Logical forms of chronicles

T. Guyet and N. Markey

TIME Symposium 2022

11/07/2022

Outline

Context: querying care pathways

- Epidemiological studies with medical databases
  - → aim to count some patients of interest suffering from disease or having a treatment
  - → diseases or treatments are high-level medical events (phenotypes) not necessarily coded
  - → need for defining phenotypes from low level databases features
- The task
  - Inputs:
    - A patient care pathway : ⟨(, 10), (, 20) . . .⟩
    - A phenotype (a *query*)
  - Output: Yes or No ... the patient matches the phenotype

Challenge: Define a framework to specify and answer to complex queries (phenotype matching) that is

- expressive enough to specify complex *phenotypes* **with temporal information**
- efficient to be applied on large numbers of care pathways

Querying with temporal pattern matching

Temporal sequence/Timed sequences

- Σ: a finite set of events,
- $\mathbb{T} = \mathbb{Q}$ or $\mathbb{R}$: temporal domain
- $S = \langle (e_1, t_1), (e_2, t_2), \ldots, (e_n, t_n) \rangle$ a finite temporal sequence where $e_i \in \Sigma$ and $t_i \in \mathbb{T}$

Temporal pattern

- A temporal pattern p represents a situation to recognize
- p is specified in a formal syntax
- p can *occur* in or *match* a sequence

$\Rightarrow$ a temporal pattern = a phenotype

$\Rightarrow$ recognize a temporal pattern = answer a query

Which domain of temporal pattern is interesting?

Many formalisms to express temporal patterns

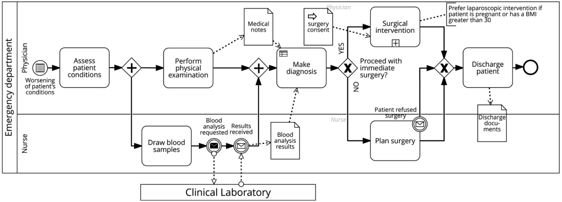
- from database community: TSQL2 [BCST96], DatalogMTL [WCGKK19]
- from business models community: BPM
- from logic: LTL [Pnu77], MTL [Koy90], TPTL [AH94], Event Calculus [Mue08], Situation calculus [LPR98]...
- from complex event processing community: simple temporal networks [DMP91], chronicles [DVD99], ONERA chronicles [KP20], ...

SELECT ShowName VALID CAST(BEGIN(VALID(A) AS DAY)) FROM NBCShows(ShowName)(PERIOD) AS A WHERE CAST(VALID(A) AS INTERVAL YEAR) >= INTERVAL '2' YEAR

Which domain of temporal pattern is interesting?

Many formalisms to express temporal patterns

- from database community: TSQL2 [BCST96], DatalogMTL [WCGKK19]
- from business models community: BPM
- from logic: LTL [Pnu77], MTL [Koy90], TPTL [AH94], Event Calculus [Mue08], Situation calculus [LPR98]...
- from complex event processing community: simple temporal networks [DMP91], chronicles [DVD99], ONERA chronicles [KP20], ...



Which domain of temporal pattern is interesting?

Many formalisms to express temporal patterns

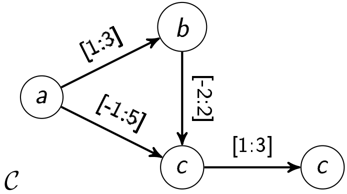
- from database community: TSQL2 [BCST96], DatalogMTL [WCGKK19]
- from business models community: BPM
- from logic: LTL [Pnu77], MTL [Koy90], TPTL [AH94], Event Calculus [Mue08], Situation calculus [LPR98]...
- from complex event processing community: simple temporal networks [DMP91], chronicles [DVD99], ONERA chronicles [KP20], ...

$$\phi = x.\Diamond(b \wedge \Diamond(c \wedge x \leq 2))$$

Which domain of temporal pattern is interesting?

Many formalisms to express temporal patterns

- from database community: TSQL2 [BCST96], DatalogMTL [WCGKK19]
- from business models community: BPM
- from logic: LTL [Pnu77], MTL [Koy90], TPTL [AH94], Event Calculus [Mue08], Situation calculus [LPR98]...
- from complex event processing community: simple temporal networks [DMP91], chronicles [DVD99], ONERA chronicles [KP20], ...



Why chronicles?

