



## Describing disease processes using a probabilistic logic of qualitative time



Maarten van der Heijden<sup>a,b,\*</sup>, Peter J.F. Lucas<sup>a</sup>

<sup>a</sup> Institute for Computing and Information Sciences, Radboud University Nijmegen, PO Box 9010, 6500GL Nijmegen, The Netherlands

<sup>b</sup> Department of Primary and Community Care, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500HB Nijmegen, The Netherlands

### ARTICLE INFO

#### Article history:

Received 29 April 2013

Received in revised form

25 September 2013

Accepted 25 September 2013

#### Keywords:

Knowledge representation

Temporal reasoning

Probabilistic logic

Dynamic Bayesian networks

Chronic obstructive pulmonary disease

### ABSTRACT

**Background:** Clinical knowledge about progress of diseases is characterised by temporal information as well as uncertainty. However, precise timing information is often unavailable in medicine. In previous research this problem has been tackled using Allen's qualitative algebra of time, which, despite successful medical application, does not deal with the associated uncertainty.

**Objectives:** It is investigated whether and how Allen's temporal algebra can be extended to handle uncertainty to better fit available knowledge and data of disease processes.

**Methods:** To bridge the gap between probability theory and qualitative time reasoning, methods from probabilistic logic are explored. The relation between the probabilistic logic representation and dynamic Bayesian networks is analysed. By studying a typical, and clinically relevant problem, the detection of exacerbations of chronic obstructive pulmonary disease (COPD), it is determined whether the developed probabilistic logic of qualitative time is medically useful.

**Results:** The probabilistic logic extension of Allen's temporal algebra, called *Qualitative Time CP-logic* provides a tool to model disease processes at a natural level of abstraction and is sufficiently powerful to reason with imprecise, uncertain knowledge. The representation of the COPD disease process gives evidence that the framework can be applied functionally to a clinical problem.

**Conclusion:** The combination of qualitative time and probabilistic logic offers a useful framework for modelling knowledge and data to describe disease processes in clinical medicine.

© 2013 Elsevier B.V. All rights reserved.

### 1. Introduction

In solving clinical problems such as diagnosis or prognosis, concerning the signs and symptoms of a disease, one often has to take into account the time when a particular event has occurred or is expected to occur. In many cases, the actual temporal details about when events have occurred are not available, or at least imprecise, whereas one is more certain about the order of the events. AI researchers have traditionally used *Allen's interval algebra* [1] to model imprecise temporal events. It forms the foundation of a temporal logic that supports reasoning about temporal events in a qualitative fashion [2]. Work by Shahar [3] indicates the usefulness of Allen's algebra for describing temporal events in medicine. However, Allen's algebra does not allow expressing uncertainty about the occurrence of the events or their qualitative, temporal

relationships. Yet, uncertainty is a feature of many problems where precise temporal information is missing, such as in clinical medicine.

In the work described in this paper it is investigated in what way Allen's interval algebra can be extended to incorporate uncertainty reasoning. Recently developed probabilistic logics can be used as a basis for such a more general language and in particular we build upon the work on CP-logic [4]. The aim is to design a framework that allows describing disease processes in a way similar to what is found in the clinical literature, i.e. imprecise yet with uncertainty made explicit.

Developing sophisticated decision-support systems for realistic clinical problems requires one to handle both the imprecision and uncertainty of medical knowledge and data. The methods developed in this paper are expected to contribute to meeting these challenges. Evidence that this is justified comes from examples taken from important clinical problems, one of which is the management of chronic obstructive pulmonary disease (COPD) for which we have developed a smartphone-based decision-support system [5].

We claim that to model disease progression it is often easier to start from imprecise notions of temporal ordering compared

\* Corresponding author at: Institute for Computing and Information Sciences, Radboud University Nijmegen, PO Box 9010, 6500GL Nijmegen, The Netherlands. Tel.: +31 243652291.

E-mail addresses: [m.vanderheijden@cs.ru.nl](mailto:m.vanderheijden@cs.ru.nl) (M. van der Heijden), [peterl@cs.ru.nl](mailto:peterl@cs.ru.nl) (P.J.F. Lucas).

to directly constructing for example a dynamic Bayesian network. That is, the kind of information readily available from questioning a patient and from general clinical knowledge about the disease, offers a good starting point. From there on we can add uncertainty to obtain a temporal probabilistic model. This way of modelling clarifies the structure of the process and makes it more explicit what kind of assumptions are made.

The topic of temporal reasoning has already received much attention, also in a medical context, exemplified by the book of Combi et al. [6]. However the work described there is almost completely symbolic in nature, focussing on logic properties. We argue that in a realistic clinical setting one cannot and should not ignore the uncertainties that arise due to hidden complexity, lack of information or measurement error. Whereas pure probabilistic frameworks tackle only the uncertainty, our qualitative time probabilistic logic seeks to model temporal uncertain processes at a practical level of abstraction.

This paper is organised as follows. In the next section we introduce two motivating examples, typical for the problems encountered in biomedicine, which will be used later to validate the framework being developed. In Section 3 we discuss related work and in Section 4 we provide some preliminaries. In Section 5 we describe the probabilistic temporal framework that seems suitable for the description of disease processes, and give some properties of the language. In Section 6 we return to our examples and show that the developed framework can be usefully applied to describe the temporal, uncertain evolution of disease processes.

## 2. Motivation

### 2.1. HIV drug resistance

Our first example, from [7], is modelling HIV mutations and drug resistance. In contrast to the COPD example which will be introduced next, HIV mutations are uncertain temporal events without recurrence. In [7] they represent the problem using temporal nodes Bayesian networks (TNBN), where random variables take time intervals as values to denote time frames in which a mutation could occur. In Section 6 we show that TNBNs can be represented in our framework.

### 2.2. The management of COPD

Throughout the paper we will use the management of chronic obstructive pulmonary disease, COPD for short, to motivate the developed methods. COPD is a progressive lung disease characterised by a mixture of chronic bronchitis and emphysema, leading to decreased respiratory capacity and potentially to respiratory failure and death. Although there are a number of causes, exposure to (tobacco) smoke is the most prevalent. Because COPD is a progressive disease, its temporal development is quite important and even more so because of the occurrence of *exacerbation* events – a worsening of symptoms with possibly a large negative influence on health status.

To model exacerbations we have to take into account that patients are usually in a home care situation which lacks precise high-frequency measurement-equipment. As a consequence we have to rely on imprecise information. Probability theory provides a tool to quantify our uncertainty about the state of the system, in this case the health status of a COPD patient. We focus on a limited number of random variables that characterise the events of the uncertain process: two main symptoms, *dyspnea* and *cough*, a common cause *infection* and the outcome *exacerbation*. As we will

be using a probabilistic logic the relations between these variables can be specified in terms of causal rules:

$dyspnea \leftarrow infection.$

$cough \leftarrow infection.$

$exacerbation \leftarrow dyspnea \wedge cough.$

Limited information also leads to temporal uncertainty, yet it is often possible to determine qualitative temporal relations. We use time intervals to model the duration of symptoms and other variables of interest. Allen's algebra then provides a language to state ordering relations between symptoms in time. The rules above can be extended to incorporate the temporal information:

$dyspnea(I) \leftarrow infection(J), \text{ allen}(I, J).$

$cough(K) \leftarrow infection(J), \text{ allen}(K, J).$

$exacerbation(L) \leftarrow dyspnea(I) \wedge cough(K) \wedge \text{ allen}(L, I) \wedge \text{ allen}(L, K).$

where  $I, J, K, L$  denote the intervals which we associate with the events and *allen* denotes some qualitative time relation between the relevant intervals. To predict exacerbations, we have to quantify the uncertainty of temporal relations between symptoms. This will be the main example, and in Section 6 we will study the situation introduced here in detail.

## 3. Related work

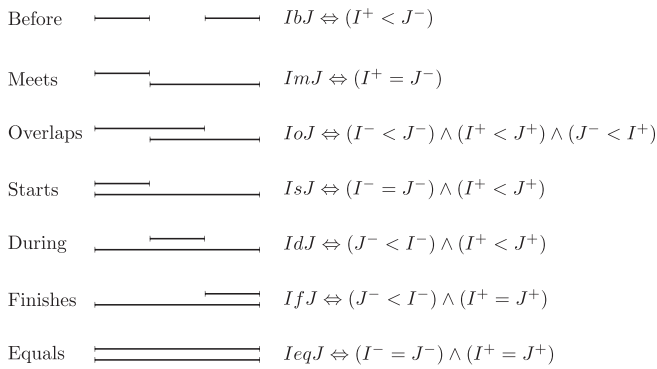
Allen's algebra has received much attention over the years and finds applications in fields from planning to clinical medicine and many more. However there are also numerous other temporal representation and reasoning frameworks. A comprehensive overview lies outside the scope of a related work section, but we mention some important work. For an overview of topics related to time in medicine we refer the reader to [6].

Besides Allen's work [1,2], McDermott's work on temporal logic [8] is well-known. Interesting to note is McDermott's observation that quite a few problems result as a consequence of uncertainty, and that no formal framework exists that satisfactorily combines logic and probability. Fortunately this has changed in recent years, leading to our current work on temporal reasoning in probabilistic logic. Combi et al. [9] describe an extension of Allen's logic that generalises to different temporal granularities, which is often necessary in a clinical context.

Also important to mention is the work on temporal constraint networks by Dechter et al. [10] and on combined metric and Allen constraints by Kautz and Ladkin [11]. Jonsson and Bäckström [12] later showed that disjunctive linear relations subsume these and a number of other temporal constraint formulations. An application of temporal constraints in a medical context can be found in modelling clinical guidelines, see e.g. [13].

In the probabilistic model field there has also been interest in temporal models, and although not directly connected to qualitative time representations, dynamic Bayesian networks [14,15] are a well-known instance of temporal probabilistic models. Many variants and extensions exist such as the networks of probabilistic events [16] where nodes are associated with events and the value of a node represents an occurrence of the event at a particular time point; temporal Bayesian networks of events [17] which are similar and use a subset of Allen relations to relate (quantified) temporal nodes; and the work by Tawfik and Neufeld [18] on temporal reasoning in Bayesian networks. These frameworks do not focus on representing qualitative time as we do in this paper.

