Declarative Process Specifications: Reasoning, Discovery, Monitoring

Claudio Di Ciccio[1][0000−0002−0955−6940] and Marco Montali[2][0000−0002−8021−3430]

[1] Sapienza University of Rome, Rome, Italy
claudio.diciccio@uniroma1.it
[2] Free University of Bozen-Bolzano, Bolzano, Italy
montali@inf.unibz.it

Abstract. The declarative specification of business processes is based upon the elicitation of behavioural rules that constrain the legal executions of the process. The carry-out of the process is up to the actors, who can vary the execution dynamics as long as they do not violate the constraints imposed by the declarative model. The constraints specify the conditions that require, permit or forbid the execution of activities, possibly depending on the occurrence (or absence) of other ones. In this chapter, we review the main techniques for process mining using declarative process specifications, which we call *declarative process mining*. In particular, we focus on three fundamental tasks of (1) reasoning on declarative process specifications, which is in turn instrumental to their (2) discovery from event logs and their (3) monitoring against running process executions to promptly detect violations. We ground our review on Declare, one of the most widely studied declarative process specification languages. Thanks to the fact that Declare can be formalized using temporal logics over finite traces, we exploit the automata-theoretic characterization of such logics as the core, unified algorithmic basis to tackle reasoning, discovery, and monitoring. We conclude the chapter with a discussion on recent advancements in declarative process mining, considering in particular multi-perspective extensions of the original approach.

1 Introduction

Finding a suitable balance between flexibility and control is a long-standing problem in the management of work processes [89]. Among the different approaches striving to achieve this balance, *flexibility by design* suggests to infuse flexibility in the process modeling language at hand. Declarative process modeling languages take this to the extreme: they support the specification of *what* are the relevant constraints on the temporal evolution of the process, without explicitly indicating *how* process instances should be routed to satisfy such constraints. In comparison with imperative approaches that produce "closed" representations (i.e., only those process executions explicitly foreseen in the model are allowed), declarative approaches yield "open" representations (i.e., every process execution is implicitly allowed, as long as it does not incur in the violation of some constraint).

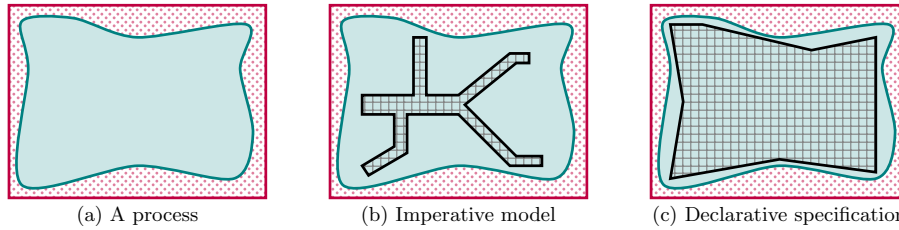(a) A process   (b) Imperative model   (c) Declarative specification

Fig. 1: Intuitive representation of the difference between imperative process models and declarative process specifications in the space of all execution traces. Diagram (a) represents a real process, which isolates the *allowed* (green, solid fill) behaviors from the *forbidden* (red, dotted fill) ones. Diagram (b) shows an *imperative process model* that stays within the boundaries of the process, but misses many allowed behaviors. Diagram (c) shows a *declarative process specification* that well approximates the boundaries of the process: it accepts only traces that are allowed by the process, and includes all the traces accepted by the imperative model in (b).

Figure 1 depicts an intuitive representation of the difference between classical imperative process models and declarative process specifications, considering execution traces that are forbidden by the real process, allowed by the real process, and captured by the designed process specification. Imperative models (such as those based on Petri nets and related formalisms) are suited to explicitly capture control-flow patterns like sequences, choices, concurrent sections, and loops. Those patterns, in turn, lend themselves to characterize a subset of the allowed traces, but struggle in covering the whole space of execution paths in the case of loosely structured, flexible processes. In other words, they favor control over flexibility. Contrariwise, declarative specifications strive to balance flexibility and control by attempting to characterize constraints that well-separate the allowed behaviors from the forbidden ones. In other words, declarative process specifications allow us to capture not only *what is expected* to occur, but also *what should not happen*. This helps in better approximating the boundaries of the real process, containing (and extending) those captured via imperative process models.

The idea of adopting a constraint-based, declarative approach to regulate dynamic systems has been originally brought forward in different communities: in data management, to express cascaded transactional updates [30]; in multiagent systems, to regulate agent interaction protocols [94]; and in business process management, to capture subprocesses that foresee loosely-coupled control-flow conditions on their activities [91]. This idea was further developed within BPM in consequent years, leading to a series of declarative, constraint-based process modeling languages, with two prominent exponents: Declare [83] and Dynamic Condition-Response Graphs [54]. Common to all such approaches is the usage of *linear temporal/dynamic logics* (i.e., temporal/dynamic logics for sequences of events) to formally describe specifications, and the exploitation of corresponding

reasoning mechanisms to tackle a variety of concrete tasks along the entire process lifecycle, from design and model analysis to runtime execution and data analysis.