Sufficient for specifying most of the temporal relations needed in phenotype expressions

- "*for 3 to 12 months*"
- "*at most 0 to 2 days*"

Good properties

- interpretability: the graphical representation makes them easy to manipulate / interpret
- efficiency: there are efficient algorithms to recognize/enumerate occurrences of a chronicle in long sequences [GBS+20]
- versatile usage: online monitoring, pattern matching, pattern mining
- expressiveness: complex temporal arrangement can be specified

Our research question: better qualify the model expressiveness

why chronicles are qualified as "expressive"?? (*initial intuition*)

→ Metric temporal constraints ... like most of the other models

→ A chronicle does not enforce the order of occurrence of events ... to be investigated

- it is possible thanks to positive and negative boundaries on the delay between occurrences

But no formal result in the state of the art ...

⇒ Our research question: can chronicles be expressed with classical metric temporal logics, namely MTL or TPTL?

Let $\mathscr{C}$ be a chronicle, does exist a formula $\varphi_{\mathscr{C}}$ such that, for all timed sequence s:

$$\mathscr{C} \in \mathsf{s} \overset{?}{\Leftrightarrow} \mathsf{s} \models \varphi_{\mathscr{C}}$$

Outline

Chronicles: formal definition

Definition (Chronicle)

A **chronicle** is a pair $(\mathcal{E}, \mathcal{T})$ where

- $\mathcal{E}$ is a **multiset** over Σ, i.e. $\mathcal{E}$ is of the form $\{\!\{c_1, \ldots, c_m\}\!\}$ such that $c_i \in \Sigma$ for $i = 1, \ldots, m$ and $c_1 \leq_\Sigma \cdots \leq_\Sigma c_m$.
- $\mathcal{T}$ is a set of **temporal constraints**, i.e. expressions of the form $(c, o_c)[t^-, t^+](c', o_{c'})$ such that
  1. $c, c' \in \mathcal{E}$ and
  2. $t^-, t^+ \in \mathbb{Q} \cup \{-\infty, +\infty\}$ and *[metric constraints]*
  3. $o_c, o_{c'} \in [m]$ and $o_c < o_{c'}$ and *[edge direction]*
  4. $c_{o_c} = c$ and $c_{o_{c'}} = c'$.

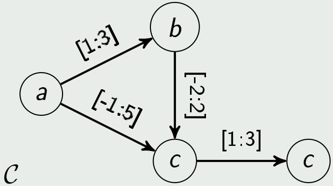
The size of a chronicle $(\mathcal{E}, \mathcal{T})$ is the size m of its multiset $\mathcal{E}$.

Chronicles

Chronicle [DGG93]: Example and graphical representation

Chronicle $(\mathcal{E}, \mathcal{T})$ where

- $\mathcal{E} = \{\!\{a, b, c, c\}\!\}$
- $\mathcal{T} = \{(a, 1)[1 : 3](b, 2),$
  $(a, 1)[-1 : 5](c, 3),$
  $(b, 2)[-2 : 2](c, 3),$
  $(c, 3)[1 : 3](c, 4)\}$



Additional comments (see [BG22] for more details)

- all constraints and conjunctives
- at most one temporal constraint between each pair of events
- arrows are directed according to the order of Σ
- like for simple temporal networks temporal constraints can be "minimized" (*but not of interest for our research question*)
  - eliminate redundancies
  - identify inconsistent network

Occurrence of chronicles (semantics)

Definition (Occurrence of a chronicle)

Let $s = \langle(\sigma_1, \tau_1), (\sigma_2, \tau_2), \ldots, (\sigma_n, \tau_n)\rangle$ be a timed sequence of length n, and $\mathscr{C} = (\mathcal{E} = \{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T})$ be a chronicle of size m.
Chronicle $\mathscr{C}$ is said to occur in s if, and only if, there exists an injective function $\varepsilon\colon [m] \to [n]$, hereafter called *embedding*, such that:

1. for all $1 \leq i \leq m$, $\sigma_{\varepsilon(i)} = c_i$, [*event mapping*]
2. for all $1 \leq i, j \leq m$, $\tau_{\varepsilon(j)} - \tau_{\varepsilon(i)} \in [t^-, t^+]$ whenever $(c_i, i)[t^-, t^+](c_j, j) \in \mathcal{T}$, [*temporal constraints*]
3. for all $1 \leq i < m$, $\tau_{\varepsilon(i)} < \tau_{\varepsilon(i+1)}$ whenever $c_i = c_{i+1}$. [*implicit order*]

Then,

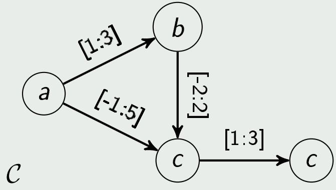
- $\tilde{s} = \{(\sigma_{\varepsilon(1)}, \tau_{\varepsilon(1)}), \ldots, (\sigma_{\varepsilon(m)}, \tau_{\varepsilon(m)})\}$ is an *occurrence* of $\mathscr{C}$ in s.
- Chronicle $\mathscr{C}$ is said to *match* the sequence s if, and only if, there is at least one occurrence of $\mathscr{C}$ in s.

