^t, A^t})$, where $A_{\Theta^t, A^t} \subseteq A^t$ consists of all temporal arcs in the reduced temporal network.

As an example of a structure of a DBN consider Figure 1, where timeslices are depicted by rectangles. It holds that $G_1 = (VAP_1, X\text{-}chest_1, PaO_2FiO_{2_1}, Temp_1\}, \{(PaO_2FiO_{2_1}, X\text{-}chest_1)\})$ and $(X\text{-}chest_1, X\text{-}chest_2) \in A^t$. An atemporal trail is $\tau^a = PaO_2FiO_{2_1} \rightarrow X\text{-}chest_1$ and a temporal one is $\tau^t = X\text{-}chest_1 \rightarrow X\text{-}chest_2 \leftarrow PaO_2FiO_{2_2}$.

## 4.2 Atemporal and temporal d-separation

As a DBN includes temporal and atemporal elements, the question is how to distinguish between these relations. In this section, we define the necessary independence reading-off methods.

We start by considering time-independent relations by ignoring the role played by temporal arcs. To obtain these relations, we apply the *atemporal d-separation* criterion, as follows. Let $U, W, Z \subseteq V_T$ be distinct vertex sets, then, if each atemporal trail connecting any vertex in $U$ with any vertex in $W$ is blocked by the set $Z$, then $U$ and $W$ are said to be *atemporally d-separated*, denoted by $\perp\!\!\!\perp_G$, given $Z$ written $U \perp\!\!\!\perp_G W \mid Z$; otherwise, they are said to be *atemporally d-connected*, formally $U \not\perp\!\!\!\perp_G W \mid Z$. Atemporal d-separation among vertices belonging to only one timeslice $G_t$ is denoted by $\perp\!\!\!\perp_{G_t}$; similarly, atemporal d-connection is in that case denoted by $\not\perp\!\!\!\perp_{G_t}$.

*Temporal d-separation* is applied to obtain the time-dependent conditional independences of the temporal network as follows. If each temporal trail between any vertex in $U$ and any vertex in $W$ is blocked by the set $Z$, then, $U$ and $W$ are said to be *temporally d-separated* given $Z$, written $U \perp\!\!\!\perp_{N_{|\Theta^t}} W \mid Z$; otherwise, they are called *temporally d-connected*, denoted by $U \not\perp\!\!\!\perp_{N_{|\Theta^t}} W \mid Z$. The entire set of independence relations in the reduced temporal network is denoted by $\perp\!\!\!\perp_{N_{|\Theta^t}}$.

Finally, we also apply d-separation defined in Section 3 to the entire DBN to obtain all (conditional) independences. The resulting relation $\perp\!\!\!\perp_N$ denotes the set of independences in network $N$, and is also interpreted as minimal I-map of $P$.

# 5 The join operator

In the previous section, DBNs and their (a)temporal relations have been defined. In this section, we define the join operator applied to correctly join atemporal and temporal independence relations of DBNs. In Section 5.1, two significant properties are introduced, which have to be included in the definition of the join operator. In Section 5.2, the join operator and its properties are considered.

## 5.1 Dependence preservation and independence concatenation

In this section, we introduce the dependence preservation and independence concatenation properties, which are necessary to join two independence relations in the *right* way. Dependence preservation will allow us to join an independence and a dependence relation in a correct way, whereas independence concatenation joins two independence relations taking into account that these relations after the join process may give rise to a dependence, as we will see.

We begin with the discussion of dependence preservation. The reason that dependence preservation is required for joining independence statements can be explained in terms of the concepts of consistency and dominance as follows.

Let the independence relations $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ be defined on the same vertex set $V$. Then, if there exist $U \perp\!\!\!\perp W \mid Z$ and $U \not\perp\!\!\!\perp' W \mid Z$ for arbitrary, mutually disjoint sets of vertices $U, W, Z \subseteq V$, then these independence statements and therefore independence relations $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ are said to be *inconsistent*. Otherwise, the statements are *consistent*. Our purpose is to join independence relations together, however, in the case, when two independence statements and therefore relations are inconsistent, a choice has to be made between the independence and dependence effecting the inconsistence. In other words, one has to *dominate* the other one. Thus, if the relations $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ are inconsistent due to the statements $U \perp\!\!\!\perp W \mid Z$ and $U \not\perp\!\!\!\perp' W \mid Z$ then $U \not\perp\!\!\!\perp' W \mid Z$ is said to *dominate* $U \perp\!\!\!\perp W \mid Z$. Since dominance has to be taken into account during joining independences, the following property is defined.