In this chapter, we focus on *declarative process mining*, that is, process mining where the input or output models are specified using declarative, constraint-based languages. Concretely, we employ the DECLARE language, but all the presented ideas seamlessly apply any language that can be formalized using logics over *finite* traces [34], which are indeed at the core of DECLARE. Focusing on finite traces reflects the intuition that every process instance is expected to complete in a finite number of steps. This aspect has a significant impact on the corresponding operational techniques, as these logics admit an automata-theoretic characterization that is based on standard *finite-state automata* [34,31], instead of automata on infinite structures, which are needed when such logics are interpreted over infinite traces.

Leveraging automata-based techniques paired with suitable measures relating traces, events and constraints, we review three interconnected fundamental declarative process mining tasks:

Reasoning – to uncover relationships among different constraints, and check key properties of DECLARE specifications;
Discovery – to extract a DECLARE specification that suitably characterizes the traces contain in an event log;
Monitoring – to provide operational decision support [67] by checking at run-time whether a running process execution satisfies a DECLARE specification, promptly detecting and reporting violations.

All the presented techniques are integrated in the MINERful process discovery technique [3] [43] and the RuM toolkit[4] [8].

The chapter is organized as follows. Section 2 introduces the declarative process specification language DECLARE alongside a running example to which we will refer throughout the remainder of the chapter. Section 3 provides the fundamental notions upon which the core techniques for reasoning, discovery and monitoring on declarative specifications are based. We define the formal semantics of DECLARE and discuss the core reasoning tasks for declarative specifications in Section 4. Section 5 explains the core notions of declarative process discovery and monitoring. Section 6 discusses the latest advances in the field of declarative process specification mining. Finally, Section 7 concludes this chapter with final remarks and a summary of the core concepts illustrated herein.

2 Declare: A Gentle Introduction

DECLARE is a language and graphical notation providing an extendible repertoire of templates to formulate constraints. The origin of the approach traces back to the PhD work by Pesic [81], and the parallel and consequent study in the PhD work by

---

[3] https://github.com/cdc08x/MINERful
[4] https://rulemining.org

Table 1: A set of DECLARE constraints among those that are typically used for process mining, with their textual description, graphical notation, and examples fulfilling or violating them.