Chronicles

Chronicle [DGG93]: Example and graphical representation

Chronicle $(\mathcal{E}, \mathcal{T})$ where

- $\mathcal{E} = \{\!\{a, b, c, c\}\!\}$
- $\mathcal{T} = \{(a, 1)[1:3](b, 2),$
  $(a, 1)[-1:5](c, 3),$
  $(b, 2)[-2:2](c, 3),$
  $(c, 3)[1:3](c, 4)\}$



Occurrences of a chronicle

SID	Sequence
s_1	$(a,1)$, $(b,3)$, $(a,4)$, $(b,4.7)$, $(c,5)$, $(c,6)$, $(d,7)$
s_2	$(b,2)$, $(d,4)$, $(a,5)$, $(c,7)$
s_3	$(a,1)$, $(b,4)$, $(c,5)$, $(b,6)$, $(c,8)$,$(d,9)$
s_4	$(b,4)$, $(a,6)$, $(e,8)$, $(c,9)$
s_5	$(b,1)$, $(a,3)$, $(c,4)$
s_6	$(c,4)$, $(a,5)$, $(c,7)$,$(b,8)$, $(c,9)$

Outline

Reminder about $\text{TPTL}_\Diamond$

Syntax of $\text{TPTL}_\Diamond$

$\text{TPTL}_\Diamond$ is a fragment of TPTL with $\Diamond$ operator only:

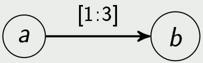
$$\text{TPTL} \ni \varphi ::= \sigma \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \Diamond\varphi \mid x.\varphi \mid x \sim c$$

where $\sim \in \{\leq, <, =, >, \geq\}$, $c \in \mathbb{Q}$, $\sigma \in \Sigma$.
We also define $\bar{\Diamond}\varphi$ that stands for $\varphi \vee \Diamond(\varphi)$.

Semantics of $\text{TPTL}_\Diamond$

- $x.\varphi$ (*clock resets*): store the current time in clock x

Example with a simple chronicle



$$\bar{\Diamond}(a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq 1))$$

$$s = (b, 2), (a, 4), (c, 5), (b, 7)$$

▸ Details

Reminder about TPTL$_\Diamond$

Syntax of TPTL$_\Diamond$

TPTL$_\Diamond$ is a fragment of TPTL with $\Diamond$ operator only:

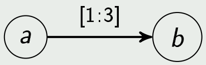
$$\text{TPTL} \ni \varphi ::= \sigma \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \Diamond\varphi \mid x.\varphi \mid x \sim c$$

where $\sim \in \{\leq, <, =, >, \geq\}$, $c \in \mathbb{Q}$, $\sigma \in \Sigma$.
We also define $\bar{\Diamond}\varphi$ that stands for $\varphi \vee \Diamond(\varphi)$.

Semantics of TPTL$_\Diamond$

- $x.\varphi$ (*clock resets*): store the current time in clock x

Example with a simple chronicle

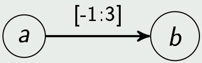


$$\bar{\Diamond}(a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq 1))$$

Is it possible the construct such a formula for any chronicle?

Implicit order of events' occurrences in TPTL$_\Diamond$ formulae

Not a systematic construction ...



$\Diamond(a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq -1))$

This formula enforces to have b after a in a timed sequence.

Leads to propose to construct an equivalent formula in two steps

1. decompose a chronicle into a disjunction of *linear chronicles*
2. construct an equivalent TPTL$_\Diamond$ formula for each *linear chronicle*

Linear chronicles

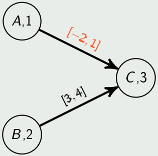
Definition (Linear chronicle)

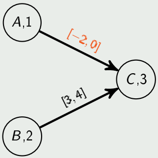
A linear chronicle is a triple $\mathscr{L} = (\{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T}, \pi)$, where $(\{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T})$ is a chronicle and π is a permutation of $[m]$.
$\mathscr{L} = (\{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T}, \pi)$ occurs in a timed sequence s whenever there exists an embedding $\varepsilon \colon [m] \to [n]$ witnessing that the chronicle $(\{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T})$ occurs in s, and such that $\varepsilon \circ \pi$ is increasing.