**Definition 4 (*dependence preservation*)** *Let* $\perp\!\!\!\perp$, $\perp\!\!\!\perp'$ *and* $\perp\!\!\!\perp''$ *be independence relations all defined on* $V$. *Then, if* $U \not\perp\!\!\!\perp W|Z$ *or* $U \not\perp\!\!\!\perp' W|Z$ *and it holds that* $U \not\perp\!\!\!\perp'' W|Z$ *for all* $U, W, Z \subseteq V$, *then it is said that* $\perp\!\!\!\perp''$ *satisfies the* dependence preservation *property with regard to* $\perp\!\!\!\perp$ *and* $\perp\!\!\!\perp''$.

Next, we define the independence concatenation property. This takes into account that when independence relations are combined an independence may change into a dependence.
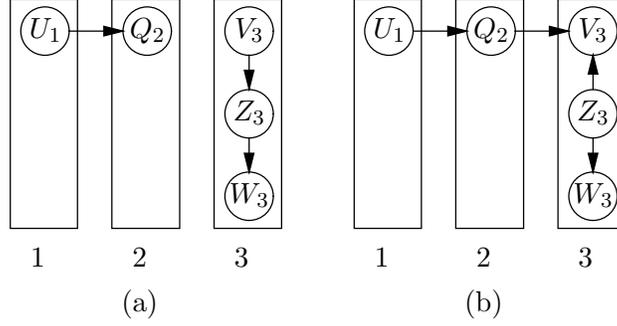
Figure 2: Temporal networks (a) and (b).

This is demonstrated in Figure 2. In both temporal networks (a) and (b) we have that $U_1 \perp\!\!\!\perp_{N_{|\Theta^t, A_T^a}} W_3 \mid V_3$ and $U_1 \perp\!\!\!\perp_{N_{|\Theta^t, A^t}} W_3 \mid V_3$. However, in temporal network (a) it holds that $U_1 \perp\!\!\!\perp_{N_{|\Theta^t}} W_3 \mid V_3$ and in temporal network (b) we have $U_1 \not\perp\!\!\!\perp_{N_{|\Theta^t}} W_3 \mid V_3$. Observe that if joining two independences results in dependence, then this dependence is always represented by a temporal trail, which *has* to consist of at least one atemporal and at least one temporal subtrail. Therefore, the (in)dependence relations can be examined separately using blockage in terms of atemporal and temporal subtrails, providing the following proposition, as basis for the definition of independence concatenation.

**Proposition 1** *Let $U, W, Z \subseteq V_T$ be disjoint sets of vertices and let temporal trail $\tau^t$ connect vertices $u$ and $w$, $u \in U, w \in W$, in the reduced temporal network $N_{|\Theta^t}$. If (i) one of the (a)temporal subtrails of $\tau^t$ is blocked by $Z$ or (ii) each (a)temporal subtrail of $\tau^t$ is not blocked by $Z$ but there is at least one convergent connection of two consecutive (a)temporal subtrails not included in $Z$ then the temporal $\tau^t$ is blocked by $Z$. Otherwise $\tau^t$ is not blocked by $Z$.*

Note that Proposition 1 provides the basis for considering the case, when joining two independences results in a dependence. To ensure us that these dependences are taken in the new joined independence relation the independence concatenation property is defined.

**Definition 5** *(**independence concatenation**) Let $U, W, Z \subseteq V_T$ be disjoint vertex sets in a temporal network $N$, and let $G$, $G'$ and $G''$ be reduced subgraphs of $N$ with corresponding sets of trails $\Theta, \Theta'$ and $\Theta''$. If each trail $u \mapsto w$, $u \in U, w \in W$ in subgraphs $G$ and $G'$ is blocked by $Z$ and if for each $q \in V_T \setminus (U \cup W)$ one of the trails $u \mapsto q \in \Theta$ and $q \mapsto w \in \Theta'$ is blocked by $Z$ or these two trails do not constitute a convergent connection at $q$ included in $Z$, then if for $U \perp\!\!\!\perp_G W \mid Z$ and $U \perp\!\!\!\perp_{G'} W \mid Z$ it holds that $U \perp\!\!\!\perp_{G''} W \mid Z$, it is said that $\perp\!\!\!\perp_{G''}$ satisfies the* independence concatenation *property with regard to $\perp\!\!\!\perp_G$ and $\perp\!\!\!\perp_{G'}$.*

## 5.2 The join operator

In this section, the join operator is defined and significant properties of this operator are given (other properties are omitted by the space limitation). We start to define the join operator.