Constraint	Explanation	Examples				Notation
Existence constraints						
INIT(a)	a is the *first* to occur	✓ ⟨a, c, c⟩	✓ ⟨a, b, a, c⟩	× ⟨c, c⟩	× ⟨b, a, c⟩	
ATLEASTONE(a)	a occurs at least *once*	✓ ⟨b, c, a, c⟩	✓ ⟨b, c, a, a, c⟩	× ⟨b, c, c⟩	× ⟨c⟩	
ATMOSTONE(a)	a occurs at most *once*	✓ ⟨b, c, c⟩	✓ ⟨b, c, a, c⟩	× ⟨b, c, a, a, c⟩	× ⟨b, c, a, c, a, a⟩	
END(a)	a is the *last* to occur	✓ ⟨b, c, a⟩	✓ ⟨b, a, c, a⟩	× ⟨b, c⟩	× ⟨b, a, c⟩	
Relation constraints						
RESPONDEDEXISTENCE(a, b)	If a occurs in the trace, then b occurs as well	✓ ⟨b, c, a, a, c⟩	✓ ⟨b, c, c⟩	× ⟨c, a, a, c⟩	× ⟨a, c, c⟩	
RESPONSE(a, b)	If a occurs, then b occurs after a	✓ ⟨c, a, a, c, b⟩	✓ ⟨b, c, c⟩	× ⟨c, a, a, c⟩	× ⟨b, a, c, c⟩	
ALTERNATERESPONSE(a, b)	Each time a occurs, then b occurs afterwards, and no other a recurs in between	✓ ⟨c, a, c, b⟩	✓ ⟨a, b, c, a, c, b⟩	× ⟨c, a, a, c, b⟩	× ⟨b, a, c, a, c, b⟩	
CHAINRESPONSE(a, b)	Each time a occurs, then b occurs immediately afterwards	✓ ⟨c, a, b, b⟩	✓ ⟨a, b, c, a, b⟩	× ⟨c, a, c, b⟩	× ⟨b, c, a⟩	
PRECEDENCE(a, b)	b occurs only if preceded by a	✓ ⟨c, a, c, b, b⟩	✓ ⟨a, c, c⟩	× ⟨c, c, b, b⟩	× ⟨b, a, c, c⟩	
ALTERNATEPRECEDENCE(a, b)	Each time b occurs, it is preceded by a and no other b can recur in between	✓ ⟨c, a, c, b, a⟩	✓ ⟨a, b, c, a, a, c, b⟩	× ⟨c, a, c, b, b, a⟩	× ⟨a, b, b, a, b, c, b⟩	
CHAINPRECEDENCE(a, b)	Each time b occurs, then a occurs immediately beforehand	✓ ⟨a, b, c, a⟩	✓ ⟨a, b, a, a, b, c⟩	× ⟨b, c, a⟩	× ⟨b, a, a, c, b⟩	
Mutual relation constraints						
COEXISTENCE(a, b)	If b occurs, then a occurs, and vice versa	✓ ⟨c, a, c, b, b⟩	✓ ⟨b, c, c, a⟩	× ⟨c, a, c⟩	× ⟨b, c, c⟩	
SUCCESSION(a, b)	a occurs if and only if it is followed by b	✓ ⟨c, a, c, b, b⟩	✓ ⟨a, c, c, b⟩	× ⟨b, a, c⟩	× ⟨b, c, c, a⟩	
ALTERNATESUCCESSION(a, b)	a and b if and only if the latter follows the former, and they alternate each other in the trace	✓ ⟨c, a, c, b, a, b⟩	✓ ⟨a, b, c, a, b, c⟩	× ⟨c, a, a, c, b, b⟩	× ⟨b, a, c⟩	
CHAINSUCCESSION(a, b)	a and b occur if and only if the latter immediately follows the former	✓ ⟨c, a, b, a, b⟩	✓ ⟨c, c, c⟩	× ⟨c, a, c, b⟩	× ⟨c, b, a, c⟩	
Negative relation constraints						
NOTCOEXISTENCE(a, b)	a and b never occur together	✓ ⟨c, c, c, b, b, b⟩	✓ ⟨c, c, a, c⟩	× ⟨a, c, c, b, b⟩	× ⟨b, c, a, c⟩	
NOTSUCCESSION(a, b)	b cannot occur after a	✓ ⟨b, b, c, a, a⟩	✓ ⟨c, b, b, c, a⟩	× ⟨a, a, c, b, b⟩	× ⟨a, b, b⟩	
NOTCHAINSUCCESSION(a, b)	a and b cannot occur contiguously	✓ ⟨a, c, b, a, c, b⟩	✓ ⟨b, b, a, a⟩	× ⟨a, b, c, a, b⟩	× ⟨c, a, b, c⟩	

Montali [73]. Notably, Declare actually stems from three initial lines of research, respectively focused on the declarative specification of business processes (cf. the ConDec language [82]), service choreographies (cf. the DecSerFlow language [5,76]), and clinical guidelines (cf. the CigDec language [78]). These lines were then unified into a single research thread. The term Declare was used for the first time in [83].
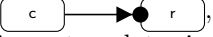
Table 1 shows a set of Declare constraints we use throughout this chapter. The whole, core set of Declare templates has been inspired by a catalogue of temporal logic patterns used in model checking for a variety of dynamic systems from different application domains [46].
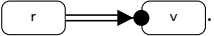
Formally, we define a declarative process specification as follows.

Definition 1 (Declarative process specification). *A declarative process specification is a tuple* $DS = (\text{Rep}, \text{Act}, K)$ *where*
- Rep *is a finite non-empty set of* templates, *where each template is a predicate* $\text{K}(x_1, \ldots, x_m) \in \text{Rep}$ *on variables* $x_1, \ldots, x_m$ *(with* $m \in \mathbb{N}$ *the* arity *of* K*),*
- Act *is a finite non-empty set of* activities,
- K *is a finite set of* constraints, *namely pairs* $(\text{K}(x_1, \ldots, x_m), \kappa)$ *where* $\text{K}(x_1, \ldots, x_m)$ *is a template from* Rep, *and* κ *is a mapping that, for every* $i \in \{1, \ldots, m\}$ *assigns variable* x_i *with an activity* $\kappa(x_i) = a_i \in \text{Act}$; *we compactly denote such a constraint with* $\text{K}(a_1, \ldots, a_m)$. ◁

Example 1 (A Declare process specification). Figure 2 portrays an example of declarative specification for the admission process of an international Bachelor's program. This example considers the Declare repertoire of templates. The process begins with the creation of an account in the university portal (henceforth, c). To specify that c is the initial task, we write Init(c), graphically depicted with the Init label in the tag on top of the activity box. Init is a unary template and Init(c) assigns its variable with activity c. Unary templates in Declare are also known as *existence* templates. We indicate that not more than one account can be created per process run with AtMostOne(c). In the diagram, it is indicated with the 0..1 label in the tag.

To register for a selection round (r), an account must have been created before Precedence(c, r). We express this rule with Precedence(c, r), which is a shorthand notation for $(\text{Precedence}(x_1, x_2), \{(x_1, \text{c}), (x_2, \text{r})\})$. Precedence is a binary template and Precedence(c, r), graphically depicted as , assigns c and r to its first and second variable, respectively. Binary templates in Declare are commonly named as *relation* templates.