Proposition

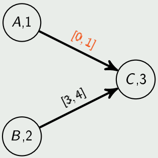
For any chronicle, there exists an equivalent disjunctive collection of linear chronicles.

Example (Equivalent collection of linear chronicles)

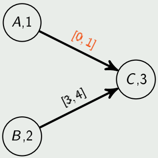




$\pi = B \prec C \prec A$



$\pi = B \prec A \prec C$



$\pi = A \prec B \prec C$

Inductive construction of TPTL$_\Diamond$ formula

Let $\mathscr{L} = (\mathcal{E} = \{\!\{c_1, \ldots, c_m\}\!\}, \mathcal{T}, \pi)$, we define a formulae $\varphi_{\mathscr{L}} = \bar{\Diamond}\varphi^1_{\mathscr{L}}$ where the collection of formulae $(\varphi^i_{\mathscr{L}})_{i=1\ldots m}$ is defined as follows: if $m = 1$, then $\varphi^1_{\mathscr{L}} = c_{\pi(1)}$; otherwise,

$$\varphi^1_{\mathscr{L}} = \left(c_{\pi(1)} \wedge x_{\pi(1)}.\Diamond\varphi^2_{\mathscr{L}}\right) \tag{1}$$

and for all $2 \leq i \leq m-1$,

$$\varphi^i_{\mathscr{L}} = c_{\pi(i)} \wedge \mathscr{T}_i(\mathscr{L}) \wedge x_{\pi(i)}.\Diamond\varphi^{i+1}_{\mathscr{L}} \tag{2}$$

and finally

$$\varphi^m_{\mathscr{L}} = c_{\pi(m)} \wedge \mathscr{T}_m(\mathscr{L}) \tag{3}$$

where

$$\mathscr{T}_i(\mathscr{L}) = \bigwedge_{\substack{(c_{\pi(k)},\pi(k))[l,u](c_{\pi(i)},\pi(i))\in\mathcal{T} \\ \pi(k)<\pi(i)}} \left(l \leq x_{\pi(k)} \leq u\right) \wedge$$

$$\left((\pi(i) > 1 \wedge c_{\pi(i)} = c_{\pi(i)-1}\right) \rightarrow x_{\pi(i)-1} > 0) \tag{4}$$

Inductive construction of TPTL$_\Diamond$ formula

Let $\mathscr{L} = (\mathcal{E} = \{\!\{c_1, \dots, c_m\}\!\}, \mathcal{T}, \pi)$, we define a formulae $\varphi_{\mathscr{L}} = \bar{\Diamond}\varphi^1_{\mathscr{L}}$ where the collection of formulae $(\varphi^i_{\mathscr{L}})_{i=1\dots m}$ is defined as follows: if $m = 1$, then $\varphi^1_{\mathscr{L}} = c_{\pi(1)}$; otherwise,

$$\varphi^1_{\mathscr{L}} = \left(c_{\pi(1)} \wedge x_{\pi(1)}.\Diamond\varphi^2_{\mathscr{L}}\right) \tag{1}$$

and for all $2 \leq i \leq m-1$,

$$\varphi^i_{\mathscr{L}} = c_{\pi(i)} \wedge \mathscr{T}_i(\mathscr{L}) \wedge x_{\pi(i)}.\Diamond\varphi^{i+1}_{\mathscr{L}} \tag{2}$$

and finally

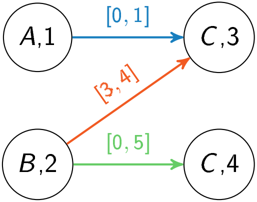
$$\varphi^m_{\mathscr{L}} = c_{\pi(m)} \wedge \mathscr{T}_m(\mathscr{L}) \tag{3}$$

where

$$\mathscr{T}_i(\mathscr{L}) = \bigwedge_{\substack{(c_{\pi(k)},\pi(k))[l,u](c_{\pi(i)},\pi(i))\in\mathcal{T} \\ \pi(k)<\pi(i)}} \left(l \leq x_{\pi(k)} \leq u\right) \wedge$$

$$\left(\left(\pi(i) > 1 \wedge c_{\pi(i)} = c_{\pi(i)-1}\right) \rightarrow x_{\pi(i)-1} > 0\right) \tag{4}$$