**Definition 6** *(**join operator**) Let $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ be two independence relations defined on the same vertex set $V$. The join operator is denoted by $\circ$. The* join *of these two relations, denoted by $\perp\!\!\!\perp \circ \perp\!\!\!\perp' = \perp\!\!\!\perp''$, is then again an independence relation, $\perp\!\!\!\perp''$, defined on $V$, satisfying the dependence preservation and the independence concatenation properties.*

The dependence preservation property implies the following proposition.

**Proposition 2** *Let $G = (V, A), G' = (V, A')$ and $G'' = (V, A'')$ be three ADGs, where $A \cup A' \subseteq A''$. Then, it holds that $\perp\!\!\!\perp_{G''} \subseteq \perp\!\!\!\perp_G \circ \perp\!\!\!\perp_{G'}$.*

Clearly, dependence preservation results in the property that I-mappedness is preserved. In Proposition 2 we have used the subset relation $\subseteq$, as the resulting ADG need not precisely consist of the union of the set of arcs of graphs $G$ and $G'$. If the graph $G''$ equals the union and $\perp\!\!\!\perp_{G''} = \perp\!\!\!\perp_G \circ \perp\!\!\!\perp_{G'}$ holds, the join operator is said to be *minimally* dependence preserving. The dependence preservation implies the following proposition.

**Proposition 3** *Let $\perp\!\!\!\perp$ and $\perp\!\!\!\perp'$ be two independence relations defined on the same vertex set $V$. Then for the independence relation $\perp\!\!\!\perp'' = \perp\!\!\!\perp \circ \perp\!\!\!\perp'$ it holds that $\perp\!\!\!\perp'' \subseteq (\perp\!\!\!\perp \cap \perp\!\!\!\perp')$.*

# 6   Temporal and atemporal independence: their interaction

In Section 4, we have defined DBNs, with a graphical representation consisting of atemporal and temporal independence relations. In Section 5, the join operator has been defined such that it satisfies dependence preservation and independence concatenation. In this section, based on the two previous sections, we investigate how to employ the join operator for joining the temporal and atemporal independence relations underlying temporal networks and therefore DBNs, where we distinguish between the temporal and atemporal independence relations for supporting modelling of *non-repetitive* DBNs. In Section 6.1, we show that the join operator can be used for joining the atemporal and temporal parts of the reduced temporal network. In Section 6.2, we show the same for the atemporal and temporal relations in the entire temporal network.

## 6.1   Joining atemporal and reduced temporal networks

In this section, we start by the consideration of the relations $\perp\!\!\!\perp_{G_t}$ and $\perp\!\!\!\perp_G$. Here, the following proposition establishes the connection between the d-separation relations of the individual timeslices $G_t$ and of the set of timeslices $G$.

**Proposition 4** *Let $\mathcal{DBN} = (N, P)$, with temporal network $N = (V_T, A)$, $G = (V_T, A_T^a)$, $A = A_T^a \cup A^t$ and JPD P, then: (i) $\perp\!\!\!\perp_G = \cap_{t \in T} \perp\!\!\!\perp_{G_t}$, and (ii) $\not\perp\!\!\!\perp_G = \cup_{t \in T} \not\perp\!\!\!\perp_{G_t}$.*

**Proof**: Case $(i)$. The conditional independence relation $\perp\!\!\!\perp_{G_t}$ consists of the complete set of conditional independences regarding its own vertex set $V_t$, which is, except when $G_t$ is an empty graph, not completely included in any other $\perp\!\!\!\perp_{G_{t'}}$, $t \neq t'$. Therefore, the complete set of conditional independences of $G$ is obtained by the intersection. Furthermore, the atemporal relations *always* describe independences between vertex sets of different timeslices. These sets of independences are the same for each relation $\perp\!\!\!\perp_{G_t}$, and these independence statements are also included in $\perp\!\!\!\perp_G$ by means of intersection.
Case $(ii)$. $\not\perp\!\!\!\perp_G = \overline{\cap_{t \in T} \perp\!\!\!\perp_{G_t}} = \cup_{t \in T} \not\perp\!\!\!\perp_{G_t}$.                        □

Observe that according to Proposition 4, the join operator $\circ$ is interpreted as the intersection of the independence relations, which does not hold in general.