Every registration to a selection round (r) gives access to a uniquely corresponding evaluation phase (v). After r, v eventually follows and no other registrations are allowed until v completes. We write AlternateResponse(r, v), graphically depicted as . The evaluation requires r to be completed before and v will not recur unless a new registration is issued: AlternatePrecedence(r, v), . Typically, if both AlternateResponse(r, v) and AlternatePrecedence(r, v) hold true, we
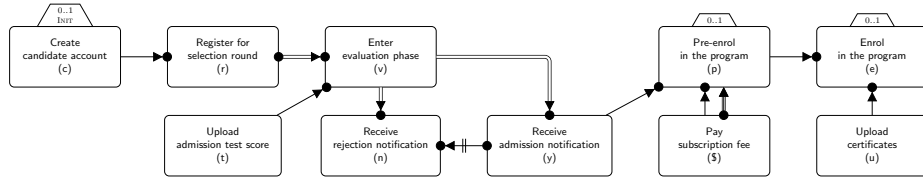
Fig. 2: The DECLARE map of the admission process at a university.

compactly represent them jointly with the *mutual* relation constraint
ALTERNATESUCCESSION(r, v) [r ●━━▶● v]. An admission test score has
to be uploaded in the platform to access the evaluation phase: PRECEDENCE(t, v).
Evaluation phases are necessary for the committee to return rejections (n)
and notifications of admission (y), thus ALTERNATEPRECEDENCE(v, y) and
ALTERNATEPRECEDENCE(v, n) hold.

After the admission has been notified, the candidate will not receive a re-
jection any longer – NOTRESPONSE(y, n), drawn in Fig. 2 as [y ●━━∦▶● n].
NOTRESPONSE(y, n) falls under the category of the *negative* relation constraints,
as the occurrence of y *disables* n in the remainder of the process execution.

Only if candidates receive a notification of admission, they will be en-
titled to pre-enrol in the program (PRECEDENCE(y, p)). The candidates are
considered as pre-enrolled immediately after they pay the subscription fee
(CHAINRESPONSE($, p), shown as follows in the diagram: [$ ●═══▶ p]).
Also, candidates cannot be considered as pre-enrolled if they have not paid the
subscription fee: PRECEDENCE($, p). Not more than one pre-enrolment is allowed
per candidate: ATMOSTONE(p). To enrol in the program (e), the candidate must
have pre-enrolled – PRECEDENCE(p, e) – and uploaded the necessary school and
language certificates – PRECEDENCE(u, e).

So far, we have been attaching an informal semantics to DECLARE and
its templates. In the next section, we provide a more systematic and formal
characterization.

3   Formal background

Considering that DECLARE templates have been originally defined starting from
a catalogue of Linear Temporal Logic (LTL) patterns [46], it is not surprising
that temporal logics have been used to characterize the semantics of DECLARE
since the very beginning. However, the fact that DECLARE specifications are
interpreted over finite-length executions calls for the use of Linear Temporal Logic
on Finite Traces (LTL_f) [34]. This indeed leads to a setting that is radically
different, both semantically and algorithmically, from the traditional one where
formulae are interpreted using LTL over infinite, recurring behaviors [33].

A complete formalization of DECLARE templates, also including an alternative
formalization using a logic programming-based approach, can be found in [74]. It

was later refined in [33]. In his PhD thesis, Di Ciccio was the first to provide a semantics based on regular expressions [39]. These two themes were later unified in [32], leading to a richer framework that is able to declaratively capture constraints and metaconstraints, that is, constraints predicating over the possible/certain satisfaction and violation of other constraints.

In this section, we provide some necessary background on LTL_f and its extension with past-tense temporal operators, as well as on the automata-theoretic characterization for this logic. We then use this framework to formalize DECLARE and reason automatically on DECLARE specifications. Thereupon, we reflect upon the most recent advances of research in attempting at capturing not only the formal semantics of constraints, but also how they pragmatically interact with relevant events.

3.1   Linear Temporal Logic on finite traces

LTL_f has the same syntax of LTL [86], but is interpreted on finite traces. In this chapter, in particular, we consider the LTL dialect including past modalities [62] for declarative process specifications as in [22].

From now on, we fix a finite set Σ representing an alphabet of propositional symbols describing (names of) activities available in the domain under study. A (finite) *trace* $t = \langle a_1, \ldots, a_n \rangle \in \Sigma$ of length $|t| = n$ is a finite sequence of activities, where the presence of activity a_i at instant i of the trace represents an *event* that witnesses the occurrence of a_i at instant i – which we also write $t(i) = a_i$. Notice that *at each instant we assume that one and only one activity occurs*. Using standard notation from regular expressions, the set Σ^* denotes the overall set of traces whose constitutive events refer to activities in Σ.