Other work on combining probabilistic and qualitative temporal reasoning includes probabilistic temporal interval networks [19], a probabilistic extension of the interval constraints network often used with Allen's algebra. This is similar to our current proposal, but



**Fig. 1.** Graphical representation of the seven basic relations that can hold between two time intervals. These relations and their inverse make up Allen’s algebra.

our work also allows modelling uncertainty in the events. Probabilistic temporal networks, as defined by Santos and Young [20], are network models that incorporate Allen’s constraints on conditional probabilities. They make the assumption however that the intervals of interest are known beforehand and can be specified explicitly, therefore not allowing uncertainty in what intervals are or will be of interest.

Then, recently, probabilistic logic has been applied to represent stochastic processes [21]. The proposed language CPT-L extends CP-logic to represent fully-observable homogeneous Markov processes. Our work differs in focussing on using CP-logic for qualitative time stochastic processes. Distributional clauses [22] extend the probabilistic logic approach (based on Sato’s distribution semantics [23] like CP-logic) to continuous distributions.

#### 4. Preliminaries

##### 4.1. Allen’s interval algebra

Allen’s algebra builds upon qualitative relations between time intervals. An interval implicitly refers to an event that takes place during that interval. In Fig. 1 the relations that can hold between two intervals are shown graphically.

We define intervals as right-open  $I = [I^-, I^+)$  on a linearly ordered time line of points  $(\mathcal{T}, \leq)$ , with  $\mathcal{T}$  a subset of the set of real numbers  $\mathbb{R}$  or, if necessary, restricted to a finite subset of the natural numbers  $\mathbb{N}$ . The special points  $I^-, I^+$  are defined as  $I^- = \inf I$  and  $I^+ = \sup I$ .

We can now define relations, letting  $\mathcal{I}$  be the set of all intervals of  $\mathcal{T}$ . A binary temporal interval relation  $R$  is defined as  $R \subseteq \mathcal{I} \times \mathcal{I}$ . Instead of temporal interval relations, we shall in the following use a predicate logic representation, where the logical atom  $R(I, J)$ , having the meaning of  $(I, J) \in R$  in the relational form, will be denoted in infix form as  $IRJ$ .

Allen defined a set of seven basic interval relations on two time intervals. Together with the inverses of these seven relations we obtain a minimal set of relations that can express any qualitative relation between two intervals. This set of relations with respect to intervals in  $\mathcal{I}$  will be denoted  $\mathcal{B}$  and the thirteen relations therein are  $\mathcal{B} = \{b, \bar{b}, m, \bar{m}, o, \bar{o}, s, \bar{s}, d, \bar{d}, f, \bar{f}, eq\}$ , where  $\bar{r}$  is the inverse relation of  $r$ , which for intervals  $I, J$  is defined as  $\bar{r}I \equiv JrI$ . Fig. 1 gives the definition of the relations in terms of interval endpoints, with  $I^-$  denoting the start point and  $I^+$  the end point of interval  $I$ , and similarly for  $J$ . The basic relations in the set  $\mathcal{B}$  are mutually exclusive and collectively exhaustive with respect to the possible relations between two intervals.

In the following examples we will consider events  $E$  with an interval index  $I$ , written as  $E_I$ , to denote that event  $E$  occurs in interval  $I$ , instead of pure interval expressions, as this makes it easier to convey ideas on how to put time intervals to use. We will also use

the shorthand notation  $E_I R E'_J$  for the formally correct notation  $E_I \wedge E'_J \wedge (IRJ)$  meaning that event  $E$  occurs at time interval  $I$  and  $E'$  at interval  $J$  and that the relationship between these intervals is expressed by the Allen expression  $IRJ$ .

**Example 1.** A certain group of COPD patients tends to have relatively frequent exacerbations – events of worsening of symptoms – that are usually caused by airway infections. Using the basic temporal relations we can describe that an infection in interval  $I$  at least partially precedes the increase in symptoms in interval  $J$ . We then obtain the expression:

$$\text{Inf}_I \circ \text{Sym}_J,$$

which means that symptoms can outlast the infection. Since an exacerbation is defined as an increase of the relevant symptoms in the interval we can say for an interval  $K$  associated with the exacerbation:

$$\text{Exa}_K \text{ eq } \text{Sym}_J.$$

With the basic relations, the full set of Allen’s relations can be constructed. Any of these relations is defined as the disjunction of a subset of the basic relations for the same intervals. Thus, one gets expressions such as  $(INJ) \equiv (IRJ) \vee (IR'J)$ , where  $N$  is the new relation, and  $R, R'$  are basic relations. However, instead of introducing new names for the resulting new relations, Allen uses set notation for the new relations, giving rise to the following definition:

**Definition 1.** An Allen relation is defined as a disjunction of basic interval relations, represented as a set. The power set of the basic relations contains all Allen relations and is denoted  $\mathcal{A} = \wp(\mathcal{B})$ . An interval formula is then of the form  $IRJ$  with  $I, J$  intervals and  $R \in \mathcal{A}$ .

Thus, for the example above,  $(INJ)$  would be represented by  $(I \{R, R'\} J)$ . Because we will be using Allen’s relations as logical relations in what follows, it is useful to notice the effects of Boolean operations on basic relations. The definition above states that Allen’s relations are disjunctions of basic relations. From mutual exclusiveness it follows that conjunctions of basic relations are false by definition (at most one relation can hold between any two intervals). For the negation of a basic relation  $R \in \mathcal{B}$  we obtain  $\neg R = \mathcal{B} \setminus \{R\}$ . Note that the negation is thus different from the inverse  $\bar{R}$ .

**Example 2.** COPD patients often have what is called ventilation-perfusion inequality – a mismatch between air flow and blood flow through the lung – which may develop during an exacerbation due to increased airway obstruction. When an exacerbation occurs we have a temporal event  $Vpi_I$  which is during, finishes or is overlapped by  $Exa_J$ . Without any further information the relation between ventilation-perfusion inequality and exacerbation can thus be described by:

$$\text{Vpi}_I(\bar{o}, d, f) \text{Exa}_J.$$

Given a set of interval relations, other temporal relations can be derived, using the transitivity of time ordering. For example if  $IbK$  and  $KbJ$ , then by transitivity it holds that  $IbJ$ . Similar rules can be constructed for all pairs of basic relations. In many cases the result will not be a basic relation however, but a set of possible relations, that is an element from  $\mathcal{A}$ . Allen [1] provides a table of all the inference rules.

More in general we can say that given a finite set of interval relations, the relations that are entailed by this set can be derived by repeated application of the following closure operator until a fixed point is reached [24]:

**Definition 2.** Let  $C$  be a finite set of interval relations. The closure operator  $\Gamma$  maps interval formulas to interval formulas using the

operations of inversion ( $\bar{\cdot}$ ), intersection ( $\cdot \cap \cdot$ ) and composition ( $\cdot \circ \cdot$ ).  $\Gamma(C)$  is the smallest set satisfying:

- 1  $C \subseteq \Gamma(C)$
- 2 For each  $I, J$  that appear in a formula in  $C : I \bar{B} J \in \Gamma(C)$
- 3 For each  $IRJ \in \Gamma(C) : J \bar{R} I \in \Gamma(C)$
- 4 For each  $(IRJ), (I'R'J) \in \Gamma(C) : I(R \cap R')J \in \Gamma(C)$
- 5 For each  $(IRK), (KR'J) \in \Gamma(C) : I(R \circ R')J \in \Gamma(C)$

The inversion operation is the inverse of each basic relation in  $R$ , the intersection operation is simply the set intersection of the relations  $R, R'$  and composition is the result of resolving the transitivity of the basic relations. Formally, composition is defined as  $\forall I, J : I(R \circ R')J \Leftrightarrow \exists K : IRK \wedge KR'J$ .

If there are multiple paths via which a relation between two intervals can be derived we are interested in the strongest relation that holds. So from  $\Gamma(C)$  we can derive the *reduced closure*  $\Gamma'(C)$ , where for each  $R$  with  $IRJ \in \Gamma(C)$  for given intervals  $I, J$  it holds that  $IR'J \in \Gamma'(C)$  if  $R' \subseteq R$ . In other words, the intersection of the relations between two intervals is contained in the reduced closure, because the intersection is the strongest relation that follows from  $C$ . See also [24].

**Example 3.** The temporal relation between the intervals of infection and exacerbation is:

$$\text{Inf}_I \circ \text{Exa}_J$$

which when combined with the relation,

$$\text{Vpi}_K(\bar{o}, d, f) \text{Exa}_J$$

can be used to infer the relation between ventilation-perfusion inequality and infection by application of the closure operator:

$$C = \{\text{Inf}_I \circ \text{Exa}_J, \text{Vpi}_K(\bar{o}, d, f) \text{Exa}_J\}$$

$$\Gamma'(C) = C \cup \{\text{Vpi}_K(\bar{b}, \bar{m}, \bar{o}, d, f) \text{Inf}_I, \text{Inf}_I(b, m, o, \bar{d}, \bar{f}) \text{Vpi}_K, \text{Exa}_J \bar{o} \text{Inf}_I, \text{Exa}_J(o, \bar{d}, \bar{f}) \text{Vpi}_K\}$$

#### 4.2. Logical reasoning with the interval algebra

As Allen showed [2], this qualitative algebra is well suited to reason about time in a logic context. Allen's relations are then represented by temporal predicates. The logic we will be using derives from the logic programming tradition of using Horn clauses,

$$H \leftarrow B_1, \dots, B_n$$

where  $H$  is the head of the clause,  $B_1, \dots, B_n$  the body and  $H$  and the  $B_i$ s are logical atoms. Variables are denoted with upper case and are implicitly universally quantified, conjunctions are denoted by commas ',' and a semicolon ';' denotes a disjunction, as in Prolog.

Also instead of using a reified logic approach as Allen does (i.e. using meta-predicates like HOLDS, OCCURS), we opt for the arguably simpler framework of temporal arguments [25] (see also [26] for a discussion on the (dis)advantages of both methods). This means that temporal predicates have a temporal argument specifying the relevant time interval. Note that this implies a typed logic with temporal and atemporal terms, which we will leave implicit as this can always be translated to first order logic at the cost of notational convenience.

