A Probabilistic Logic of Qualitative Time

Maarten van der Heijden[1,2] and Peter J.F. Lucas[2]

[1] Institute for Computing and Information Sciences,
Radboud University Nijmegen, The Netherlands
[2] Department of Primary and Community Care,
Radboud University Nijmegen Medical Centre, The Netherlands
{m.vanderheijden, peterl}@cs.ru.nl

Abstract. Representing and reasoning over dynamic processes can benefit from using a qualitative representation of time when precise timing information is unavailable. Allen's interval algebra provides a framework to do so. However, a more precise model can be obtained by modelling uncertainty in the process. In this paper we propose an extension of Allen's algebra with uncertainty, using probability theory. Specifically, we use the expressive power of CP-logic, a probabilistic logic, to represent and reason with uncertain knowledge in combination with Allen's interval algebra to reason qualitatively about time.

1 Introduction

In solving problems, one often has to take into account the time when a particular event has occurred or is expected to occur. Typically, the actual temporal details about when events have occurred are not available, or at least imprecise, whereas one is more certain about the actual order of the events. Medicine is a field where much of the information about patients has such a temporal, yet imprecise dimension. AI researchers have traditionally used Allen's interval algebra [1] to model situations where there is much imprecision about the temporal evolution of events. Although Allen's algebra supports qualitative reasoning about time, it does no allow expressing uncertainty about the qualitative, temporal relationships. Yet, uncertainty is a typical characteristics of many problems where precise temporal information is missing; medicine can again be taken as a prototypical domain for which this is true. Work by Shahar [2] clearly indicates the usefulness of Allen's algebra for describing temporal events in medicine, and provides a use case in the form of temporal abstraction. For example, in hospital intensive care, interpreting the large amounts of (temporal) data becomes more manageable if we abstract from individual time points to a more qualitative representation. In this paper we aim to develop a happy marriage between Allen's interval logic and uncertainty reasoning by making use of probabilistic logic as a unifying formalism.

Frameworks that combine logic and uncertainty have garnered quite some attention in the past few years. Specifically, various authors have proposed probabilistic logics, combining the expressive power of (a subset of) predicate logic representations with probabilistic uncertainty calculus. Examples of such temporal logics include probabilistic Horn abduction [3], Bayesian logic programs [4], and more recently ProbLog [5] and CP-logic [6], among others. Many of these probabilistic logics could serve as

the basis for an extension of Allen's algebra; we argue that at a conceptual level CP-logic is already well aligned and is thus a natural choice. Allen [7] provides a temporal logic based on his interval algebra to model processes over time. We want to be able to reason about when certain events happen and how they relate to other events. It is then quite reasonable to take a causal viewpoint, as time and causality are closely related – causes precede their effects – and describing a temporal process as a causal mechanism seems an intuitive representation. CP-logic is short for 'causal probabilistic logic' and its semantics favours descriptions from a causal viewpoint, which meshes well with process descriptions of the kind you would want to specify in Allen's logic. So if we are able to reason with Allen's interval algebra within CP-logic, we obtain something that is conceptually pleasing while being more expressive than Allen's logic.

Throughout the paper we will use a medical example to illustrate the developed concepts, because as mentioned, clinical medicine is a typical environment in which uncertainty and time play an important role. Specifically, we look at examples related to chronic obstructive pulmonary disease, COPD for short, and related lung diseases and complications. COPD is a progressive lung disease which is 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, (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. In modelling these kinds of situations many factors are uncertain. Often you do not know whether, for example, an exacerbation will occur, and even if you do you may not know when exactly. Allen's algebra consists of qualitative relations that partially model the temporal uncertainty, yet Allen also recognised that modelling interesting processes that develop through time we need more than just temporal relations. The logic he proposed [7] combines the temporal relations with operators akin to predicate logic in order to state structural properties about the domain. Basically it is logic that provides the expressiveness to model things like causal connections, while the interval relations express temporal information and at least some of the uncertainty involved. However, the uncertainties of, for instance, predicting whether an exacerbation will occur given some observations of patient symptoms related in time, requires more extensive modelling capabilities. Advances in probabilistic logic provide us with tools that might help in modelling these uncertainties.

This paper is organised as follows. First, we go over some preliminaries, specifically, reviewing Allen's temporal algebra somewhat more formally in Section 2.1 and 2.2, followed in Section 2.3 by a description of CP-logic, the probabilistic logic we use. Then in Section 3 we describe the extension to probabilistic temporal logic.

2 Preliminaries

2.1 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. When specifying an

event, the interval when the event happens is made explicit. To start, we will review the various possible relations between time intervals. Later we will discuss how such intervals, and their relationships, can be used to specify the evolution of events.