Definition 2 (Syntax of LTL_f). *Well-formed formulae are built from Σ, the unary temporal operators $\bigcirc$ (next) and $\ominus$ (yesterday), and the binary temporal operators $\mathbf{U}$ (until) and $\mathbf{S}$ (since) as follows:*

$$\varphi ::= \mathsf{a} \mid (\neg\varphi) \mid (\varphi_1 \wedge \varphi_2) \mid (\bigcirc\varphi) \mid (\varphi_1 \ \mathbf{U} \ \varphi_2) \mid (\ominus\varphi) \mid (\varphi_1 \ \mathbf{S} \ \varphi_2)$$

where $\mathsf{a} \in \Sigma$. ◁

Definition 3 (Semantics of LTL_f, satisfaction, validity, entailment).
An LTL_f formula φ is inductively satisfied *in some instant i $(1 \leq i \leq n)$ of a trace t of length $n \in \mathbb{N}$, written $t, i \vDash \varphi$, if the following holds:*
- $t, i \vDash \mathsf{a}$ *iff $t(i)$ is assigned with a;*
- $t, i \vDash \neg\varphi$ *iff $t, i \nvDash \varphi$;*
- $t, i \vDash \varphi_1 \wedge \varphi_2$ *iff $t, i \vDash \varphi_1$ and $t, i \vDash \varphi_2$;*
- $t, i \vDash \bigcirc\varphi$ *iff $i < n$ and $t, i + 1 \vDash \varphi$;*
- $t, i \vDash \ominus\varphi$ *iff $i > 1$ and $t, i - 1 \vDash \varphi$;*
- $t, i \vDash \varphi_1 \ \mathbf{U} \ \varphi_2$ *iff $t, j \vDash \varphi_2$ with $i \leq j \leq n$, and $t, k \vDash \varphi_1$ for all k s.t. $i \leq k < j$;*
- $t, i \vDash \varphi_1 \ \mathbf{S} \ \varphi_2$ *iff $t, j \vDash \varphi_2$ with $1 \leq j \leq i$, and $t, k \vDash \varphi_1$ for all k s.t. $j < k \leq i$.*

A formula φ is satisfied by *a trace t (equivalently, t* satisfies φ), written $t \vDash \varphi$, *iff* $t, 1 \vDash \varphi$. *A formula* φ *is: (i)* satisfiable *if it has a satisfying trace from* Σ^*; *(ii)* valid *if every trace in* Σ^* *satisfies it. A formula* φ_1 entails *formula* φ_2, *written* $\varphi_1 \models \varphi_2$, *if, for every trace t of length* $n \in \mathbb{N}$ *and every i s.t.* $1 \leq i \leq n$, *if* $t, i \models \varphi$ *then* $t, i \models \psi$. ◁

Since LTL_f is closed under negation, it is easy to see that a formula φ is valid if and only if $\neg\varphi$ is unsatisfiable.

It is worth noting that, in LTL_f, the next operator is interpreted as the so-called *strong* next: $\bigcirc \varphi$ requires that the next instant exists within the trace, and that at such next instant φ holds. This has an important consequence: differently from LTL, in LTL_f formula $\neg \bigcirc \varphi$ is *not* equivalent to $\bigcirc \neg\varphi$. This is because $\neg \bigcirc \varphi$ is true in an instant of a finite trace either when that instant has no successor, or the next instant exists and in such a next instant φ does not hold. More on this can be found in [33].

From the basic operators above, the following can be derived:
- Classical boolean abbreviations **true**, **false**, $\vee$, $\rightarrow$;
- Constant **end** $\equiv \neg \bigcirc$ **true**, denoting the last instant of a trace;
- Constant **start** $\equiv \neg \ominus$ **true**, denoting the first instant of a trace;
- $\Diamond \varphi \equiv$ **true** $\mathbf{U}$ φ indicating that φ eventually holds true in the trace (hence, before or at **end**);
- φ_1 $\mathbf{W}$ $\varphi_2 \equiv (\varphi_1$ $\mathbf{U}$ $\varphi_2) \vee \square\varphi_1$, which relaxes $\mathbf{U}$ as φ_2 may never hold true;
- $\Diamondblack \varphi \equiv$ **true** $\mathbf{S}$ φ indicating that φ holds true at some instant before the current one (i.e., after **start** in the trace);
- $\square \varphi \equiv \neg \Diamond \neg\varphi$ indicating that φ holds true from the current instant till **end**;
- $\boxminus \varphi \equiv \neg \Diamondblack \neg\varphi$ indicating that φ holds true from **start** to the current instant.