A connection between an interval relation representation and predicate logic can be made by introducing a predicate that represents the temporal relation between intervals. Concretely, using a Prolog like syntax, we define the predicate  $r/3$ , with the first two arguments representing the intervals and the last argument a basic temporal relation; and the predicate  $\text{allen}/3$ , with two intervals and a list of basic relations as arguments such that  $\text{allen}(I, J, L)$  holds if the disjunction over elements  $X$  in list  $L$ ,  $\bigvee_{X \in L} r(I, J, X)$  holds.

**Example 4.** Consider again our COPD example which we can now represent somewhat more structured:

$$\text{exacerbation}(P, J) \leftarrow \text{patient}(P), \text{infection}(P, I), r(I, J, o).$$

$$\text{vpi}(P, J) \leftarrow \text{patient}(P), \text{exacerbation}(P, I), \text{allen}(I, J, [\bar{o}, d, f]).$$

Here  $\text{vpi}$  stands for ventilation-perfusion inequality. As the previous example showed the temporal relation between infection and  $\text{vpi}$  can be derived. In this case the result is a disjunction of five possible relations:

$$\text{vpi}(P, J) \leftarrow \text{patient}(P), \text{infection}(P, I), \text{allen}(I, J, [\bar{b}, \bar{m}, \bar{o}, d, f]).$$

#### 4.3. CP-logic

To represent and reason with probabilistic knowledge, we will use the probabilistic logic language CP-logic [4]. This language is based on the theory of logic programming extended with a probabilistic semantics. The main intuition is that probabilistic logic formulae represent causal rules, that is a logic clause gives a relation from some cause to a set of possible outcomes (each with some probability). A CP-logic program thus describes a causal process.

**Definition 3.** A causal probabilistic rule has the form:

$$(H_1 : \alpha_1); \dots; (H_n : \alpha_n) \leftarrow B$$

where  $\alpha_i$  is the (non-zero) probability of outcome  $H_i$  such that  $\sum_{i=0}^n \alpha_i \leq 1$ ;  $H_i$  are logical atoms and  $B$  is the body of the clause.

In other words, a causal rule gives a probability distribution over possible effects of some cause  $B$ . Deterministic effects can be modelled simply by a non-probabilistic logic formula  $H \leftarrow B$  which is equivalent to  $H : 1 \leftarrow B$ . Unconditional probabilistic rules can be seen as a prior distribution over logical facts. CP-logic is restricted to finite domains, so although one can write quantified formulae, these are interpreted as a finite set of ground instances.

The probabilistic semantics are based on the work by Shafer [27] on probability trees. The main idea behind this is that probabilistic processes are best described by a dynamic unfolding of events. Each node in the tree represents some state of the domain, transitions between nodes are probabilistic events as described by the causal probabilistic rules in the knowledge base, and each outgoing edge is some alternative outcome labelled with its probability. The leaves of a probability tree each describe a possible outcome of the events modelled in the knowledge base. All events are considered to be independent and dependencies have to be modelled explicitly in the rules. As each causal rule fires independently, it follows that the probability of a leaf node  $l$  is the product of the labels on the edges from  $l$  to the root of the tree. Since there may be multiple series of events that lead to the same final state the probability of an interpretation is the sum over all the leaves in the tree that share the same interpretation. Each rule is independent, which means that multiple rules with the same outcome are independent causes. In CP-logic, these independent causes are interpreted as a noisy-OR. See Vennekens et al. [4] for details and [28] for alternative interpretations.

**Example 5.** In Fig. 2 a CP-logic event tree is shown, representing the situation of whether a COPD patient suffers an exacerbation caused by either an infection or by breathing in a noxious substance. The tree follows from these CP-rules:

$$\text{exacerbation} : 0.6 \leftarrow \text{infection}.$$

$$\text{exacerbation} : 0.2 \leftarrow \text{noxious\_substance}.$$

$$\text{infection} : 0.05.$$

$$\text{noxious\_substance} : 0.01.$$

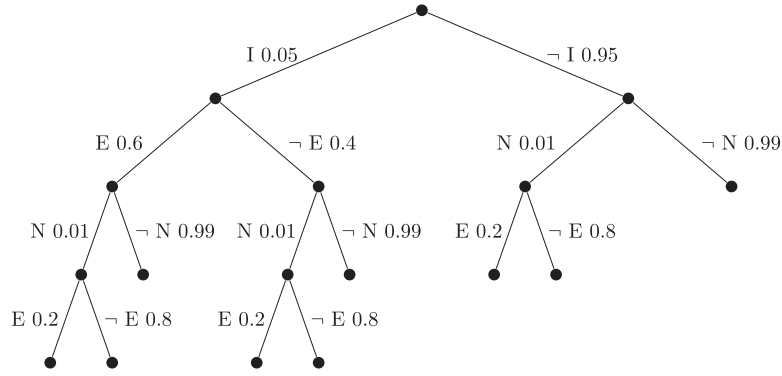


Fig. 2. A probability tree, where *I* is short for the *infection* event, *N* denotes *noxious substance* and *E* is *exacerbation*.

The probability of an exacerbation can be computed by summing over the leaves *l* that contain *E* (short for exacerbation) on the path from the root to *l*. Note that on some branches *E* occurs twice, which is allowed because there are two rules that have *E* as a consequence. The probability of, for example, the leftmost path is  $0.05 \cdot 0.6 \cdot 0.01 \cdot 0.2 = 0.00006$  and the probability of an exacerbation is:

$$0.00006 + 0.00024 + 0.0297 + 0.00004 + 0.0019 = 0.03194.$$

#### 4.4. Markov processes

Markov processes are commonly used as representation of uncertainty and time. The following definitions will be useful later on. Suppose we have a Markov process represented by the chain

$$X_1 \rightarrow X_2 \rightarrow \dots \rightarrow X_t \rightarrow \dots$$

The joint probability up to some point *T* is

$$\begin{aligned} P(X_{1:T}) &= P(X_T | X_{T-1}, X_{T-2}, \dots)P(X_{T-1}, X_{T-2}, \dots) \\ &= P(X_1) \prod_{t=2}^T P(X_t | X_{1:t}), \end{aligned}$$

where the notation  $1:t$  is an abbreviation of the sequence from 1 to *t*. The factorised notation is useful as we are primarily interested in Bayesian network representations. A common independence assumption is the first order Markov property  $P(X_{t+1} | X_{1:t}) = P(X_{t+1} | X_t)$ , which states that the future does not depend on the past given the present. The joint probability then simplifies to

$$P(X_{1:T}) = P(X_1) \prod_{t=2}^T P(X_t | X_{t-1}).$$

A generalisation of Markov chains is dynamic Bayesian networks (DBNs) that allow modelling independences in the state space description. The state space is then described by a graph  $G = (V, A)$  in which vertices *V* represent variables *X* at each time point, arcs *A* represent dependencies either within or between time slices and the joint probability of the process variables *X* factorises over the graph. The joint probability is given by:

$$\begin{aligned} P(X_{1:T}, 1:n) &= P(X_{1,1:n}) \prod_{t=2}^T P(X_{t,1:n} | X_{t-1,1:n}) \\ &= \prod_{t=1}^T \prod_{i=1}^n P(X_{t,i} | \text{pa}(X_{t,i})) \end{aligned}$$

Note that  $\text{pa}(X)$  denotes the parents in the graph of *X*, which can either be in the same time slice or in time slice  $t - 1$ , again assuming the first order Markov property.

It turns out that we can describe Markov processes in CP-logic as is shown by the work of Thon et al. [21] on CPT-L, which represents observable homogeneous Markov processes in CP-logic.

### 5. A probabilistic logic of qualitative time

When modelling real world situations, qualitative time is useful for those processes for which precise timing information is unavailable. However, it may be possible to obtain likelihood information, telling us that some event is more likely to happen at a particular time, even when the timing information is imprecise. This leads to temporal process descriptions – as represented with Allen’s logic – extended with probabilistic information. The expressive power of CP-logic will appear sufficient to act as a basis for such an extended, qualitative temporal and uncertain logic.

#### 5.1. A framework for uncertain temporal reasoning

To model uncertain processes we are primarily interested in the occurrence of events. In our context we consider events that are uniquely associated with intervals, and this association is expressed by means of a time-interval index. We interpret events as taking place throughout their associated intervals.

Now to define temporal events, we index facts with time intervals from the set of all time intervals  $\mathcal{I}$ .

**Definition 4.** Let  $\mathcal{E}$  denote the event space containing all probabilistic events of interest. A temporal uncertain event  $E_I$  is defined as a probabilistic event  $E \subseteq \mathcal{E}$  that occurs in time interval  $I \in \mathcal{I}$ .

The Boolean algebra of temporal events  $B(\mathcal{E}_{\mathcal{I}})$ , where  $\mathcal{E}_{\mathcal{I}}$  is defined as  $\mathcal{E}_{\mathcal{I}} = \{E_I | E \subseteq \mathcal{E}, I \in \mathcal{I}\}$ , should obey certain rules, taking into account the time interval indices of the events. The elements of the Boolean algebra are obtained by constructing conjunctions of events  $(E_I \wedge E'_I)$ , disjunctions  $(E_I \vee E'_I)$  and negations  $\neg E_I$ , with events  $E_I, E'_I \in \mathcal{E}_{\mathcal{I}}$ .

With all the basic ingredients defined above, it is now possible to define a framework for uncertain temporal reasoning.

**Definition 5.** A probabilistic qualitative time algebra PQT is defined as a 4-tuple  $PQT = \langle \mathcal{I}, \mathcal{E}_{\mathcal{I}}, \mathcal{B}, P \rangle$ , where

- $\mathcal{I}$  is the set of all time intervals;
- $\mathcal{E}_{\mathcal{I}}$  is the set of temporal uncertain events;
- $\mathcal{B}$  is the set of basic Allen relations;
- *P* is an associated joint probability distribution.

Below, we will use CP-logic as a practical language to implement the framework.

## 5.2. On events and intervals

There are certain properties of temporal events which require further attention, and which will be considered subsequently.