Two intervals can have a number of qualitative relations, some event can for example happen before another or overlap with it. The special treatment to relate processes for which exact timing information is unavailable, is offered by Allen's algebra. Allen's algebra defines 13 basic interval relations: *before*, *meets*, *overlaps*, *starts*, *during*, *finishes*, *equals* and their inverses. The inverse of a relation should be interpreted as the relation that holds when the intervals are interchanged, for example if interval i_1 is before i_2, i_2 is after i_1, thus *before* and *after* are each other's inverse. In Figure 1 the relations are shown graphically. They are mutually exclusive and complete in the sense that any two intervals can be assigned exactly one relation.
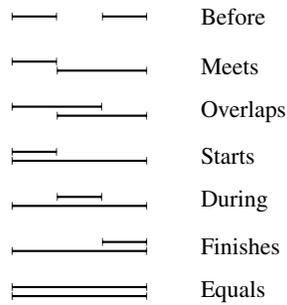

Fig. 1. Graphical representation of the seven basic relations that can hold between two time intervals. These relations and their inverse make up Allen's algebra.

Formally, we define intervals on a linearly ordered time line of points $(\mathcal{T}, \leq)$, which we take to be a subset of the set of the natural numbers $\mathbb{N}$. In the following we use the common abbreviations $<, \leq, =, \geq, >, \neq$ with their usual meaning to denote ordering and equality relations between elements of $\mathcal{T}$.

Definition 1. *An* interval I *is defined as a pair of time points* $I = [l, u)$, *with* $l, u \in \mathcal{T}$, *using the convention of right-open intervals. We then define two special points* $I^- = l = \inf I$ *and* $I^+ = u = \sup I$ *to distinguish the start and end of* I.

The following properties follow from the definitions given above:

Property 1: (trichotomy law) Only one of either $t < t'$, $t = t'$ or $t > t'$ holds.
Property 2: For each interval $I^- < I^+$.
Property 3: The number of intervals from $\mathcal{T}$ is equal to $\binom{|\mathcal{T}|}{2}$

Property 1 follows from the linear order. Property 2 follows from the definition of intervals and their nonempty nature. Finally, property 3 follows from the fact that each choice of two time points constitutes an interval, and that the start of an interval is always less than the end. Thus, the number of possible intervals is equal to the number

of ordered pairs taken from the set $\mathcal{T}$, where one number is always less than or greater than the other number. We can now define relations on intervals.

Definition 2. *Let* $\mathcal{I}$ *be the set of all intervals of* $\mathcal{T}$. *A binary* temporal interval relation R *is defined as* $R \subseteq \mathcal{I} \times \mathcal{I}$. *In the following we use the notation* IRJ *for* $(I, J) \in R$.

Allen defined a set of seven basic interval relations on two time intervals. Together with the inverses of these relations we obtain a minimal set of relations that can express any qualitative relation between two intervals. This set of relations with respect to the fixed set of intervals $\mathcal{I}$ 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.

Definition 3. *Let* $\mathcal{B}$ *be the set of* basic relations *on any two intervals* $I, J \in \mathcal{I}$, *with* I^- *denoting the start point and* I^+ *the end point of interval* I, *and similarly for* J:

$$IbJ \Leftrightarrow (I^+ < J^-) \qquad \qquad \textit{Interval I is before interval J}$$
$$ImJ \Leftrightarrow (I^+ = J^-) \qquad \qquad \textit{Interval I meets J}$$
$$IoJ \Leftrightarrow (I^- < J^-) \wedge (I^+ < J^+) \wedge (J^- < I^+) \qquad \textit{Interval I overlaps J}$$
$$IsJ \Leftrightarrow (I^- = J^-) \wedge (I^+ < J^+) \qquad \qquad \textit{Interval I starts J}$$
$$IdJ \Leftrightarrow (J^- < I^-) \wedge (I^+ < J^+) \qquad \qquad \textit{Interval I is during J}$$
$$IfJ \Leftrightarrow (J^- < I^-) \wedge (I^+ = J^+) \qquad \qquad \textit{Interval I finishes J}$$
$$IeqJ \Leftrightarrow (I^- = J^-) \wedge (I^+ = J^+) \qquad \qquad \textit{Interval I is equal to J}$$

The inverse of a relation R *is denoted* $\bar{R}$, *and is defined as* $I\bar{R}J \equiv JRI$, *the relation* equals *is thus its own inverse.*

In the examples we use events with an interval index instead of pure interval expressions, as this simplifies the exposition. The formal details are deferred until Section 3.

Example 1. Consider our lung disease example. 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 Inf_I at least partially precedes the increase in symptoms in interval Sym_J, where the notation indicates the event and the interval in which it occurs. We then obtain the statement: $\mathrm{Inf}_I o \mathrm{Sym}_J$, 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: $\mathrm{Exa}_K eq \mathrm{Sym}_J$.