The connection between the join operator for reduced temporal networks is established by the following proposition.

**Proposition 5** *Let the reduced temporal networks $N_{|\Theta^t, A_T^a}$ and $N_{|\Theta^t, A^t}$ be the atemporal and temporal parts of the reduced temporal network with independence relations $\perp\!\!\!\perp_{N_{|\Theta^t, A_T^a}}$ and $\perp\!\!\!\perp_{N_{|\Theta^t, A^t}}$, respectively. Then, the independence relation of the reduced temporal network $N_{|\Theta^t}$ can be obtained by $\perp\!\!\!\perp_{N_{|\Theta^t}} = \perp\!\!\!\perp_{N_{|\Theta^t, A_T^a}} \circ \perp\!\!\!\perp_{N_{|\Theta^t, A^t}}$.*

**Proof**: Since the join operator satisfies independence concatenation and dependence preservation, joining two independence relations, the join operator has inserted all the independences, which are obtained applying the d-separation criterion. Furthermore, these two properties ensure us that there is also no uncorrect relation in $\perp\!\!\!\perp_{N_{|\Theta^t}}$. $\qquad\square$

We have joined two independence relations, however, we still have to show the correct working of the join operator, for which soundness and completeness are the necessary and sufficient conditions. The join operator $\circ$ is called *sound* if all $r \in (\perp\!\!\!\perp \circ \perp\!\!\!\perp')$ is correct in the joined independence relation $\perp\!\!\!\perp''$. The join operator $\circ$ is *complete* if for all $r \in \perp\!\!\!\perp''$ it holds that $r \in (\perp\!\!\!\perp \circ \perp\!\!\!\perp')$.

The following theorem implies that the join operator is sound and complete for joining the atemporal and temporal independence relations of the reduced temporal network.

**Theorem 1** *The join operator applied as $\perp\!\!\!\perp_{N_{|\Theta^t}} = \perp\!\!\!\perp_{N_{|\Theta^t, A_T^a}} \circ \perp\!\!\!\perp_{N_{|\Theta^t, A^t}}$ is sound and complete.*

**Proof**: Soundness: By the temporal d-separation criterion, since the two independence relations between two vertex sets will only be joined into an independence relation, if each trail connecting these two vertex sets is blocked by the conditioning vertex set, independence in the resulting relation is implied. Completeness: We prove the completeness of the join operator by proving that if $N_{|\Theta^t, A_T^a}$ and $N_{|\Theta^t, A^t}$ are I-maps then $N_{|\Theta^t}$ is also an I-map. The preservation of I-mappedness of the reduced temporal network follows from Proposition 2. Furthermore, the join operator satisfies also the minimally dependence preserving property, since $N_{|\Theta^t}$ does not contain any extra arcs. $\qquad\square$

## 6.2 Joining it all together

In this subsection, the temporal and atemporal independence relations are joined together obtaining the relation $\perp\!\!\!\perp_N$. Recall that the atemporal relations are defined by the atemporal properties of the graph. The relation $\perp\!\!\!\perp_G$ is obtained by the application of the concept of atemporal d-separation. The temporal relations are meant to be the relations which are recovered by the temporal d-separation criteria and are denoted by $\perp\!\!\!\perp_{N_{|\Theta^t}}$. The following proposition implies the correct way to join these relations.

**Proposition 6** *Let $\perp\!\!\!\perp_G$ and $\perp\!\!\!\perp_{N_{|\Theta^t}}$ be atemporal and temporal independences of the temporal networks. Then, the entire set of independence relations of the temporal network can be obtained by using the join operator as $\perp\!\!\!\perp_G \circ \perp\!\!\!\perp_{N_{|\Theta^t}} = \perp\!\!\!\perp_N$.*

**Proof**: Let $U, W, Z \subseteq V_T$ be disjoint vertex sets in the temporal network $N$. As each trail $u \mapsto w$, $u \in U, w \in W$ is blocked by $Z$ then $U \perp\!\!\!\perp_G W \mid Z \circ U \perp\!\!\!\perp_{N_{|\Theta^t}} W \mid Z = U \perp\!\!\!\perp_N W \mid Z$ holds. $\qquad\square$

Figure 3 summarises the way to join independence relations defined in propositions 4, 5 and 6.