For  $E=E'$  with temporal events  $E_I$  and  $E'_J$ , there is an interaction between the Boolean operations on temporal events and the Allen relationships of the time intervals. For example, when  $lbj$  holds, then  $(E_I \wedge E'_J)$  cannot be simplified; however, if  $leqj$  holds, then  $(E_I \wedge E'_J) = E_I = E'_J$ . In addition, when two intervals  $I$  and  $J$  meet, i.e.  $lmj$  holds, then  $(E_I \wedge E'_J) = E_K$  with  $K=I \cup J$ . This expresses that event  $E$  actually occurred during interval  $K$ .

**Proposition 1.** For intervals  $I, J \in \mathcal{I}$  and a relation  $IRJ$  it holds that

$$(E_I \wedge E'_J) = E_{I \cup J} \quad (5.1)$$

if  $R \in \mathcal{B} \setminus \{b, \bar{b}\}$ .

In other words, it holds for all cases except when  $I$  and  $J$  are disconnected intervals.

**Proof.** The proposition can be understood by splitting the intervals into the intersection  $K=I \cap J$ , and the remainders  $I'=I \setminus K$  and  $J'=J \setminus K$ . Note that  $\{b, \bar{b}, m, \bar{m}\}$  are the only relations for which the intersection  $K$  is empty, but that for  $\{m, \bar{m}\}$  we have that  $\sup I = \inf J$  and hence  $I, J$  are connected. Then by realising that  $(I' \cup K) \cup (K \cup J') = I' \cup K \cup J'$  and that the same event in connected intervals cannot be distinguished, we obtain  $E_{I \cup J}$ .  $\square$

Note that Eq. (5.1) is true for any temporal relation  $R$  if we define the time-interval index  $I$  of a temporal event  $E_I$  as a set of time intervals, rather than as a single time interval, with the singleton set as a special case. This would generalise Allen's algebra of temporal relations to relations between sets of time intervals. As it is not our intention in this paper to change the algebraic basis of Allen's algebra, we will not pursue this idea further.

For a disjunction  $E_I \vee E'_J$ , we can apply the results from the proof of Proposition 1 to obtain that

$$(E_I \vee E'_J) = E_K \wedge (E_{I'} \vee E_{J'}) \quad (5.2)$$

where  $K, I'$  and  $J'$  are defined as in the proof above, for  $IRJ$ , with  $R \in \mathcal{B} \setminus \{b, \bar{b}\}$ .

## 5.3. Uncertainty

There are various ways in which uncertainty can be incorporated into the language:

- Uncertainty can be expressed with respect to the temporal events  $E_I$ .
- Uncertainty can be expressed with regard to the relation between time intervals  $IRJ$ .

When combined for two temporal events  $E_I$  and  $E'_J$ , these assumptions give rise to a joint probability distribution of the form

$$P(E_I, E'_J, IRJ),$$

with  $R$  an Allen relation. Note that this could also be the special case where we have a single event  $E$  associated with both interval  $I$  and interval  $J$ .

Often one is interested in more than two events, which can be modelled with a joint probability over all the events and pairwise temporal relations between the events. For example for three events we have the joint probability

$$P(E_I, E'_J, E''_K, IRJ, IR'K, JR''K).$$

Sometimes this joint probability will simplify due to independences between the events. In the following we focus on situations with

two events, yet similar arguments can also be made for multiple events.

### 5.3.1. Certain Allen relations

By conditioning on  $IRJ$ , one removes part of the uncertainty, yielding:

$$P(E_I, E'_J \mid IRJ).$$

In this case, only events are uncertain, and relations between interval are considered to be part of the (deterministic) logic specification. We can now define a distribution over temporal events.

**Definition 6.** The probability of a Boolean expression of events is given by the probability function  $P : \mathcal{B}(\mathcal{E}_{\mathcal{I}}) \rightarrow [0, 1]$ , where  $\mathcal{B}(\mathcal{E}_{\mathcal{I}})$  denotes the Boolean algebra over the set of temporal events  $\mathcal{E}_{\mathcal{I}}$ .

Here uncertainty is introduced at the level of events attached to particular intervals, which means we are looking at what happens at a greater level of detail than the qualitative relations describe. However, the nature of Allen's qualitative algebra dictates that relationships between the intervals are more important than the exact intervals. Consequently, the following parameter invariance appears appropriate to restrict arbitrary models to those models where the qualitative relations are the key primitives for temporal uncertainty:

$$\forall R, I, J, K, L : P(E_I, E'_J \mid IRJ) = P(E_K, E'_L \mid KRL), \quad (5.3)$$

i.e. when the Allen relation  $R$  between (potentially) different intervals is the same, the probability of joint occurrence of the associated temporal events is also the same. This invariance ensures that the influence of time on probabilities of events is governed by the temporal relations.

This invariance is actually related to a common assumption for dynamic Bayesian networks, where one assumes parameter invariance over time to limit the number of parameters. There is however a subtle difference. The invariance that is commonly assumed for DBNs is that  $P(X_t \mid X_{t-1})$  is equal for all  $t$ . The invariance in Eq. (5.3) however states that as long as two events are related through the same temporal relation their probabilities are equal, which is not about repetition invariance. Yet the relation with Markov processes is interesting and we will now show that the qualitative time framework can be used to model Markov processes.

The important properties of a Markov process as described in Section 4.4 for a representation in terms of Allen's relations are the ordering of events and the factorisation over time slices. Hence a chain of events, connected with the relations  $R_i \in \{b, m\}$  represents the correct ordering:

$$E_I R_1 E'_J R_2 E''_K R_3 \dots$$

Since this results in disjoint time slices, the factorisation property also holds under the first order Markov assumption.

To generalise to representing a DBN with Allen relations we have to somehow map events to time slices depending on the relations between them. A division in classes of relations turns out to be useful here:

$$\mathcal{B} = \mathcal{P} \cup \bar{\mathcal{P}} \cup \mathcal{C} \cup \mathcal{O}$$

where the subsets are the precedence relations  $\mathcal{P} = \{b, m\}$ , the inverse precedence relations  $\bar{\mathcal{P}} = \{\bar{b}, \bar{m}\}$ , the concurrent relations  $\mathcal{C} = \{s, \bar{s}, d, \bar{d}, f, \bar{f}, eq\}$  and the overlap relations  $\mathcal{O} = \{o, \bar{o}\}$ .

**Definition 7.** A temporal partition  $\Pi$  of a set of temporal events  $\mathcal{E}_{\mathcal{I}}$  is a partition such that for all sets  $L \in \Pi$  it holds that:  $E_I R E'_J$  with  $R \in \mathcal{C}$  iff  $E_I, E'_J \in L$ ; and  $E_I R E'_J$  with  $R \in \mathcal{P} \cup \bar{\mathcal{P}}$  iff  $E_I \in L, E'_J \in M, M \in \Pi$  and  $L \neq M$ .

First, we need to show that the uncertain temporal events can be linearly ordered. Recall that a linear order  $\leq$  is transitive, antisymmetric and total.

**Lemma 1.** *Let  $\mathcal{E}_T$  be a set of temporal events and  $\Pi$  a temporal partition of  $\mathcal{E}_T$ , then, the temporal events can be linearly ordered.*

**Proof.** The sets  $L \in \Pi$  are linearly ordered by the fact that the precedence relation induces a linear order  $\leq$ . Hence, temporal events in different sets are also linearly ordered. Events in each set  $L$  are equivalent as they are neither  $<$  nor  $>$ , showing antisymmetry. Between any two temporal events in a temporal partition a relation in  $\mathcal{B} \setminus \mathcal{O}$  holds, so totality follows.  $\square$

From the linear order that can be imposed on the temporal events it follows that one can map temporal events to a DBN.

**Proposition 2.** *A set of temporal events  $\mathcal{E}_T$  with a temporal partition  $\Pi$  can be mapped to a DBN with the first order Markov assumption.*

**Proof.** A DBN with the first order Markov assumption has arcs either within a time slice or from time slice  $t$  to  $t+1$ . The temporal partition results in a linear order by Lemma 1. Now let each  $L \in \Pi$  correspond to a time slice. Two time slices are connected by arcs in the DBN graph if  $L < M$  with  $L, M \in \Pi$ ,  $E_j R E_l$  holds with  $R \in \overline{\mathcal{P}}$ ,  $E_j \in M$  and  $E_l \in L$ , and there is no set  $N \in \Pi$  such that  $L < N < M$ .  $\square$

In other words, we can construct a first order Markov DBN for those temporal networks that only contain relations in  $\mathcal{C}$  and  $\mathcal{P}$ , resulting in intra and inter time-slice arcs respectively. If we also want to take into account the overlap relations  $\mathcal{I} \cup \mathcal{O}$  the construction is less straightforward. The overlapping events can be considered an indivisible unit, which allows us to assign the combined interval  $I \cup J$  to a single time slice. An undesirable consequence is that a chain of overlapping events would result in a single time slice for the complete chain. However we can give a restriction that allows us to deal with some overlapping events.

**Proposition 3.** *Let  $E_l \mathcal{O} E_j'$  with  $E_l, E_j' \in \mathcal{E}_T$ .  $\mathcal{E}_T$  corresponds to a DBN if for all  $E_K'' \in \mathcal{E}_T$  we have  $E_K'' R E_l$  and  $E_K'' R' E_j'$  with  $R, R' \in \mathcal{C}$ ; or  $R, R'$  either both in  $\mathcal{P}$  or both in  $\overline{\mathcal{P}}$ .*

**Proof.** If  $R, R' \in \mathcal{C}$  the joined interval  $I \cup J$  is also only related via  $\mathcal{C}$ , which by Proposition 2 results in a single time slice. If  $R, R' \in \mathcal{P}$  or  $R, R' \in \overline{\mathcal{P}}$  the joined interval will by the same proposition result in a separate time slice for the events.  $\square$

### 5.3.2. Uncertain Allen relations

The joint probability  $P(E_l, E_j, IRJ)$  represents the combined uncertainty in events and their temporal relations. Quantifying uncertainty in the relations by means of probability generalises the use of disjunctions in Allen's algebra to model for example uncertainty in the ordering of events. Conditioning the joint probability on the temporal relation, we obtain:

$$P(E_l, E_j, IRJ) = P(E_l, E_j | IRJ)P(IRJ).$$

For fixed intervals  $I, J$ , the following holds:

$$P(I\mathcal{B}J) = P\left(\bigvee_{r \in \mathcal{B}} IrJ\right) = \sum_{r \in \mathcal{B}} P(IrJ) = 1,$$

which is due to the mutual exclusivity of the basic relations. Thus, for any subset  $R \subseteq \mathcal{B}$  it holds that:

$$P(IRJ) = P\left(\bigvee_{r \in R} IrJ\right) = \sum_{r \in R} P(IrJ).$$

So although the set  $\mathcal{A}$  is large ( $|\mathcal{A}| = 2^{13}$ ) the probability of a relation  $R \in \mathcal{A}$  is simply the sum over the basic relations in  $R$ .