Definition 4. *An* Allen relation *is defined as a disjunction of basic interval relations, represented as a set. The power set of the basic relations (all Allen relations) is denoted* $\mathcal{A} = \wp(\mathcal{B})$. *An* interval formula *is then of the form* IRJ *with* I, J *intervals and* $R \in \mathcal{A}$.

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. By the completeness of the basic relations we have 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 an interval 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: $\text{Vpi}_I\{\bar{o}, d, f\}\text{Exa}_J$.

2.2 Logical reasoning with the interval algebra

As Allen showed [7], this qualitative algebra is well suited to reason about time in a logic context. We are thus abstracting somewhat from the relational perspective above, and proceed to use a logical framework, that is, Allen's relations are represented by temporal predicates. The logic we will be using derives from the logic programming tradition of using Horn clauses, $H \leftarrow B_1, \ldots, B_n$, where H is the head of the clause and $B_1, \ldots, B_n$ the body and H and the B_i 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), we opt for the arguably simpler framework of temporal arguments [8]. This means that temporal predicates have a temporal argument specifying the relevant time interval. Note that this implies a typed logic, which we will leave implicit as this can always be translated to first order logic at the cost of notational convenience.

Example 3. Consider again our COPD example, now in logic representation:

$\text{exacerbation}(P, I') \leftarrow \text{patient}(P), \text{infection}(P, I), o(I, I').$
$\text{vpi}(P, I') \leftarrow \text{patient}(P), \text{exacerbation}(P, I), (\bar{o}(I', I); d(I', I); f(I', I)).$

Here vpi stands for ventilation-perfusion inequality.

2.3 CP-logic

To represent and reason with probabilistic knowledge, we will use the probabilistic logic language CP-logic [6]. This language is based on Prolog – providing the logic part of the language – extended with probabilistic semantics. The main intuition is that probabilistic logic statements represent causal laws, that is a logic clause gives a relation from some cause to a set of possible outcomes (each with some probability).

Definition 5. *A causal probabilistic law has the form:* $(H_1 : \alpha_1) ; \ldots ; (H_n : \alpha_n) \leftarrow B$ *where α_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 law gives a distribution over possible effects of a cause B. CP-logic is restricted to finite domains, so although you can write quantified rules, these are expanded to a set of ground instances for reasoning. The probabilistic semantics of CP-logic can be described as follows. As is common in logic programming the semantics are defined in terms of Herbrand interpretations, that is the domain is the set of constants of the theory and a symbol is interpreted as itself. The Herbrand universe is the set of all ground terms and the Herbrand base the set of ground atoms.

Definition 6. *Let H_U denote the Herbrand universe. A probabilistic process over H_U is a pair $\langle \mathbf{T}, \mathbf{I} \rangle$, where $\mathbf{T} = \langle V, E \rangle$ is a tree with each edge $e \in E$ labelled with a probability $P((v, w))$ and where for each node $v \in V$ the probabilities of the outgoing edges of v sum to 1: $\sum_{w:(v,w)\in E} P((v, w)) = 1$; $\mathbf{I}$ is an interpretation function, mapping nodes in $\mathbf{T}$ to Herbrand interpretations, i.e. subsets of the Herbrand base.*

Each transition between nodes is a probabilistic event, described by a causal law.

Definition 7. *A causal law c fires in a node v of $\mathbf{T}$ if v has child nodes $v_1, \ldots, v_{n+1}$ such that for $1 \leq i \leq n$: $\mathbf{I}(v_i) = \mathbf{I}(v) \cup \{H_i\}$ and the label of the edge (v, v_i) is α_i; $\mathbf{I}(v_{n+1}) = \mathbf{I}(v)$ and the label of the edge (v, v_{n+1}) is $1 - \sum_i \alpha_i$.*

The leaves of a probability tree each describe a possible outcome of the events modelled by causal laws. The probability of a leaf node l is the product of the labels on the edges from l to the root of the tree. By the fact that each causal law fires independently, the product over the probabilities of the outcomes on the path is indeed the probability of the state in l given the events on the path. 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. See also Vennekens et al. [6].

Example 4. In Figure 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-laws:

exacerbation : $0.6 \leftarrow$ infection.
exacerbation : $0.2 \leftarrow$ noxious_substance.
infection : 0.05.
noxious_substance : 0.01.