The join operator also satisfies the conditions for soundness and completeness.
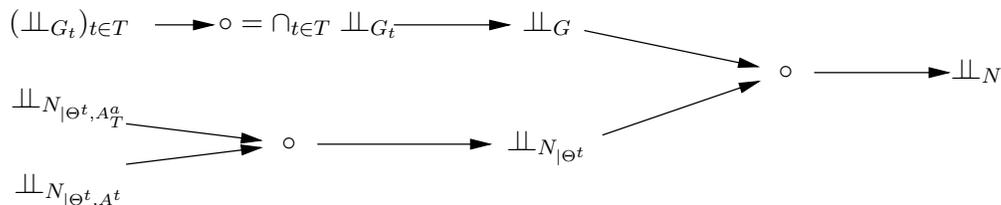
Figure 3: Joining temporal and atemporal independence relations.

**Theorem 2** *The join operator joining $G$ and $N_{|\Theta^t}$ into temporal network $N$ is sound and complete.*

**Proof**: Soundness follows from the temporal d-separation criterion. If each trail connecting the two joined vertex sets is blocked by the conditioning vertex set, independence in the resulting relation is implied. The proof of completeness is also similar to the proof of Theorem 1. If the relations $\perp\!\!\!\perp_G$ and $\perp\!\!\!\perp_{N_{|\Theta^t}}$ are I-maps then the relation $\perp\!\!\!\perp_N$ is also an I-map according to Proposition 2. $\qquad\qquad\Box$

Finally, the various independence relations can be compared to each other.

**Proposition 7** *The independence sets $\perp\!\!\!\perp_G$, $\perp\!\!\!\perp_{N_{|\Theta^t}}$, $\perp\!\!\!\perp_{N_{|\Theta^t,A_T^a}}$, $\perp\!\!\!\perp_{N_{|\Theta^t,A^t}}$ and $\perp\!\!\!\perp_N$ also satisfy the following properties:*

- $\perp\!\!\!\perp_{N_{|\Theta^t}} \subseteq \perp\!\!\!\perp_{N_{|\Theta^t,A_T^a}}$, $\perp\!\!\!\perp_{N_{|\Theta^t}} \subseteq \perp\!\!\!\perp_{N_{|\Theta^t,A^t}}$;

- $\perp\!\!\!\perp_N \subseteq \perp\!\!\!\perp_G$, $\perp\!\!\!\perp_N \subseteq \perp\!\!\!\perp_{N_{|\Theta^t}}$.

# 7 Conclusions

The aim of the research described in this paper was to study how the modelling non-repetitive DBNs can be simplified by distinguishing between time-independent and time-dependent independence relations. As this gave rise to various separate, but linked, independence relationships, the usual property that independence and dependence complement each other no longer holds. We introduced a join operator with special semantics to overcome this problem. Using the join operator allows one to build DBNs in a modular fashion, hence the title of the paper. As far as we know, this paper offers the first systematic method for buiding non-repetitive DBNs. Much work still needs to be done to bridge the gap between the theoretical work in this paper and practice.

# References

[1] E. Castillo, J.M. Gutiérrez, and A.S Hadi. Expert Systems and Probabilistic Network Models. Springer-Verlag, New York, 1997.

[2] M. Deviren and K. Daoudi. Continuous speech recognition using dynamic Bayesian networks: a fast decoding algorithm. In: Proc PGM'02, Cuenca, Spain, 2002, pp. 54–60.

[3] N. Friedman, K. Murphy, and S. Russell. Learning the structure of dynamic probabilistic networks. In: Proc 14th UAI, 1998, pp. 139–147.

[4] F.V. Jensen. Bayesian Networks and Decision Graphs. Springer, New York, 2001.

[5] K. P. Murphy Dynamic Bayesian Networks: Representation, Inference and Learning Dissertation, 2002

[6] U. Kjaerulff. A computational scheme for reasoning in dynamic probabilistic networks. In: Proc UAI'92, 1992, pp. 121–129.

[7] Y. Xiang, and V. Lesser. Justifying multiply sectioned Bayesian networks. In: Proc 6th Int Conf on Multiagent Systems, Boston, 2000, pp. 349–356.

[8] A. Tucker and X. Liu. Learning Dynamic Bayesian Networks from Multivariate Time Series with Changing Dependencies. The 5th International Symposium on IDA, 2003.