Inductive construction of TPTL$_\Diamond$ formula: example



$B \prec A \prec C_3 \prec C_4$

$$\mathscr{T}_1 = \emptyset$$
$$\mathscr{T}_2 = \emptyset$$
$$\mathscr{T}_3 = x_1 \leq 1 \wedge x_1 \geq 0 \wedge x_2 \leq 4 \wedge x_2 \geq 3$$
$$= x_1 \leq 1 \wedge x_2 \leq 4 \wedge x_2 \geq 3$$
$$\mathscr{T}_4 = x_2 \leq 5 \wedge x_2 \geq 0 \wedge x_3 > 0$$
$$= x_2 \leq 5 \wedge x_3 > 0$$

$$\varphi_{\mathscr{L}} = \bar{\Diamond}(B \wedge x_2.\Diamond(A \wedge \mathscr{T}_2 \wedge x_1.\Diamond(C \wedge \mathscr{T}_3 \wedge x_3.\Diamond(C \wedge \mathscr{T}_4))))$$

$$\varphi_{\mathscr{L}} = \Diamond(B \wedge x_2.\Diamond(A \wedge x_1.\Diamond(C \wedge x_1 \leq 1 \wedge x_2 \leq 4 \wedge x_2 \geq 3 \wedge x_3.\Diamond(C \wedge x_2 \leq 5 \wedge x_3 > 0))))$$
$$\vee (B \wedge x_2.\Diamond(A \wedge x_1.\Diamond(C \wedge x_1 \leq 1 \wedge x_2 \leq 4 \wedge x_2 \geq 3 \wedge x_3.\Diamond(C \wedge x_2 \leq 5 \wedge x_3 > 0))))$$

Inductive construction of TPTL$_\Diamond$ formula: properties

Proposition

For any linear chronicle $\mathscr{L}$, $\varphi_{\mathscr{L}}$ is an equivalent TPTL$_\Diamond$ formula for $\mathscr{L}$.

Theorem

Any chronicle $\mathscr{C}$ admits an equivalent TPTL$_\Diamond$ formula $\varphi_{\mathscr{C}}$.

Size of the formulae

- A chronicle of size m requires to encode at most $m!$ linear chronicles
- For a linear chronicle of size m, the size of the equivalent formula is at most $2 \times (2 \times m(m-1) + m)$
- For a linear chronicle of size m, the number of clocks is at most $m-1$

Outline

1. Introduction
2. Chronicles
3. Comparison with TPTL
4. Comparison with MTL
5. Conclusions and perspectives

Reminder about MTL

Syntax

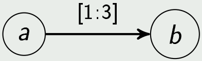
$$\mathsf{MTL} \ni \varphi ::= \sigma \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \mathcal{U}_I \varphi_2$$

where σ ranges over Σ, and $I = [l, u]$ is a temporal interval in $\mathbb{Q}$.
We also define the classical shorthand $\Diamond_I \varphi$, which stands for $\top \mathcal{U}_I \varphi$.

Pointwise semantics

- $(\rho, i) \models \varphi_1 \mathcal{U}_I \varphi_2 \iff \exists j \geq i$ s.t. $(\rho, j) \models \varphi_2$ and $\tau_j - \tau_i \in I$ and $\forall i < k < j,\ (\rho, k) \models \varphi_1$

Example with a simple chronicle

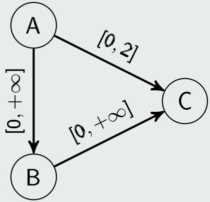


$$\Diamond(a \wedge \bar{\Diamond}_{[1,3]} b)$$

Is is possible to construct such a formula for any chronicle?

Simple chronicle with no equivalent MTL formulae

Example (Example of simple chronicle without MTL equivalent)



- It has only three events!
- It has two non-trivial permutations as A and B can occur at the same time
- Equivalent simplified TPTL$_\Diamond$ formula with permutation order $A \prec B \prec C$:

$$\bar{\Diamond}(a \wedge x.\Diamond(b \wedge y.\Diamond(c \wedge x \leq 2)))$$

Previous results on expressiveness of TPTL and MTL [BCM10]

- TPTL$_\Diamond$ is not comparable with MTL (different expressiveness)
- An example of a TPTL formula that has no equivalent MTL formula in the pointwise semantics is: $\phi = x.\Diamond(b \wedge \Diamond(c \wedge x \leq 2))$

$\Rightarrow$ the same kind of reasoning applies to exhibit some problematic sequences