### Closure operations

When using probabilistic relations it is natural to think about whether the properties of Allen's algebra can be generalised. Returning for a moment to relations instead of a logic representation, we should reconsider the operations from Definition 2: *inverse*, *intersection* and *composition*. In [19] some methods were proposed, explained below. However, in our opinion the motivation is lacking and we propose an alternative for the probabilistic equivalent of the intersection operation.

We first introduce the notation  $P[I\mathcal{B}J]$  to define a distribution over the elements of  $\mathcal{B}$ . So for example if for  $IRJ$  we have  $P(IeqJ) = 0.3$ ,  $P(IbJ) = 0.7$  and  $P(IrJ) = 0$  for all  $r \in \mathcal{B} \setminus \{eq, b\}$ ; we write

$$P[I\mathcal{B}J] = [eq : 0.3, b : 0.7].$$

Now, let the following probability distribution for  $JR'K$  be given:

$$P[J\mathcal{B}K] = [d : 0.5, o : 0.5].$$

We can now define the probabilistic operations.

*Inverse.* The inverse of  $P[I\mathcal{B}J]$  is found by elementwise inversion  $P[J\mathcal{B}I] = [eq : 0.3, \bar{b} : 0.7]$ .

*Composition.* If we compute the composition  $I(R \circ R')K$  we obtain  $I\{bmosd\}K$ , for which we can compute the probabilities via the procedure in [19]

$$\begin{aligned} P(d) &= eq \circ d = 0.3 \cdot 0.5 = 0.15 \\ P(o) &= eq \circ o = 0.3 \cdot 0.5 = 0.15 \\ P(b) &= b \circ d = 0.7 \cdot 0.5 = 0.35 \\ P(bmosd) &= b \circ o = 0.7 \cdot 0.5 = 0.35 \end{aligned}$$

Note that  $P(Ir''K) = P(IrJ)P(Jr'K)$  with  $r'' = r \circ r'$ . The composition  $b \circ o$  results in multiple relations, which means that the probability 0.35 has to be divided over  $\{bmosd\}$ . A non-informative choice would be a uniform assignment, but in general we have

$$\begin{aligned} P(I\{b, o, d, (b \vee m \vee o \vee s \vee d)\}K) &= 1 \\ 0.15 &\leq P(IdK) \leq 0.15 + a \\ 0.15 &\leq P(IoK) \leq 0.15 + b \\ 0.35 &\leq P(IbK) \leq 0.35 + c \\ P(ImK) + P(IsK) + a + b + c &= 0.35 \end{aligned}$$

which implies  $a, c, b \in [0, 0.35]$ .

*Intersection.* For the intersection operation we first consider conditioning. Assume that we have the following probability distribution:

$$P[I\mathcal{B}K] = [m : 0.3, o : 0.7]$$

When we now also learn that  $I\{m, s\}K$  holds this means that according to Allen's algebra only  $ImK$  can hold as  $\{m\} = \{m, o\} \cap \{m, s\}$ . Probabilistically, this can be viewed as a conditioning event which leads to the conclusion that *meets* holds almost surely, as follows:

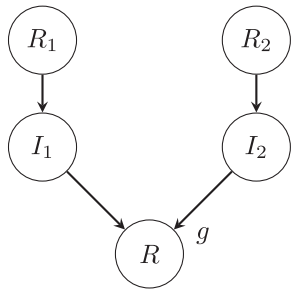
$$P(ImK | I\{m, s\}K) = \frac{P(ImK)}{P(I\{m, s\}K)} = \frac{P(ImK)}{P(ImK) + P(IsK)} = 1$$

as  $P(IsK) = 0$ , and

$$P(IoK | I\{m, s\}K) = \frac{P(\perp)}{P(I\{m, s\}K)} = 0$$

Here we assume that we know the initial distribution and obtain additional information on which we can condition. In this particular situation we can solve the problem and compute the probabilities as in the example. This is different from the approach taken in [19], where two distributions  $P_{R'}(IR'J)$  and  $P_{R''}(IR''J)$  are combined into a single distribution  $P(IRJ)$ :

$$\forall X \in R : P_R(IxJ) = \frac{1}{Z} \frac{P_{R'}(IxJ)P_{R''}(IxJ)}{P_{R'}(IxJ) + P_{R''}(IxJ)},$$



**Fig. 3.** Causal independence model to combine the distributions over the relations  $R_1, R_2$  via the interaction function  $g$ .

where  $Z$  is the normalisation  $Z = \sum_{z \in R} P_R(IzJ)$ . There appears to be little theoretical justification for this approach however. Instead, we opt to use the framework of causal independence, which offers a principled approach to model interactions between probabilistic events. Modelling the intersection operation then results in the graphical model shown in Fig. 3.

The behaviour of the causal independence model depends on the choice of interaction function,  $g$  in the figure. Using some variant of the noisy-AND appears appropriate, due to the similarity to the logical case. Since the relations are not binary valued the standard approach does not apply, but the generalisation by Srinivas [29] offers an alternative. The probability of  $R$  given  $R_1$  and  $R_2$  is given by:

$$P(R | R_1, R_2) = \sum_{(I_1, I_2 | g(I_1, I_2) = R)} \prod_i P(I_i | R_i)$$

where the probabilities  $P(I_i | R_i)$  are given by:

$$P(I_i = x | R_i = y) = \begin{cases} P(\text{pass}(I_i)) + P(\text{fail}(I_i = x)) & \text{if } x = y \\ P(\text{fail}(I_i = x)) & \text{otherwise} \end{cases}$$

where  $\text{fail}(I_i = x)$  indicates that  $I_i$  takes on the value  $x$  irrespective of the value of  $R_i$ , whereas  $\text{pass}(I_i)$  indicates that no failure occurs and hence

$$P(\text{pass}(I_i)) = 1 - \sum_{z \in \text{val}(R_i)} P(\text{fail}(I_i = z)).$$

For the interaction function  $g$  an analogue to the conjunction used in Allen's algebra is the componentwise AND-function

$$R = g(I_1, I_2) \quad \text{with} \quad R(i) = I_1(i) \wedge I_2(i),$$

where the components correspond to the basic relations  $\mathcal{B}$  and  $R(i)$  denotes the  $i$ th component of  $R$  and similarly for  $I$ . This then results in the probabilities

$$P(R = r) = P(g(I_1, I_2) = r) = P(I_1 = r)P(I_2 = r),$$

which is not a normalised distribution, so after normalisation we obtain

$$P(R = r) = \frac{1}{Z} P(g(I_1, I_2) = r) = \frac{1}{Z} P(I_1 = r)P(I_2 = r),$$

where  $Z = \sum_z P(I_1 = z)P(I_2 = z)$ .

#### 5.4. The framework in CP-logic

Probabilistic logic, and specifically CP-logic, offers a way to implement the framework described above. In CP-logic events are represented by facts, which are interpreted as independent events. Relations between facts are stated by logical expressions and through these expressions, dependences between events embedded in the facts can be introduced.

We can now define our language which we call *Qualitative Time CP-logic* as a restriction of CP-logic. The language consists of CP-rules (Definition 3), where the predicates model temporal probabilistic events (Definition 4). Furthermore, the special predicates  $r/3$  and  $\text{allen}/3$  (defined in Section 4.2), are used to denote the temporal relations between the events in each rule. We allow abbreviations where the omission of an interval means that the event is static over the time of interest and the omission of a relation in a rule indicates that  $\text{allen}(\cdot, \cdot, [\mathcal{B}])$  holds. These abbreviations allow us to incorporate specific static rules in standard CP-logic syntax. The probabilistic semantics are inherited from CP-logic, which leads to the result in the following section.

#### 5.5. Reasoning with temporal events

Now that we have defined uncertainty with respect to temporal events we can examine reasoning with uncertainty and time. To capture the qualitative nature of the time representation in logic, we represent intervals as atoms, foregoing the need to reason about the underlying time points.

The probability of temporally related events is of interest, which is, as we have seen, the probability  $P(E_i, E'_j, IRJ)$ . The following proposition shows how we can compute this probability via the CP-logic semantics, given that  $IRJ$  is a logical relation between the intervals associated with the events.

**Proposition 4.** *Let  $R$  be a relation  $R \in \mathcal{A}$ , and  $E_i, E'_j \in \mathcal{E}_T$  events; the probability  $P(E_i, E'_j, IRJ)$  is then obtained as follows:*

$$\begin{aligned} P(E_i, E'_j, IRJ) &= P(E_i | E'_j, IRJ)P(E'_j | IRJ)P(IRJ) \\ &= \sum_{\ell} \mathbb{I}_{\ell}(E_i, E'_j, IRJ) \prod_{e_{\ell}} P(e_{\ell}) \end{aligned}$$

where  $\ell$  denotes a leaf node in the tree,  $e_{\ell}$  is an edge on the path from  $\ell$  to the root of the tree; and  $\mathbb{I}_{\ell}(x)$  is an indicator function that is 1 if  $x$  is true in  $\ell$ .

**Proof.** (sketch) The proposition follows from the probability tree semantics of CP-logic. An event  $E_i$  is either true with a priori probability  $p$ , or is the consequent of a rule  $E_i : p \leftarrow B$ , with  $B$  the body of the clause. A rule decomposes the joint probability  $P(H, B)$  into  $P(H | B)P(B)$ . The probability of a particular event can hence be computed via a series of rule applications, which can be represented as an event tree. As CP-rules fire independently, we obtain the product from the proposition as a path in the tree. As there are multiple series of rules (logic proofs) that can lead to the same events, the total probability  $P(E_i, E'_j, IRJ)$  is the sum over those leaves  $\ell$  in which the events are true.  $\square$

The proof relies on properties of CP-logic, the details of which can be found in [4]. Note that the temporal relations can be seen as a special kind of event. We use them only as constraints in the body of rules, but this is not a technical requirement.