The probability of an exacerbation can be computed by summing over the leaves $l \in V$ that contain E (short for exacerbation) on the path from the root to l. The probability of, for instance, the left most path is $0.05 \cdot 0.6 \cdot 0.01 \cdot 0.2 = 0.00006$ and the probability of an exacerbation is:

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

Another way to obtain the probabilities is by considering a probability distribution over ground logic programs, which is the usual interpretation of ProbLog, a probabilistic language related to CP-logic, see e.g. [5]. Each grounding of a fact c_i in the logic theory T has some probability p_i, thus given that we only consider finite theories, a finite number of substitutions θ_{ij} gives the grounded set of all logical facts from the theory $L_T = \{c_1\theta_{11}, \ldots, c_1\theta_{1j_1}, \ldots, c_n\theta_{n1}, \ldots, c_n\theta_{nj_n}\}$. The program defines a probability distribution over ground logic programs $L \subseteq L_T$:

$$P(L \mid T) = \prod_{c_i\theta_j \in L} p_i \prod_{c_i\theta_j \in L_T \setminus L} (1 - p_i) \tag{1}$$

This is equivalent with the product over edges in the event tree described above. The probability of a query q is the marginal of $P(L \mid T)$ with respect to q, that is, those

I 0.05 ¬ I 0.95

E 0.6 ¬ E 0.4 N 0.01 ¬ N 0.99

N 0.01 ¬ N 0.99 N 0.01 ¬ N 0.99 E 0.2 ¬ E 0.8
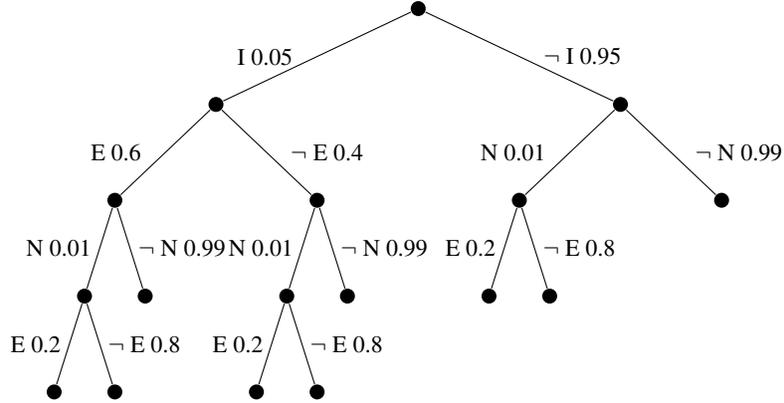
E 0.2 ¬ E 0.8 E 0.2 ¬ E 0.8

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

groundings of the program that prove the query $BK \cup L \vDash q$, where BK is the background knowledge. If we take $P(q \mid L) = 1$ if there exists a substitution θ such that $BK \cup L \vDash q\theta$ and $P(q \mid L) = 0$ otherwise, we obtain:

$$P(q \mid T) = \sum_{L \subseteq L_T} P(q \mid L)P(L \mid T), \tag{2}$$

which is again equivalent with the probability tree view, where we sum over the leaves in the tree that have the same interpretation.

3 A probabilistic extension of Allen's algebra

Probabilistic logic is a useful tool that combines logical and probabilistic reasoning, which can be used to extend Allen's logic framework for qualitative temporal reasoning with uncertainty. When modelling real world situations, qualitative time is useful for those processes for which it is difficult to obtain precise timing information. However, even when available timing information is only qualitative, events that occur are not necessarily equally likely. 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 kind of information can be represented using probabilistic logic. Temporal process descriptions – as represented with Allen's logic – can therefore be extended with probabilistic information, and the capabilities of CP-logic will appear sufficient to act as a basis for such an extended, qualitative temporal and uncertain logic.

3.1 On events and intervals

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 unique association is expressed by means of a time-interval index.

Definition 8. *Let $\mathcal{E}$ denote the* event space *containing all probabilistic events of interest. A temporal event E_I is defined as a probabilistic event $E \in \mathcal{E}$ that is temporally uncertain, expressed by the time interval $I \in \mathcal{I}$.*

In the probabilistic logic context sketched in the previous section events are represented by facts, which are interpreted in probabilistic logic 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. Now to define temporal events, we index facts with time, which in our case means that facts are augmented with an interval valued argument.

An important issue is the interpretation of events associated with intervals. At least three interpretations seem possible:

– An event implies that somewhere during the interval the modelled occurrence happens. That is the event is instantaneous but not precisely defined in time.
– An event is some occurrence that has a certain duration, for which the exact temporal extend is unknown. The interval gives the temporal bounds in which the event is contained.
– An event occurrence lasts the complete duration of the interval.

These interpretations lie on a spectrum from instantaneous events to extended events with different levels of temporal uncertainty involved. For instantaneous events uncertainty about a time interval can be derived from the uncertainty of the time points. However, in many domains, medicine among them, it is unrealistic to model events as instantaneous. When considering events with duration for which the interval gives lower and upper bounds for the start and end of the event, there can be uncertainty about the interval bounds and uncertainty about when the event occurs within the interval. For events that have a duration equal to the interval length, the temporal uncertainty lies in which interval is assigned to the event, since there is no uncertainty inside the interval.