Outline

Conclusions

Conclusion on the expressiveness of chronicles

- any chronicle can be equivalently encoded in a $\text{TPTL}_\Diamond$ formula
- some (simple) chronicles have no equivalent MTL formula

Comments about this results

- Surprisingly, the use of negative boundaries is not a difficult problem (it simply leads to longer formulae)
- Intuitively, MTL is not expressive enough because of the lack of memories (single clock)

Practical use of chronicle

- Chronicles can be used in practice: `PyChronicles` *pip* package
  - Pure Python encoding of chronicle recognition
  - Extension of Pandas dataframe to recognize patterns in sequences
- We propose a equivalent $\text{TPTL}_\Diamond$ formula construction, but its purpose is not to be used in practice

Perspectives

Equivalence with $TPTL_\Diamond$ (without negation)?

$\rightarrow$ would it be possible to express any $TPTL_\Diamond$ as a disjunction of chronicles?

Construction improvement

- Propose shorter $TPTL_\Diamond$ formula
- Propose $TPTL_\Diamond$ formula with fewer clocks

Compare the expressiveness of other temporal pattern languages

- MTL with past
- ONERA Chronicles [KP20]
- Timed automaton
- ...

References I

- Rajeev Alur and Thomas A. Henzinger, *A really temporal logic*, Journal of the ACM 41 (1994), no. 1, 181–203.
- Patricia Bouyer, Fabrice Chevalier, and Nicolas Markey, *On the expressiveness of TPTL and MTL*, Information and Computation 208 (2010), no. 2, 97–116.
- Michael H Böhlen, Jan Chomicki, Richard T Snodgrass, and David Toman, *Querying TSQL2 databases with temporal logic*, Proceedings of the International Conference on Extending Database Technology, 1996, pp. 325–341.
- Philippe Besnard and Thomas Guyet, *Chronicles*, under submission, 2022.
- Christophe Dousson, Paul Gaborit, and Malik Ghallab, *Situation recognition: Representation and algorithms*, Proc. of the 13th Int. Joint Conf. on Artificial Intelligence (IJCAI 1993) (Ruzena Bajcsy, ed.), Morgan Kaufmann, August 28 - September 3, 1993, pp. 166–172.
- Rina Dechter, Itay Meiri, and Judea Pearl, *Temporal constraint networks*, Artificial intelligence 49 (1991), no. 1-3, 61–95.
- Christophe Dousson and Thang Vu Duong, *Discovering chronicles with numerical time constraints from alarm logs for monitoring dynamic systems.*, Proceedings of the international joint conference on Artifical intelligence (IJCAI), 1999, pp. 620–629.
- Thomas Guyet, Philippe Besnard, Ahmed Samet, Nasreddine Ben Salha, and Nicolas Lachiche, *Énumération des occurrences d'une chronique*, Actes de la conférence Extraction et Gestion des Connaissances (EGC), 2020, pp. 253–260.
- Ron Koymans, *Specifying real-time properties with metric temporal logic*, Real-time systems 2 (1990), no. 4, 255–299.

References II

Romain Kervac and Ariane Piel, *A survey on chronicles and other behavior detection techniques*, Journal of Aerospace Lab 15 (2020).

Hector Levesque, Fiora Pirri, and Ray Reiter, *Foundations for the situation calculus*, Linköping Electronic Articles in Computed and Information Science 3 (1998), no. 18.

Erik T Mueller, *Event calculus*, Foundations of Artificial Intelligence 3 (2008), 671–708.

Amir Pnueli, *The temporal logic of programs*, Proceedings of the Annual Symposium on Foundations of Computer Science (SFCS), 1977, pp. 46–57.

Przemyslaw A. Wałęga, Bernardo Cuenca Grau, Mark Kaminski, and Egor V. Kostylev, *DatalogMTL: Computational complexity and expressive power*, Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2019, pp. 1886–1892.

Chronicle occurrence enumeration algorithm

Main principles

- an *on-going occurrence* of a chronicle maps each event of $\mathcal{E}$ with an interval of admissible positions

$$\forall e \in \mathcal{E},\ adm(e) = [a^-, a^+]$$

- the algorithm transverses the set of events $e \in \mathcal{E}$
  - progressively narrows intervals to single position (within admissible ones) corresponding to occurrence of e in S
  - → if does not exist, then this occurrence is discarded
  - → if multiple admissible, duplicate on-going occurrences
- for each (not yet used) constraint $(e, o_e)[t^-, t^+](e', o_{e'})$, the admissible positions are updated:

$$adm(e') = [a'^-, a'^+] := [a'^-, a'^+] \cap [t_e + t^-, t_e + t^+],$$

as soon as $adm(e') = \emptyset$ discard the on-going occurrence!