Example 2. Let $t = \langle a, b, b, c, d, e \rangle$ be a trace and φ_1, φ_2 and φ_3 three LTL_f formulae defined as follows: $\varphi_1 \doteq d$; $\varphi_2 \doteq \Diamond b$; $\varphi_3 \doteq \square(b \rightarrow \Diamond d)$. We have that $t, 1 \nvDash \varphi_1$ whereas $t, 5 \vDash \varphi_1$; $t, 1 \vDash \varphi_2$ whereas $t, 5 \nvDash \varphi_2$; $t, 1 \vDash \varphi_3$ and $t, 5 \vDash \varphi_3$ (in fact, $t, i \vDash \varphi_3$ for any instant $1 \leq i \leq n$). ◁

3.2   Finite-state automata

One of the central features of LTL_f is that a finite state automaton (FSA) [26] $\mathscr{A}(\varphi)$ can be computed such that for every trace t we have that $t \vDash \varphi$ iff t is in the language recognized by $\mathscr{A}(\varphi)$, as illustrated in [34,32,22,41]. We include the main notions next, recalling that focusing on deterministic FSAs is without loss of generality, as over finite traces every non-deterministic FSAs can be determinized [55].

Definition 4 (Finite state automaton (FSA)). *A (deterministic) finite state automaton (FSA) is a tuple* $A = (\Sigma, S, \delta, s_0, S_F)$, *where:*
- Σ *is a finite set of symbols;*
- S *is a finite non-empty set of states;*
- $\delta : S \times \Sigma \rightarrow S$ *is the* transition function, *i.e., a partial function that, given a starting state and a (labeled) transition, returns the target state;*
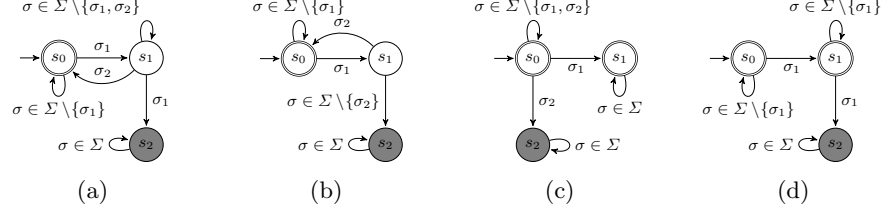
Fig. 3: Examples of constraint FSAs.

- s_0 is the initial state;
- $S_F \subseteq S$ is the set of final (accepting) states $s_F \in S_F$ ◁

In the remainder of the chapter, we assume that δ is left-total and surjective on $S \setminus \{s_0\}$, that is, the transition function is defined for every state and symbol, and every state is on a path from the initial one – with the possible exception of the initial state itself. An FSAs that is left-total is called *untrimmed*. Notice that these two requirements are without loss of generality: every FSA can be converted into an equivalent FSA that is left-total and surjective. In particular, to make an FSAs untrimmed, it is sufficient to: *(i)* introduce a non-final trap state $s_\perp$; *(ii)* for every state s and symbol a' such that $\delta(s, a')$ is *not* defined, enforce $\delta(s, a') = s_\perp$; *(iii)* connect $s_\perp$ to itself for every symbol, setting $\delta(s_\perp, a) = s_\perp$ for every $a \in \Sigma$.

Example 3. Figure 3 depicts four FSAs. States are represented as circles and transitions as arrows. Accepting states are decorated with a double line. The initial state is indicated with a single, unlabeled incoming arc. For instance, Fig. 3(a) is such that $\Sigma \supseteq \{\sigma_1, \sigma_2\}$, $S = \{s_0, s_1, s_2\}$, $S_F = \{s_0\}$, $\delta(s_0, \sigma_1) = s_1$ and $\delta(s_1, \sigma_1) = s_2$.   ◁

Definition 5 (Runs and traces of an FSA). *Let $A = (\Sigma, S, \delta, s_0, S_F)$ be an FSA as per Definition 4. A computation π of A is a finite sequence alternating states and activities $s_0 \xrightarrow{\sigma_0} \ldots \xrightarrow{\sigma_{n-1}} s_n$ that starts from the initial state s_0 is such that for every $0 \leq i < n$, we have $\delta(s_i, \sigma_i) = s_{i+1}$. If π terminates in a final state, that is, $s_n \in S_F$, then it is a* run, *and* induces *a corresponding trace $\sigma_0, \ldots, \sigma_{n-1}$ over Σ^* obtained from π by only keeping the symbols that label the transitions.*   ◁

Example 4. In Fig. 3(a), $\pi_1 = s_0 \xrightarrow{\sigma_1} s_1$, $\pi_2 = s_0 \xrightarrow{\sigma_2} s_0 \xrightarrow{\sigma_1} s_1 \xrightarrow{\sigma_1} s_2$, and $\pi_3 = s_0 \xrightarrow{\sigma_1} s_1 \xrightarrow{\sigma_2} s_2 \xrightarrow{\sigma_1} s_0$ are three examples of computations. However, only π_3 is a run because $s_0 \in S_F$ whereas $s_1, s_2 \notin S_F$. Notice that, in Fig. 3, we additionally highlight with a grey background colour those states that cannot be in a step of a run – that is, from which accepting states cannot be reached (e.g., s_2 in Fig. 3(a)).   ◁