**Example 6.** We are again interested in modelling the relation between an infection and the occurrence of an exacerbation. Given the temporal relations *overlaps* and *equals* we now obtain the CP-logic rules:

```

exacerbation(j) : 0.7; ¬exacerbation(j) : 0.3 ← infection(i), allen(i, j, [o, eq]).
r(i, j, o) : 0.8; r(i, j, eq) : 0.2.
infection(i) : 0.01.
  
```

It follows that the probability of an exacerbation is a priori  $7 \times 10^{-3}$ , and if we would have (definite) evidence for an infection the probability would be 0.7.

The distribution over possible temporal relations in this case has the same result as when  $r(i, j, o)$  or  $r(i, j, eq)$  holds with probability 1,



since the same rule for exacerbation fires in both cases. It does show however that by specifying a distribution in this way the relations are mutually exclusive also in the reasoning process, which ensures that the right probabilities are computed, as given in Proposition 4.

### 5.6. Metric relations

A number of extensions to Allen's algebra have been proposed (e.g. [10,11]) based on various kinds of constraint reasoning. The added value lies mostly in metric constraints, which allow numerical constraints in addition to the pure qualitative relations in Allen's algebra. Jonsson and Bäckström [12] proposed the framework of disjunctive linear relations, which supports tractable reasoning with metric relations, while being at least as expressive as a number of earlier proposals (see also [30] on tractability results).

For practical medical applications metric constraints seem a useful addition to qualitative relations, enabling a more precise model of the temporal relation in those cases where the necessary level of quantitative detail is attainable. The essence of adding metric information is that additional constraints between interval endpoints should be satisfied, which can be modelled in the logic part of our specification. That is, special instances of temporal relations can be defined containing a metric constraint. The metric extension is useful to model relations that cannot be expressed within the original algebra, for instance a physical constraint that some effect occurs only after a certain amount of time.

Formally, for a qualitative temporal relation  $R \in \mathcal{B}$ , a quantified version  $R_m(x, y)$  of this relation has numerical parameters  $x, y$  that bound the distance between interval endpoints. Depending on the temporal relation, this definition can still be implemented in different ways. For instance a constraint like 'I is at least two time points before J' can be specified with a single quantitative parameter, leading to an alternative definition of *before*:

$$b_m(I, J, X, Y) \leftarrow I^+ + X < J^-,$$

meaning that I is at least X time points before J. Here we did not need the parameter Y, in practice different types of metric constraints might be useful that do require two parameters, like 'I between 2 and 4 time points before J':

$$b_m(I, J, X, Y) \leftarrow X \leq J^- - I^+ \leq Y.$$

With this parametrisation we can construct a quantified set of temporal relations on a level of detail that is appropriate for the domain.

To specify a distribution over relations, notice that the metric versions of the qualitative relations define a subset of the intervals, i.e. in terms of relations as sets  $b_m \subseteq b$ , hence the probability assigned to before can be distributed over the metric relations under the condition that they are disjoint. For example, if in the original distribution the probability of *before* was  $p$  and we want to introduce the relation 'at least two time points before',  $b_{>2}$ , with probability  $\frac{1}{2}p$ , it suffices to define  $b_{>2}$  and add the rule

$$I b_{>2} J : 0.5 \leftarrow I b J.$$

## 6. Validation of the qualitative probabilistic framework

In this section, we return to the two typical biomedical cases described in Section 2 and we use them here to validate the framework. We investigate whether the framework has sufficient expressive power to capture the biomedical knowledge concerning those cases. As the same framework is used to handle different cases, this shows that the framework is generic and has potential for dealing with a wide variety of other biomedical problems as well.

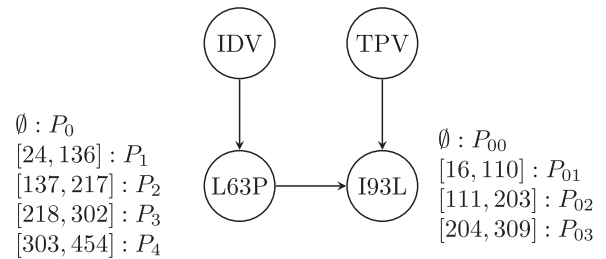


Fig. 4. Submodel of the TNBN from [7].

### 6.1. Temporal nodes Bayesian networks

We first study a particular class of temporal models called temporal nodes Bayesian networks (TNBN) [31], which represent event intervals as values of variables in a Bayesian network. This allows reasoning about events in those situations where one is not interested in dynamic changes over time but in single occurrences that are temporally uncertain. This class of models describes a worthwhile subset of the situations that we want to be able to represent.

In a recent paper [7] a TNBN was used to model interactions between antiretroviral drugs and mutations of the Human Immunodeficiency Virus, HIV, that result in drug resistance. The temporally uncertain aspect is here the occurrence of the mutations given other mutations and an antiretroviral drug or a cocktail of those drugs (which present an evolutionary pressure for particular mutations to succeed). The presence of drugs is modelled atemporally, while each mutation has a probability of occurrence in each of a set of consecutive intervals.

To represent this in our logic we first need to represent the atemporal information, which consists of a set of ground clauses of the type:

$$\text{anti\_viral}(X) : P.$$

indicating that a particular drug X occurs with probability P, although when studying a particular situation it will typically be known which drugs are present, allowing us to instantiate these clauses to either true or false. Each mutation is influenced by other mutations and drugs resulting in rules of the type:

$$\text{mutation}(X, \emptyset) : P_0; \text{mutation}(X, i_1) : P_1; \text{mutation}(X, i_2) : P_2;$$

$$\dots; \text{mutation}(X, i_n) : P_n \leftarrow \text{mutation}(Y, j), \dots, \text{anti\_viral}(Z), \dots$$

where  $i_1, i_2, \dots, i_n, j$  are ground intervals,  $\emptyset$  indicates the event did not occur and  $\sum_{k=0}^n P_k = 1$  holds as usual. In terms of Allen's relations we have the constraint

$$i_1 m i_2 m \dots m i_n,$$

for all temporal variables. However, because each variable is allowed to take on different interval values, interactions between mutations result in other temporal relations. In TNBNs these relations are modelled implicitly in the probability distributions. That is, if  $\text{mutation}(X, I) \leftarrow \text{mutation}(Y, J)$ , the  $n \cdot (m + 1)$  parameters for  $i_k, j_l$  with  $k \in \{0, \dots, n\}, l \in \{0, \dots, m\}$  encode the Cartesian product of possible combinations of I and J.

Concretely, we model the two most common mutations, L63P and I93L, and two drugs, IDV and TPV, that influence them. This is

a submodel of one of the networks described in [7], reproduced in Fig. 4. This results in the following rules:

anti\_viral(IDV).  
 anti\_viral(TPV).  
 mutation(L63P,  $\emptyset$ ) :  $P_0$ ; mutation(L63P, [24, 136]) :  $P_1$ ;  
 mutation(L63P, [137, 217]) :  $P_2$ ; mutation(L63P, [218, 302]) :  $P_3$ ;  
 mutation(L63P, [303, 454]) :  $P_4 \leftarrow$  anti\_viral(IDV).  
 mutation(I93L,  $\emptyset$ ) :  $P_{00}$ ; mutation(I93L, [16, 110]) :  $P_{01}$ ;  
 mutation(I93L, [111, 203]) :  $P_{02}$ ; mutation(I93L, [204, 309]) :  $P_{03}$ ;  
 $\leftarrow$  anti\_viral(TPV), mutation(L63P,  $\emptyset$ ).

where for brevity we have omitted the rules for the other parent configurations of I93L which are analogous to the one shown. It is easy to see that the probability of for example mutation(I93L,  $\emptyset$ ) can be computed as:

$$\begin{aligned} P(\text{mutation(I93L, } \emptyset)) &= \sum_I P(\text{mutation(L63P, } I))P(\text{mutation(I93L, } \emptyset) \mid \text{mutation(L63P, } I)) \\ &= \sum_{i=0}^4 P_i P_{i0}. \end{aligned}$$

This is the expected result of the CP-logic semantics and of TNBNs, and indeed our rule representation is equivalent to the TNBN.

With our representation we can also quite simply introduce extensions with little additional work, as the framework takes care of the probabilistic reasoning. For example it may be interesting to make the temporal representation more qualitative, instead of the concrete intervals used now. To do so we can model the interaction between mutations by means of a temporal relation, and especially *before* is a natural candidate. For instance, the influence of L63P on I93L can be modelled as two rules, for the case that L63P does not occur (not shown) and for the case that L63P occurs before I93L:

mutation(I93L,  $\emptyset$ ) :  $P_{10}$ ; mutation(I93L,  $J$ ) :  $P_{11}$ ;  
 $\leftarrow$  anti\_viral(TPV), mutation(L63P,  $I$ ),  $r(I, J, b)$ .

Another possible extension would be to also take into account time for the antiviral drugs. It appears plausible that the influence of a drug is concentrated in the interval in which it is present, with possibly limited influence some time longer due to the fact that it changed the population of viruses. We can capture this idea in the rules:

anti\_viral(IDV,  $I$ ).  
 mutation(L63P,  $\emptyset$ ) :  $P_0$ ; mutation(L63P,  $J$ ) :  $P_1$ ;  
 $\leftarrow$  anti\_viral(IDV,  $I$ ),  $r(I, J, eq)$ .  
 mutation(L63P,  $\emptyset$ ) :  $P'_0$ ; mutation(L63P,  $J$ ) :  $P'_1$ ;  
 $\leftarrow$  anti\_viral(IDV,  $I$ ),  $r(I, J, m_x)$ .

where the relation  $m_x$  is a metric version of the *meets* relation, defined as  $Im_x J \leftarrow I^* = J^- \wedge J^* - J^- \leq x$ , allowing us to limit the time of influence of the drug after it is no longer administered.

The ease with which we can incorporate ideas for extensions in the language shows the versatility of our framework. The logic offers a readable language to represent relations both atemporal and temporal which together with the probabilistic semantics results in a compelling framework to represent many clinical

situations. We now return to the problem of modelling the disease process of COPD.