▸ Details ▸ Example

Chronicle occurrence enumeration algorithm

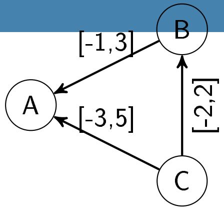
Algorithm "idea"

order event types by increasing frequency in sequence S
for all occurrence of $e_0 \in \mathcal{E}$

1. create a on-going occurrence $o = ([t_0, t_0], [-\infty, \infty], [-\infty, \infty], \dots)$
2. for events $e_i \in [e_1, \dots, e_n]$
   1. for each on-going occurrence o
      1. look for event e_i within interval $o_i = [t_i^-, t_i^+]$
      2. duplicate o_i for each of the admissible occurrences of e_i
      3. propagate constraints

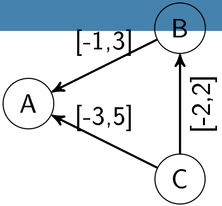
Algorithm example

$\langle(E,1)\ (B,2)\ (A,3)\ (E,4)\ (C,5)$
$(B,6)\ (A,7)\ (E,8)\ (A,9)\ (A,10)\rangle$



- Event processing order : C, B and A
- Processing of C
  - generate a single occurrence $o = ([5,5],[-\infty,\infty],[-\infty,\infty])$

Algorithm example

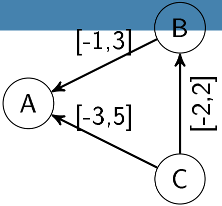


$\langle (E,1)\ (B,2)\ (A,3)\ (E,4)\ (C,5)$
$(B,6)\ (A,7)\ (E,8)\ (A,9)\ (A,10) \rangle$

- Event processing order : C, B and A
- Processing of C
  - generate a single occurrence $o = ([5,5], [-\infty,\infty], [-\infty,\infty])$
  - propagate constraints:
    - $C \rightarrow B$: $o = ([5,5], [3,7], [-\infty,\infty])$
    - $C \rightarrow A$: $o = ([5,5], [3,7], [2,10])$

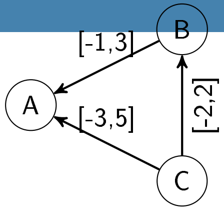
Algorithm example

$\langle (E,1)\ (B,2)\ (A,3)\ (E,4)\ (C,5)$
$(B,6)\ (A,7)\ (E,8)\ (A,9)\ (A,10)\rangle$



- Event processing order : C, B and A
- Processing of C
  - generate a single occurrence $o = ([5,5],[-\infty,\infty],[-\infty,\infty])$
  - propagate constraints:
    - $C \to B$: $o = ([5,5],[3,7],[-\infty,\infty])$
    - $C \to A$: $o = ([5,5],[3,7],[2,10])$
- Processing of B
  - narrow intervals with occurrences: $(B,2)$ is invalid ($2 \notin [3,7]$), so $o = ([5,5],[6,6],[2,10])$
  - propagate constraints:
    - $B \to A$: $o = ([5,5],[6,6],[2,10] \cap [5,9]) = ([5,5],[6,6],[5,9])$

Algorithm example



$\langle (E,1)\ (B,2)\ (A,3)\ (E,4)\ (C,5)$
$(B,6)\ (A,7)\ (E,8)\ (A,9)\ (A,10) \rangle$

- Event processing order : C, B and A
- Processing of C
  - generate a single occurrence $o = ([5,5], [-\infty,\infty], [-\infty,\infty])$
  - propagate constraints:
    - $C \rightarrow B$: $o = ([5,5], [3,7], [-\infty,\infty])$
    - $C \rightarrow A$: $o = ([5,5], [3,7], [2,10])$
- Processing of B
  - narrow intervals with occurrences: $(B,2)$ is invalid ($2 \notin [3,7]$), so $o = ([5,5], [6,6], [2,10])$
  - propagate constraints:
    - $B \rightarrow A$: $o = ([5,5], [6,6], [2,10] \cap [5,9]) = ([5,5], [6,6], [5,9])$
- Processing of A
  - narrow intervals with occurrences: $(A,3)$ and $(A,10)$ are invalid, but $(A,7)$ and $(A,9)$ are valid, two occurrences
    - $([5,5], [6,6], [7,7])$ and $([5,5], [6,6], [9,9])$