Definition 6 (Accepted trace, language of an FSA). *A trace $t \in \Sigma^*$ is accepted by FSA $A = (\Sigma, S, \delta, s_0, s_F)$ if there is a run of A inducing t. The language $\mathscr{L}(A)$ of A is the set of traces accepted by A.*   ◁

(a) Identity element   (b) Absorbing element

Fig. 4: Finite state automata acting as identity element and absorbing element for the automata cross-product operation.

Example 5. For the FSA in Fig. 3(a), the language contains the trace $t_1 = \langle \sigma_1, \sigma_2, \sigma_1 \rangle$, since a run exists over this sequence of labels (i.e., π_3 above), whereas $t_2 = \langle \sigma_2, \sigma_1 \rangle$ is not part of the language.   ◁

Automata product. FSAs are closed under the (synchronous) product operation $\times$ [87]. The (cross-)product $A \times A'$ of two FSAs A and A' is an FSA that accepts the intersection of languages (sets of accepted traces) of each operand: $\mathscr{L}(A \times A') = \mathscr{L}(A) \bigcap \mathscr{L}(A')$. It is defined as follows.

Definition 7 (Automata product). *The* product FSA *of two FSAs* $A = (\Sigma, S, \delta, s_0, S_F)$ *and* $A' = (\Sigma, S', \delta', s_0', S_F')$ *over the same alphabet* Σ *is the FSA* $A \times A' = (\Sigma, S^\times, \delta^\times, s_0^\times, S_F^\times)$, *where the set* $S^\times \subseteq S \times S'$ *of states (obtained from the cartesian product of the states in* A *and* A'*), its initial state* $s_0^\times$*, its final states* $S_F^\times$*, and the transition function* $\delta^\times$*, are defined by simultaneous induction as follows:*

- $s_0^\times = \langle s_0, s_0' \rangle \in S^\times$;
- *For every state* $\langle s_1, s_1' \rangle \in S^\times$*, state* $s_2 \in S$*, state* $s_2' \in S'$*, and label* $\ell \in \Sigma$*, if* $\delta(s_1, \ell) = s_2$ *and* $\delta'(s_1', \ell) = s_2'$ *then: (i)* $\langle s_2, s_2' \rangle \in S^\times$*, (ii)* $\delta^\times(\langle s_1, s_1' \rangle, \ell) = \langle s_2, s_2' \rangle$*, (iii) if* $s_2 \in S_F$ *and* $s_2' \in S_F'$*, then* $\langle s_2, s_2' \rangle \in S_F^\times$*.*
- *Nothing else is in* $S_F^\times$*,* $S^\times$*, and* $\delta^\times$*.*   ◁

Notice that the FSA constructed with Definition 7 can be manipulated using language-preserving automata operations, such as in particular *minimization* [55].

The product operation $\times$ is commutative and associative. The identity element for $\times$ over alphabet Σ is $A^{\mathrm{I}} = (\Sigma, \{s_0\}, s_0, \{s_0\} \times \Sigma \times \{s_0\}, \{s_0\})$ – depicted in Fig. 4(a). It accepts all traces over Σ: $\mathscr{L}(A^{\mathrm{I}}) = \mathbb{P}(\Sigma^*)$ as any sequence of transitions labeled by symbols in Σ corresponds to a run for A^{I}. The absorbing element is $A^\emptyset = (\Sigma, \{s_0\}, s_0, \{s_0\} \times \Sigma \times \{s_0\}, \emptyset)$ and is illustrated in Fig. 4(b). It does not accept any trace at all: $\mathscr{L}(A^\emptyset) = \emptyset$ as any sequence of transitions labeled by symbols in Σ corresponds to a computation ending in a non-accepting state.

4 Reasoning

Equipped with the notions acquired thus far, we can now discuss the core reasoning tasks that are associated to declarative process specifications. To this end, we begin this section by describing the semantics of DECLARE in detail.

Table 2: Semantics of some DECLARE constraints.