As it is possible to have intermittent events, that is a recurring process which could be seen as a single event, we need to define how events with multiple intervals interact. For example, a fever may abate for a day and then return, which one could still look on as a single fever event. The algebraic properties of temporal events should make clear which properties hold. The Boolean algebra of temporal events $B(\mathcal{E}_\mathcal{I})$, where $\mathcal{E}_\mathcal{I}$ is defined as $\mathcal{E}_\mathcal{I} = \{E_I \mid E \in \mathcal{E}, I \in \mathcal{I}\}$, 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 $(E_I \wedge E'_J)$, disjunctions $(E_I \vee E'_J)$ and negations $\neg E_I$, with events $E_I, E'_J \in \mathcal{E}_\mathcal{I}$.

Now, for $E = E'$ with temporal events E_I and E'_J, there is an interaction between the Boolean operations on events and the relation between the time intervals. For example, if $I eq J$ holds, then $(E_I \wedge E_J) = E_I = E_J$. In addition, when two intervals I and J meet, i.e. $I m J$ holds, then $(E_I \wedge E_J) = E_K$ with $K = I \cup J$. Thus the event E actually occurred during interval K. It turns out that $(E_I \wedge E_J) = E_{I \cup J}$ holds for $I * J$ with $* \in \mathcal{B} \setminus \{b, \bar{b}\}$, i.e. it holds for all cases except when I and J are disconnected intervals. The other operations of the Boolean algebra do not interact with the relations.

3.2 Uncertainty

Uncertainty can be incorporated into the resulting language with respect to the temporal events E_I; and with regard to the relation between time intervals IRJ. When combined for two temporal events E_I and E'_J, these assumptions would give rise to a joint probability distribution of the form $P(E_I, E'_J, IRJ)$ with R an Allen relation. Note the special case when a single event E is associated with both intervals.

By conditioning on IRJ, one removes part of the uncertainty, yielding: $P(E_I, E'_J \mid IRJ)$. We first look at this case where only events are uncertain, and relations between interval are considered to be part of the logic specification. Because events are associated with intervals, the logic relation serves as a constraint on the events and as such still influences the uncertain part. Let us now look at specifying probabilities for events.

Definition 9. *The probability of a Boolean expression of events is given by the probability function $P : B(\mathcal{E}_\mathcal{I}) \to [0, 1]$, where $B(\mathcal{E}_\mathcal{I})$ denotes the Boolean algebra over the set of temporal events $\mathcal{E}_\mathcal{I}$.*

This defines uncertainty in terms of what is sometimes called a probability algebra – a pair consisting of a Boolean algebra and a probability function. But it tells us nothing yet over the actual shape of the distribution; one could for example parametrise on interval properties such as start point and length, or some domain dependent parametrisation.

We may also be interested in answering questions about whether a certain event occurred irrespective of time. For this we need a concept of atemporal events, summarising over the intervals. This is also related to multiple granularities, where some events occur at a different time scale than others, which appears to be an intermediate level of summary over time. If we are interested in the probability that an event occurs irrespective of time, we can say that this is equivalent to the probability that the event occurs in at least one interval, which can be expressed by the disjunction over events for each possible interval. Or more formally: $P'(E) = P\left(\bigvee_{I \in \mathcal{I}} E_I\right)$, where P' denotes the distribution over atemporal events. Since the disjunction over events is a proper probability, we obtain a distribution over atemporal events.

A similar question which may also be of interest is the probability of whether an event will have occurred by a certain time $t \in \mathcal{T}$. This can be solved by considering a subset of all intervals that satisfy the time constraint. Hence, $P''(E_{<t}) = P\left(\bigvee_{I \in S} E_I\right)$ with $S = \{I \mid I \in \mathcal{I}, I < t\}$, where at least all intervals that are entirely before t should be taken into account, hence $I^+ < t$. Limiting S to these intervals gives a lower bound. The meaning of $I < t$ for intervals with $I^- < t < I^+$ depends on the interpretation one chooses for temporal events (see Section 3.1). For the interpretation of events during the whole interval, it depends on whether you are interested in events that have started but not yet finished or only in completed events. For the other two interpretations it may be possible to compute a probability of partial intervals by looking inside the intervals.

Specific distributions If we now choose to define a distribution based on the endpoints of the interval associated with an event E, we can write: $P(E_I) = P_{I^-, I^+}(E_I)$, which indicates that the distribution depends on the parameters I^- and I^+, which we could also parametrise as $P(E_I) = P_{s,l}(E_I)$, with $s \in \mathcal{T}$ the interval start point and $l =$

$I^+ - I^-$ the length of the interval. Hence, the probability of for example an exacerbation event depends on when the exacerbation starts and its duration, which appear to be reasonable parameters to model clinical events of interest.

Given the parametrisation, we still have to choose the exact shape of the distribution, which depends on the exact situation we want to model. Here we consider a fairly general case, that *when* something happens is often unrelated to its *duration*, resulting in the assumption that the start point and duration are chosen independently, according to their own distribution.