## 6.2. Describing the evolution of COPD

Our framework is particularly well-suited to the clinical task of monitoring patients with chronic obstructive pulmonary disease. As the previous examples have shown both the time and uncertainty aspects play a role in COPD management. In this section we develop a model of COPD disease progression in more detail.

We start with a clinical description of influence between the relevant variables and their qualitative temporal relations. Relevant events to construct a process description of COPD progression are the outcome variable *exacerbation* its usual cause *infection* and the observable symptoms, signs and lab data. In the context of monitoring patients at home, access to signs or lab data is limited and we focus on symptoms, specifically dyspnea (difficulty in breathing), sputum colour and volume, cough and reduced activity. The patients from whom the data have been acquired typically have around two exacerbations per year, and consequently exacerbation symptoms will be largely absent most of the time. The situations we are interested in are the episodes in which an exacerbation occurs. When an exacerbation develops, different symptoms will increase over time until either through medication or through natural recovery the symptoms return to normal levels.

After having identified the variables our logic representation makes it easy to model explicitly the structural relations in the domain. That is, we construct rules of the type:

symptom( $P, S$ )  $\leftarrow$  infection( $P, Y$ ), symptom( $S$ ), patient( $P$ ).  
 symptom(*dyspnea*).  
 symptom(*cough*).  
 symptom(*sputum*).  
 symptom(*activity*).

This indicates that a patient  $P$  with an infection  $Y$  (which would allow us to specify different kinds of infections, e.g. bacterial or viral; or even more detailed by identifying specific bacteria that often cause airway infections like *H. Influenzae*, or *S. Pneumoniae*, but here we limit ourselves to just modelling the presence of an infection), has symptoms from the set {*dyspnea*, *cough*, *sputum*, *activity*}. In addition we should take into account interdependencies of the symptoms, for example, that an increase in dyspnea can cause a decrease of activity (the fact that some symptoms indicate an increase and others a decrease can easily be represented explicitly in the logic but is omitted for simplicity):

symptom( $P, activity$ )  $\leftarrow$  symptom( $P, dyspnea$ ), patient( $P$ ).

The resulting rules are quite readable, and both obtaining and verifying the domain knowledge with help from an expert is a realistic option.

Given the structural properties the next step is modelling the temporal characteristics. It is unrealistic to have precise timing information as this depends on many patient specific and environmental factors. Allen's algebra is a useful compromise between unattainable precise temporal knowledge and disregarding the temporal dimension of the problem completely. Concretely, we can describe the temporal structure of the process by assigning relations between infection events, exacerbation events and symptoms, for example in a substantial number of cases the first symptom to increase is dyspnea, which often stays present until the exacerbation ends. So the most likely temporal relation would be:

infection( $\bar{d}o$ )dyspnea( $\bar{f}o$ )exacerbation.

Similarly we can use domain knowledge to define relations for the other symptoms with respect to infection and exacerbation events. Analogous to the structural relations, we can also take into account the temporal relations between symptoms. It should be noted however that if the desired information is hard to obtain, this just means that we have no explicit constraint on the temporal relation between events, which means that all basic relations are possible.

In general it will often not be easy to specify structural and temporal characteristics exactly, requiring the inclusion of uncertain information. The advantage of starting with logic (without probabilities) is that the main structure of the process can be determined more easily. The choice for a probabilistic logic like CP-logic allows us to subsequently refine the representation within the same framework to include uncertainty. If clinical data is available, it can be applied in this stage to estimate probabilities.

The temporal relations may also be underspecified which can be modelled with a distribution over the basic temporal relations. For example derived from relative frequencies in the data and a knowledge-based prior distribution. We then might obtain for example the following distribution for the relation between an exacerbation and dyspnea – which in CP-logic notation consists of a set of statements like  $r(I, J, X) : P$ , where  $X$  is a temporal relation and  $P$  a probability – shown here more compactly:

$$\begin{matrix} b & 0 & m & 0 & o & 0 & s & 0.01 & d & 0.05 & f & 0.35 & eq & 0.18 \\ \bar{b} & 0 & \bar{m} & 0 & \bar{o} & 0.2 & \bar{s} & 0.05 & \bar{d} & 0.1 & \bar{f} & 0.05 \end{matrix} \quad (6.1)$$

The relation *before* and *meets* have been assigned probability zero, because we are only interested in symptoms that actually co-occur with an exacerbation. And *overlaps* has zero probability since it makes little sense to consider exacerbations that start before their symptoms.

Putting it all together we end up with a representation that includes structural, temporal and uncertainty information. To give an example of the reasoning in our COPD model we restrict ourselves to the symptoms *dyspnea* and *cough*, their cause *infection* and the effect *exacerbation*. We assume the temporal relation *equals* for the relation between the symptoms and infection. Then using the predicates we introduced in Section 4.2 and somewhat more compact notation for readability (omitting the specification of the patient  $P$ , writing symptoms as predicates and some set-notation in 6.5), the rules to model this process are the following:

$$dyspnea(I) : 0.8 \leftarrow infection(J), r(I, J, eq). \quad (6.2)$$

$$cough(I) : 0.6 \leftarrow infection(J), r(I, J, \bar{d}). \quad (6.3)$$

$$\begin{matrix} exacerbation(I) : 0.95 \\ \leftarrow dyspnea(J), allen(I, J, [\bar{o}, f]), cough(K), r(I, K, X). \end{matrix} \quad (6.4)$$

$$\begin{matrix} exacerbation(I) : 0.7 \\ \leftarrow dyspnea(J), allen(I, J, [C \setminus \{f\}]), cough(K), r(I, K, X). \end{matrix} \quad (6.5)$$

$$infection(I) : 0.05 \leftarrow infection(J), r(I, J, \bar{b}). \quad (6.6)$$

$$infection(i1) : 0.05. \quad (6.7)$$

Both exacerbations and individual symptoms are recurring events but the start event of a series of occurrences is the only event for which we need to model the repetition. This is expressed by rule 6.6. To predict the probability of an exacerbation in a particular interval, we need probabilities for the dependencies between exacerbations and symptoms given a temporal relation. Instead of

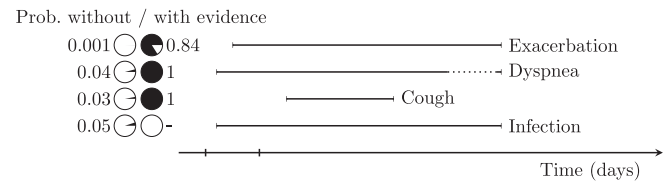


Fig. 5. Graphical representation of the example time-course and probabilities. The proportion of the circle that is shaded indicates the probability.

exhaustively enumerating all relations, we identify the dominant pattern: exacerbation is *overlapped-by* or *finishes* dyspnea. The only probabilities that have to be specified are those for the dominant temporal pattern (rule 6.4) and for the case summarising the other situations (rule 6.5). Instantiations of the relation predicates for specific intervals are omitted for conciseness.

With this representation we can now query probabilities of interest given evidence (observations). For example, let us assume the relations between infection and the symptoms,  $r(dyspnea, infection, eq)$  and  $r(cough, infection, \bar{d})$  hold. The distribution for  $r(exacerbation, cough, X)$  can be derived from  $r(dyspnea, infection, eq)$ ,  $r(infection, cough, d)$  and  $allen(dyspnea, exacerbation, [C \cup \{\bar{o}\}])$  leading to the relation  $allen(exacerbation, cough, [C \cup \emptyset])$ , with the distribution:

$$\begin{matrix} o & 0.009 & s & 0.006 & d & 0.006 & f & 0.006 & eq & 0.006 \\ \bar{o} & 0.189 & \bar{s} & 0.189 & \bar{d} & 0.572 & \bar{f} & 0.009 \end{matrix} \quad (6.8)$$

resulting from the procedure described in 5.3.2 and using the uniform distribution to split up probabilities when the composition resulted in multiple basic relations. To compute  $P(exacerbation)$  the relation is marginalised, but it nonetheless provides the insight that the probability that an exacerbation contains an episode of coughing is high.

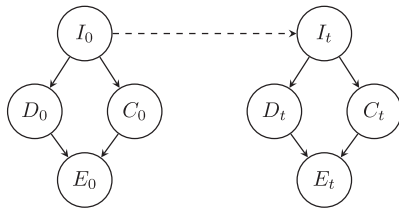
Computing  $P(exacerbation)$  from the rules and distributions above, we obtain  $P(exacerbation) = .001$ , which is dominated by the small prior probability of an infection. Then, if we have evidence (from the patient) that dyspnea and cough are present, we obtain  $P(exacerbation) = 0.84$ . See also Fig. 5.

### 6.2.1. Dynamic Bayesian network

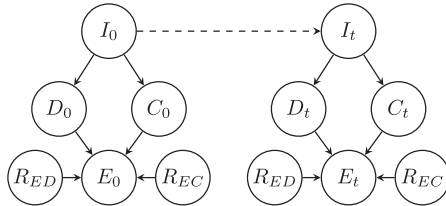
To show the different aspects of the language we can compare the situation above to a model based on a dynamic Bayesian network. As we have seen we can convert a model based on Allen's algebra to a DBN using the procedure in Section 5.3.1.

First we notice that possible infections are the repeating events (with relation *before*) that define the time slices of the DBN. Inspecting the rules above we see that the symptoms are contemporary with the infection event, resulting in connections within the time slice. Both observations are a direct consequence of Proposition 2. However, we also have that the relation between *dyspnea* and *exacerbation* is *overlaps* with probability 0.2. In Proposition 3 we gave a characterisation of situations for which we can construct a DBN. In the current situation it means that the exacerbation and its symptoms would have to be modelled as interacting in the same time slice. This seems a reasonable solution, resulting in the DBN structure shown in Fig. 6. Note that a consequence of this choice is that we lose the possibility to make a distinction between the start of the symptoms (dyspnea, cough) and the exacerbation. This shows that although a DBN can be constructed for this situation, the representation with Allen's algebra is more expressive.

With respect to the probabilistic parameters we can reuse the parameters from the logic representation. We did not display the rules that govern the probability of symptoms without an infection, which are needed to complete a conditional probability table, but they have a similar pattern as the rules shown. The



**Fig. 6.** Structure of the COPD-monitoring DBN. Dashed arcs represent inter time-slice connections. I = infection, C = cough, D = dyspnea, E = exacerbation.