TPTL$_\Diamond$ details

$s = (b, 2), (a, 4), (c, 5), (b, 7)$
$\varphi = \Diamond(a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq 1))$

$$\begin{aligned} s \models \varphi \quad & \textit{iff} \quad \varphi, 0, [x =?] \models \Diamond(a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq 1)) \\ & \textit{iff} \quad \varphi, 4, [x =?] \models a \wedge x.\Diamond(b \wedge x \leq 3 \wedge x \geq 1) \\ & \textit{iff} \quad \varphi, 4, [x = 4] \models \Diamond(b \wedge x \leq 3 \wedge x \geq 1) \\ & \textit{iff} \quad \varphi, 7, [x = 4] \models b \wedge x \leq 3 \wedge x \geq 1 \end{aligned}$$

$\varphi, 7, [x = 4] \models b \wedge x \leq 3 \wedge x \geq 1$ is true because:

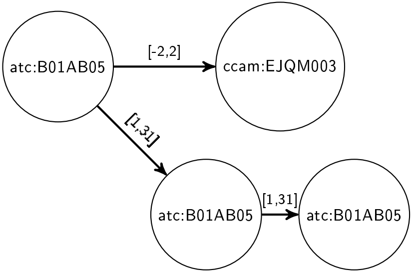
- $s(7) = b$
- $x - 7 = 4 - 7 = 3 \leq 3$
- $x - 7 = 4 - 7 = 3 \geq 1$

back

Use case: Deep Venous Thrombosis

Traitement par *enoxaparin* pour au moins 3 mois, *demarré* 2 jours avant ou après un echo-Doppler pour thrombose.

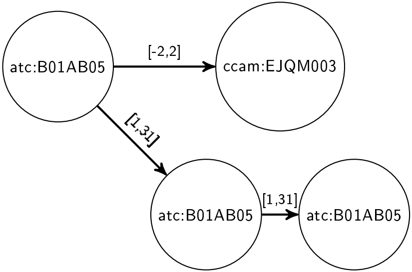
- *atc* : *B*01*AB*05: enoxaparin (anticoagulant injectable)
- *ccam* : *EJQM*003: echo Doppler



Use case: Deep Venous Thrombosis

Traitement par *enoxaparin* pour au moins 3 mois, *demarré* 2 jours avant ou après un echo-Doppler pour thrombose.

- *atc* : *B*01*AB*05: enoxaparin (anticoagulant injectable)
- *ccam* : *EJQM*003: echo Doppler



Limites

- *enoxaparin* is a type of anticoagulant ... *miss the others*
- *echo Doppler* is not the only procedure to diagnose DVT ...

⇒ need for more expressiveness about the event types

Use case: detecting patients having a Deep Venous Thrombosis (DVT) in the SNIIRAM

⇒ need for defining **high-level medical events** (*phenotypes*) from low level databases features

Deep Venous Thrombosis (DVT)

- Deep vein thrombosis (DVT) is a medical condition that occurs when a blood clot forms in a deep vein
- ICD-10 Codes: I802
- SNOMED CT Concept Id: 128053003

Detecting all patients with DVT in the SNIIRAM/DCIR

- ICD-10 codes are available only during hospitalisation and are not enough accurate (include suspicious DVT)
- Need to include additional features to better specify such a medical event (easily accessible features in SNIIRAM):
  - specific medical procedures: e.g. Doppler ultrasonography or CT-scan
  - antithrombotic deliveries / anticoagulant treatment

Use case: detecting patients having a Deep Venous Thrombosis (DVT) in the SNIIRAM

In clinical practice facing a suspicion of VTE physicians first prescribe antithrombotics and then confirm or not the diagnosis through specific medical procedures: e.g Doppler ultrasonography or CT-scan. Patients with suspected Pulmonary Embolism are often hospitalized whereas patients with suspected Deep Vein Thrombosis (DVT) are managed on an outpatient basis. On the one side, if the DVT suspicion is confirmed, antithrombotic deliveries continue for 3 to 12 months (once per month). Hence, the diagnosis (through the same medical procedures as above) is preceded or followed by initiating an anticoagulant treatment within a time window of at most 0 to 2 days. On the other side, Pulmonary Embolism suspicion leads to hospitalization during which medical procedures are performed to confirm the diagnosis and then anticoagulant delivery is observed only after the patient comes back home.

Requirements for such complex queries

- ontological reasoning / ontology mediated query answering
- temporal reasoning / temporal query