Definition 10. *Let P be a specific distribution following Definition 9. Assuming that the probability of an event is described by the probability of the start point and duration of the interval, we obtain $P(E_I) = P_s(E_I)P_l(E_I)$, with s the interval start point and l the interval length.*

We can now choose a specific distribution for interval start points and interval length. A distribution like, for example, the beta-distribution seems useful to model particular situations as it produces different shapes depending on the parameters. In the discrete case with bounded support we obtain the same flexibility by using a beta-binomial distribution, a combination of a beta-distribution:

$$f(p;\, \alpha, \beta) = \frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha, \beta)} \quad \text{with} \quad B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, \mathrm{d}x,$$

where $f(p;\, \alpha, \beta)$ is parametrised by α, β and $B(\alpha, \beta)$ is the beta-function; and the binomial with the well-known form: $f(k;\, n, p) = \binom{n}{k}p^k(1-p)^{n-k}$. When compounded the beta-distribution gives the probability of the parameter p of the binomial, which leads to:

$$f(k;\, n, \alpha, \beta) = \binom{n}{k}\frac{\int_0^1 p^{k+\alpha-1}(1-p)^{n-k+\beta-1}\, \mathrm{d}p}{B(\alpha, \beta)} = \binom{n}{k}\frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)}.$$

The advantage of this distribution is that by choosing the parameters appropriately, we obtain useful special cases like a Bernoulli distribution when $n = 1$; the discrete uniform distribution for $\alpha = \beta = 1$; and a binomial distribution for large α and β.

For any particular modelling situation other distribution may be appropriate. For instance a combination of an exponential distribution (geometric in the discrete case) for interval start and a normal distribution over durations may be useful. But in general the choice of distribution is largely domain specific.

3.3 Reasoning with probabilistic intervals

Reasoning with temporal relations can now be given the added dimension of reasoning with uncertainty by considering the probabilities over temporal events. The logic framework allows us to do temporal reasoning which due to the probabilistic semantics automatically also gives us the probabilities, for which we only need specify probability distributions over events as described above. In medicine a fairly natural question to ask is what the probability of some event is given some other event that is temporally

related, which we can write as the probability: $P(E'_J, E_I : IRJ)$, because IRJ is determined logically it should be interpreted as a constraint on the distribution rather than probabilistic conditioning, hence the notation with a colon. The following proposition shows how we can compute this probability via the CP-logic semantics, given that IRJ is a logical relation between the unknown intervals associated with the events.

Proposition 1. *Let R be a relation $R \in \mathcal{A}$, and $E, E' \in \mathcal{E}$ events; the probability $P(E, E' : IRJ)$ is then obtained as follows:*

$$P'(E, E' : IRJ) = \sum_{I,J} \mathbb{I}(IRJ)P(E_I, E'_J) = \sum_{\ell} \mathbb{I}_\ell(IRJ)\prod_{e_\ell} P(e_\ell)$$

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

Proof. The indicator function ensures that R holds. By the fact that if IRJ holds in some leaf node of the tree, the temporal events E_I, E'_J must have occurred on the path from the root to ℓ, hence the probability of the events follows from the probability tree semantics of CP-logic. This can be seen by considering the probabilistic process given in Definition 6 and the transitions between nodes of Definition 7.

Although this proposition shows the probability calculus in the tree representation, it is also informative to see how they follow from the logical reasoning. We can then restate the proposition as follows.

Proposition 2. *Let R be a temporal relation $R \in \mathcal{A}$ and C the set of clauses $\{E(I), E'(J), IRJ\}$, the probability $P(C)$ then follows from the proofs $BK \cup L \vDash C$.*

Proof. By Equation 2 we sum over all proofs of C, that is all ground substitutions θ such that $BK \cup L \vDash C\theta$. The probability of each proof is given by Equation 1. Since each proof of IRJ requires instantiations for I and J, there will be probabilistic events E_I, E'_J that are part of the proof and that adhere to the constraint IRJ. Hence we obtain Proposition 1.

Example 5. We are again interested in modelling the relation between an infection and the occurrence of an exacerbation. The temporal relation is still *overlaps*, but now we also have probabilistic information attached to intervals. By choosing the parameters of the beta-binomial distribution $\alpha = \beta = 1$ we obtain a uniform distribution, which we can make explicit by writing down the possible intervals $\mathcal{I}$ with the time-line for this example restricted to $[0, 3)$. We then obtain the following logic specification:

$o(I1, I2) \leftarrow i(I1, S1, E1), i(I2, S2, E2), S2 > S1, S2 < E1, E2 > E1.$
$i(i1, 0, 1) : P; i(i1, 0, 2) : P; i(i1, 0, 3) : P;$
$\quad i(i1, 1, 2) : P; i(i1, 1, 3) : P; i(i1, 2, 3) : P \leftarrow \mathsf{betabinom}(P).$
$i(i2, 0, 1) : Q; i(i2, 0, 2) : Q; i(i2, 0, 3) : Q;$
$\quad i(i2, 1, 2) : Q; i(i2, 1, 3) : Q; i(i2, 2, 3) : Q \leftarrow \mathsf{betabinom}(Q).$
$\mathsf{exacerbation}(i2) \leftarrow \mathsf{infection}(i1), o(i1, i2).$
$\mathsf{infection}(i1).$

For the betabinomial distribution with $\alpha = \beta = 1$ we find $P = Q = 0.111$. The astute reader then notices that the probabilities in the example do not sum to 1 as would be expected. The reason for this is a censoring effect that results from our definition of probabilities on intervals. That is, each point on the time line is the start point of an interval with a certain probability. Since an interval cannot end before it starts, the possible end points are limited to those points that follow the start point but precede the end of the time line. Although it would be possible to specify a distribution over those points, that would result in the strange situation that shorter intervals become more likely towards the end of the time line. A more natural solution is thus to consider the end of the time line as a boundary that we cannot look beyond, but which does not limit the possibility of event occurring after the boundary. This results in a truncated distribution where the probability mass that falls beyond the time line is simply discarded, hence leading to a sum over interval probabilities lower than 1.

With this representation we can now answer probabilistic queries about our temporal concepts. The reasoning mechanics of CP-logic will take care of the probabilistic part of the queries. The probability of observing for instance an exacerbation in the interval $[1, 3)$ follows from the probability of an infection in some interval that overlaps with $[1, 3)$, which is this case only leaves the interval $[0, 2)$. The probability tree that is constructed thus contains a single path consisting of the events infection and exacerbation, with the uncertainty modelled through the probabilistic choices for the intervals, i.e.: $P([0, 2)) = 0.111$ and $P([1, 3)) = 0.111$ by the uniform distribution. The final probability is the product over the probabilities of the events in the tree, hence $P(\mathsf{exacerbation}([1, 3))) = 0.111 \cdot 0.111 \cdot 1 \cdot 1 = 0.0123$.

An advantage of this representation is that it is possible to start with a logical expression and add probabilities by defining a distribution over intervals. Besides the temporal probabilistic information, the probabilistic logic framework can also be used to incorporate more general probabilistic facts. For instance, in the example above, the infection predicate can easily be assigned a prior probability, for example $\mathsf{infection}(i1) : 0.1$. This models the situation that the probability of contracting an infection is 0.1, and the timing is distributed according to distribution $i1$.

Example 6. Consider again our running example. Now say we observed an exacerbation in the interval $[1, 3)$. Given the observation we can ask the question what the probability $P(\mathsf{vpi}(i1))$ is given the evidence $\mathsf{i}(i2, 1, 3)$. The answer follows from the probability tree where the possible outcomes of $i2$ are replaced by the determined evidence $\mathsf{i}(i2, 1, 3) : 1$. Now the reasoning mechanism can simply be applied leading to

$$
\begin{aligned}
P(\mathsf{vpi}(i1)) &= P(i1)P(i2)P(d(i1, i2) \vee f(i1, i2) \vee \bar{o}(i1, i2))P(\mathsf{vpi}(i1)) \\
&= P(i1)P([1, 3))P(d(i1, [1, 3)) \vee f(i1, [1, 3)) \vee \bar{o}(i1, [1, 3))) \cdot 1 \\
&= P([2, 3))P([1, 3))P(f([2, 3), [1, 3))) = 0.111 \cdot 0.111 \cdot 1 = 0.0123
\end{aligned}
$$

Note that in these examples we assume that the specified probability distributions are valid given that some temporal relation holds, which means that we modelled the relation between infection and exacerbation within the context of overlapping time intervals. This works for some situations, but it would also be interesting to look at temporal relation as influencing a distribution, instead of as a constraint. We could for example study how, given an event E_I, the additional information IbJ changes the distribution

of E_J. It is unlikely that we can say in general what the effect of temporal information will be, as this will be domain and event specific, however some regularity is expected.

Let us look at a specific case where we have events E_I, E'_J with IbJ. We could have a number of possible situations, for example, E and E' could be 'either-or' events which means that the added information of IbJ results in the probability of E'_J becoming zero because E_I already occurred. In our running example this could be the case for the probability of an exacerbation after the infection has ended (if we leave other causes like a second infection out of consideration). Another situation could be that E_I facilitates the occurrence of E', thus increasing the probability of E'_J when IbJ holds. Yet another possibility would be two events that usually occur overlapping in time, for which the additional information IbJ makes E'_J less likely.