**Fig. 7.** Structure of the COPD-monitoring DBN with variables for the temporal relations. Dashed arcs represent inter time-slice connections. I = infection, C = cough, D = dyspnea, E = exacerbation.

uncertainty on the relations is however less easy to represent: since we aggregate all concurrent relations, an obvious solution would be to marginalise out the relations. Taking as an example the probability of *exacerbation* given that *dyspnea* and *cough* are true, this results in the parameter estimate (omitting the conditioning on  $D$ ,  $C$  on the right for readability):

$$\begin{aligned}
 P(E \mid D, C) &= \sum_{R_{ED}, R_{EC}} P(E, R_{ED}, R_{EC}) \\
 &= \sum_{R_{ED}, R_{EC}} P(E, R_{ED} \mid R_{EC}) P(R_{EC}) \\
 &= \sum_{R_{ED}=\{\bar{0}, f\}} P(E, R_{ED}) + \sum_{R_{ED}=\{C \setminus f\}} P(E, R_{ED}) \\
 &= \sum_{R_{ED}=\{\bar{0}, f\}} P(E \mid R_{ED}) P(R_{ED}) + \sum_{R_{ED}=\{C \setminus f\}} P(E \mid R_{ED}) P(R_{ED}) \\
 &= 0.95 \cdot 0.55 + 0.7 \cdot 0.45
 \end{aligned}$$

The parameter estimates for the other cases are similar.

Alternatively we can explicitly model the relations within the time slice by introducing extra variables that take the relations as values. To maintain the temporal semantics of the DBN we again only consider relations in  $C$  within a time slice. The network from Fig. 6 can then be extended to the one in Fig. 7. Since we assumed earlier that the distribution over temporal relations is independent of the values of the (event) variables involved, the relation variables do not have parents in the graph. The parameters of the relation variables are equal to the distribution in (6.1). For *exacerbation*, in this particular example  $R_{EC}$  has no influence and we can encode the joint distribution of *dyspnea*, *cough* and  $R_{ED}$  easily in the CPT. We obtain  $P(E \mid D = 1, C = 1, R_{ED} = r) = p$  where  $p = 0.95$  when  $r \in \{\bar{0}, f\}$  and  $p = 0.7$  when  $r \in \{C \setminus f\}$ . This follows directly from the rules we gave earlier.

Summarising, we can see that comparing the two representations the DBN has a straightforward interpretation in terms of events that happen at the same time and those that do not. Allen's algebra allows a richer temporal structure without enforcing too much temporal detail. The logic representation can be used when only ordering information is available and allows us to model both structural and temporal information. Furthermore, since it is easy to define new relations for a particular domain situation, much more

complex relations can also be specified easily. In order to gain the same level of detail in a DBN representation one has to introduce additional variables, losing some of the convenience of the compact representation of DBNs.

The best choice of temporal model is a balancing act where it depends on the importance and availability of temporal information whether a DBN or a logic representation has more value. In the case of COPD monitoring, a DBN is useful as a simple classifier that takes some temporal information into account. However, to gain a better understanding of the nature of the disease process it is worthwhile to use a richer representation that models different kinds of temporal relations between symptoms and between a symptom and an exacerbation. Using qualitative relations we can model the additional temporal structure, without requiring unrealistic temporal detail.

## 7. Conclusion

In this paper we reformulated and extended Allen's algebra in a probabilistic logic, leading to a framework called *Qualitative Time CP-logic*. Medical processes are often well suited to be represented in a qualitative time framework, and the representation explored in this paper appears to capture knowledge that fits clinical reasoning well. Specifically, the framework allows us to construct models stepwise, starting with structural properties represented by logic, adding qualitative temporal information and finally representing uncertainty with probabilities. Compared to directly constructing a dynamic Bayesian network, we gain the expressiveness of probabilistic logics in general and a more intuitive time representation. If desired, the model can still be converted to a DBN, but this requires some assumptions on how to simplify the temporal information represented with Allen's algebra. Our COPD model indicates that we can specify clinically relevant processes in an intuitive fashion within our framework, while the generic nature of the framework implies that we can also deal with many other biomedical situations.

## References

- [1] Allen J. Maintaining knowledge about temporal intervals. *Communications of the ACM* 1983;26(11):832–43.
- [2] Allen J. Towards a general theory of action and time. *Artificial Intelligence* 1984;23(2):123–54.
- [3] Shahar Y. A framework for knowledge-based temporal abstraction. *Artificial Intelligence* 1997;90(1–2):79–133.
- [4] Vennekens J, Denecker M, Bruynooghe M. CP-logic: a language of causal probabilistic events and its relation to logic programming. *Theory and Practice of Logic Programming* 2009;9(3):245–308.
- [5] van der Heijden M, Lucas P, Lijnse B, Heijdra Y, Schermer T. An autonomous mobile system for the management of COPD. *Journal of Biomedical Informatics* 2013;46(3):458–69.
- [6] Combi C, Keravnou E, Shahar Y. *Temporal information systems in medicine*. New York, NY, USA: Springer Publishing Company Incorporated; 2010.
- [7] Hernandez-Leal P, Rios-Flores A, Ávila-Rios S, Reyes-Terán G, González J, Orihuela-Espina F, et al. Discovering human immunodeficiency virus mutational pathways using temporal Bayesian networks. *Artificial Intelligence in Medicine* 2013;57(3):185–95.
- [8] McDermott D. A temporal logic for reasoning about processes and plans. *Cognitive Science* 1982;6(2):101–55.
- [9] Combi C, Pinciroli F, Pozzi G. Managing different time granularities of clinical information by an interval-based temporal data model. *Methods of Information in Medicine* 1995;34(5):458–74.
- [10] Dechter R, Meiri I, Pearl J. Temporal constraint networks. *Artificial Intelligence* 1991;49(1–3):61–95.
- [11] Kautz H, Ladkin P. Integrating metric and qualitative temporal reasoning. In: Dean T, McKeown K, editors. *AAAI'91: proceedings of the 9th national conference of the American Association for Artificial Intelligence*. AAAI Press; 1991. p. 241–6.
- [12] Jonsson P, Bäckström C. A unifying approach to temporal constraint reasoning. *Artificial Intelligence* 1998;102(1):143–55.
- [13] Terenziani P, Carlini C, Montani S. Towards a comprehensive treatment of temporal constraint in clinical guidelines. In: Artale A, Fisher M, editors. *TIME'02: proceedings of the 9th international symposium on temporal representation and reasoning*. Los Alamitos, CA, USA: IEEE Computer Society; 2002. p. 20–7.

- [14] Dean T, Kanazawa K. A model for reasoning about persistence and causation. *Computational Intelligence* 1989;5(3):142–50.
- [15] Murphy KP. *Dynamic Bayesian networks: representation, inference and learning*. Berkeley, USA: University of California; 2002 [Ph.D. thesis].
- [16] Galán S, Aguado F, Díez F, Mira J. NasoNet, modelling the spread of nasopharyngeal cancer with networks of probabilistic events in discrete time. *Artificial Intelligence in Medicine* 2002;25(3):247–64.
- [17] Arroyo-Figueroa G, Sucar L. Temporal Bayesian networks of events for diagnosis and prediction in dynamic domains. *Applied Intelligence* 2005;23(2):77–86.
- [18] Tawfik A, Neufeld E. Temporal reasoning and Bayesian networks. *Computational Intelligence* 2000;16(3):349–77.
- [19] Ryabov V, Trudel A. Probabilistic temporal interval networks. In: Combi C, Ligozat G, editors. *TIME'04: proceedings of the 11th international symposium on temporal representation and reasoning*. Los Alamitos, CA, USA: IEEE Computer Society; 2004. p. 64–7.
- [20] Santos Jr E, Young J. Probabilistic temporal networks: a unified framework for reasoning with time and uncertainty. *International Journal of Approximate Reasoning* 1999;20(3):263–91.
- [21] Thon I, Landwehr N, de Raedt L. Stochastic relational processes: efficient inference and applications. *Machine Learning* 2011;82(2):239–72.
- [22] Gutmann B, Thon I, Kimmig A, Bruynooghe M, de Raedt L. The magic of logical inference in probabilistic programming. *Theory and Practice of Logic Programming* 2011;11:663–80.
- [23] Sato T. A statistical learning method for logic programs with distribution semantics. In: Sterling L, editor. *ICLP'95: proceedings of the 12th international conference on logic programming*. Cambridge, MA, USA: The MIT Press; 1995. p. 715–29.
- [24] Nebel B, Bürckert H-J. Reasoning about temporal relations: a maximal tractable subclass of Allen's interval algebra. *Journal of the ACM* 1995;42(1):43–66.
- [25] Bacchus F, Tenenber J, Koomen J. A non-reified temporal logic. *Artificial Intelligence* 1991;52(1):87–108.
- [26] Vila L. A survey on temporal reasoning in artificial intelligence. *AI Communications* 1994;7(1):4–28.
- [27] Shafer G. *The art of causal conjecture*. Cambridge, MA, USA: The MIT Press; 1996.
- [28] Hommersom A, Lucas P. Generalising the interaction rules in probabilistic logic. In: Walsh T, editor. *IJCAI'11: proceedings of the 22nd international joint conference on artificial intelligence*. AAAI Press; 2011. p. 912–7.
- [29] Srinivas S. A generalization of the noisy-or model. In: Heckerman D, Mamdani E, editors. *UAI'93: proceedings of the 9th conference on uncertainty in artificial intelligence*. San Francisco, CA, USA: Morgan Kaufmann; 1993. p. 208–18.
- [30] Krokchin A, Jeavons P, Jonsson P. Reasoning about temporal relations: the tractable subalgebras of Allen's interval algebra. *Journal of the ACM* 2003;50(5):591–640.
- [31] Arroyo-Figueroa G, Sucar L. A temporal Bayesian network for diagnosis and prediction. In: Laskey K, Prade H, editors. *UAI'99: proceedings of the 15th conference on uncertainty in artificial intelligence*. San Francisco, CA, USA: Morgan Kaufmann; 1999. p. 13–20.