The pattern that emerges shows some resemblance to the qualitative influences in qualitative probabilistic networks [9], where the temporal relation IbJ has a positive or negative influence on the probability of events given the relation. That is the probability of E'_J increases (decreases) when E_I has a positive (negative) influence given that the relation holds. The problem with such a characterisation is that it is hard to imagine what we can do with this in practice. Knowing that the probability increases does not tell us how exactly we should change the probability distribution. Nevertheless, studying this kind of patterns might be useful from a knowledge representation viewpoint; by defining specific patterns of temporal influence on distributions we acquire an additional modelling tool. So although we cannot ascertain the effect of temporal relations in general, it may be useful to add specific cases to out modelling language.

Uncertainty on relations As mentioned in Section 3.2 we may also be interested in the situation where not only events are uncertain but also the temporal relations themselves. The uncertainty on relations conveys that it may not be known what the order of events is. Allen's algebra uses disjunctions to model this situation, but with probabilities we can use additional likelihood information, if available, and otherwise a uniform distribution can be used to regain the deterministic case. The joined probability that is now relevant is: $P(E_I, E_J, IRJ) = P(E_I, E_J \mid IRJ)P(IRJ)$. Compared to the previous situation we thus need probabilities defined on relations, for which a discrete distribution constructed by hand seems adequate as the number of probabilities that has to specified is at most 13 (the cardinality of $\mathcal{B}$). Often a few relations will be more likely than others limiting the number of probabilities further. Note that although the set $\mathcal{A}$ is large ($|\wp(\mathcal{B})| = 2^{13}$) the probability of a relation $R \in \mathcal{A}$ is the sum over the basic relations in R, due to the mutual exclusivity of the basic relations.

Reasoning with uncertain relations requires little additional machinery. It results in additional probabilistic facts in our logic, but these behave like any other fact. The combined logic and probabilistic reasoning thus proceeds as described before. How a prior probability on relations should be interpreted, independent of events, is less clear however. One solution would be using domain knowledge to count how many times certain relations occur when marginalising over all events of interest. This requires a relevant data set, that then provides prior probabilities for the occurrence of temporal relations that are meaningful within the specific domain.

4 Related work

Allen's algebra has had much attention over the years and finds applications in various fields, from planning to clinical medicine and many more. Besides Allen's work [1, 7], well-known contemporary work is the temporal logic by McDermott [10]. Both authors deal with much broader concepts than those considered here, like continuous change, actions and plans, but 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.

Probabilistic temporal interval networks [11] are a probabilistic extension of the network representation often used to specify the consistency problem in Allen's algebra. The relations between two intervals are weighted with probabilities. This is thus a generalisation of the uncertainty that can exist about what relation holds between two intervals from disjunctions to a distributions over relations. Probabilistic temporal networks [12], are network models that incorporate Allen's constraints on conditional probabilities. They make the assumption that the intervals of interest are known beforehand and can be specified explicitly, thus not allowing uncertainty in what intervals are or will be of interest. Then, recently, probabilistic logic has been applied to represent stochastic processes [13]. The proposed language CPT-L extends CP-logic to represent fully-observable homogeneous Markov processes, and allows efficient reasoning.

5 Conclusion

In this paper we have looked at the interaction of qualitative time and probability, to obtain a temporal representation with uncertainty. The representation language seems usable to reason with uncertain temporal information, although work remains to further study its properties and applicability.

References

1. Allen, J.: Maintaining knowledge about temporal intervals. Commun ACM (1983)
2. Shahar, Y.: A framework for knowledge-based temporal abstraction. Artif Intell (1997)
3. Poole, D.: Probabilistic Horn abduction and Bayesian networks. Artif Intell (1993)
4. Kersting, K., de Raedt, L.: Bayesian logic programs. TR 151, University of Freiburg (2000)
5. Kimmig, A., Demoen, B., de Raedt, L., Santos Costa, V., Rocha, R.: On the implementation of the probabilistic logic programming language ProbLog. Theor Pract Log Prog (2010)
6. Vennekens, J., Denecker, M., Bruynooghe, M.: CP-logic: A language of causal probabilistic events and its relation to logic programming. Theor Pract Log Prog (2009)
7. Allen, J.: Towards a general theory of action and time. Artif Intell (1984)
8. Bacchus, F., Tenenberg, J., Koomen, J.: A non-reified temporal logic. Artif Intell (1991)
9. Wellman, M.: Fundamental concepts of qualitative probabilistic networks. Artif Intell (1990)
10. McDermott, D.: A temporal logic for reasoning about processes and plans. Cog Sc (1982)
11. Ryabov, V., Trudel, A.: Probabilistic temporal interval networks. In: TIME. (2004)
12. Santos Jr., E., Young, J.: Probabilistic temporal networks: A unified framework for reasoning with time and uncertainty. Int J Approx Reason (1999)
13. Thon, I., Landwehr, N., de Raedt, L.: Stochastic relational processes: Efficient inference and applications. Mach Learn (2011)