Template	LTL_f expression [34,22]	Activation	Target
Existence constraints			
ATLEASTONE(x)	$\square\,(\mathbf{start} \to \Diamond\, x)$	**start**	$\Diamond\, x$
ATMOSTONE(x)	$\square(x \to \neg \bigcirc \Diamond\, x)$	x	$\neg \bigcirc \Diamond\, x$
INIT(x)	$\square\,(\mathbf{start} \to x)$	**start**	x
END(x)	$\square\,(\mathbf{end} \to x)$	**end**	x
Relation constraints			
RESPONDEDEXISTENCE(x,y)	$\square\,(x \to \Diamond\, y \vee \ominus\!\!\!\diagdown\, y)$	x	$\Diamond\, y \vee \ominus\!\!\!\diagdown\, y$
RESPONSE(x,y)	$\square\,(x \to \Diamond\, y)$	x	$\Diamond\, y$
ALTERNATERESPONSE(x,y)	$\square\,(x \to \bigcirc(\neg x \;\mathbf{U}\; y))$	x	$\bigcirc(\neg x \;\mathbf{U}\; y)$
CHAINRESPONSE(x,y)	$\square\,(x \to \bigcirc y)$	x	$\bigcirc y$
PRECEDENCE(x,y)	$\square\,(y \to \ominus\!\!\!\diagdown\, x)$	y	$\ominus\!\!\!\diagdown\, x$
ALTERNATEPRECEDENCE(x,y)	$\square\,(y \to \ominus(\neg y \;\mathbf{S}\; x))$	y	$\ominus(\neg y \;\mathbf{S}\; x)$
CHAINPRECEDENCE(x,y)	$\square\,(y \to \ominus x)$	y	$\ominus x$
Negative relation constraints			
NOTRESPONDEDEXISTENCE(x,y)	$\square(x \to (\square\, \neg y \wedge \boxminus \neg y))$	x	$\square\, \neg y \wedge \boxminus \neg y$
NOTRESPONSE(x,y)	$\square(x \to \square\, \neg y)$	x	$\square\, \neg y$
NOTCHAINRESPONSE(x,y)	$\square(x \to \neg \bigcirc y)$	x	$\neg \bigcirc y$
NOTPRECEDENCE(x,y)	$\square(y \to \boxminus \neg x)$	y	$\boxminus \neg x$
NOTCHAINPRECEDENCE(y,x)	$\square(y \to \neg \ominus x)$	y	$\neg \ominus x$

4.1 Semantics of Declare

The semantics of a DECLARE template $\text{K}(x_1,\ldots,x_m)$ is given as an LTL_f formula $\varphi_{\text{K}(x_1,\ldots,x_m)}$ defined over variables $x_1,\ldots,x_m$ instead of activities. Given the free variables x and y, e.g., RESPONSE(x,y) corresponds to $\square(x \to \Diamond\, y)$, witnessing that whenever x occurs, then y is expected to occur at some later instant. Table 2 shows the LTL_f formulae of some templates of the DECLARE repertoire. The formalization of a constraint is then obtained by grounding the LTL_f formula of its template.

Definition 8 (Constraint formula, satisfying trace). *The* formula *of constraint* $\text{K}(a_1,\ldots,a_m)$, *written* $\varphi_{\text{K}(a_1,\ldots,a_m)}$, *is the* LTL_f *formula obtained from* $\varphi_{\text{K}(x_1,\ldots,x_m)}$ *by replacing* x_i *with* a_i *for each* $1 \leq i \leq m$. *A trace* t satisfies $\text{K}(a_1,\ldots,a_m)$ *if* $t \models \varphi_{\text{K}(a_1,\ldots,a_m)}$; *otherwise, we say that* t violates $\text{K}(a_1,\ldots,a_m)$. ◁

Example 6. Considering Table 2, we have $\varphi_{\text{RESPONSE}(\mathsf{a},\mathsf{b})} = \square(\mathsf{a} \to \Diamond\,\mathsf{b})$, and $\varphi_{\text{RESPONSE}(\mathsf{b},\mathsf{c})} = \square(\mathsf{b} \to \Diamond\,\mathsf{c})$. Traces $\langle \mathsf{b}\rangle$ and $\langle \mathsf{a},\mathsf{b},\mathsf{a},\mathsf{c},\mathsf{b}\rangle$ satisfy RESPONSE(a,b), while $\langle \mathsf{a}\rangle$ and $\langle \mathsf{a},\mathsf{b},\mathsf{a},\mathsf{a},\mathsf{c}\rangle$ do not. ◁

A DECLARE specification is then formalized by conjoining all its constraint formulae, thus obtaining a direct, declarative notion of *model trace*, that is, a trace that is accepted by the specification.

Definition 9 (Specification formula, model trace). *The* formula *of* DE-CLARE *specification* $\mathrm{DS} = (\mathrm{REP}, \mathrm{Act}, K)$*, written* φ_{DS}*, is the* LTL_f *formula* $\bigwedge_{\mathrm{K} \in K} \varphi_{\mathrm{K}}$*. A trace* $t \in \mathrm{Act}^*$ *is a* model trace *of* DS *if* $t \models \varphi_{\mathrm{DS}}$*; in this case, we say that* t *is* accepted by DS*, otherwise that* t *is* rejected by DS*.*   ◁

Constructing constraint and specification formulae is, however, not enough. When one reads $\square(\mathsf{a} \to \Diamond\, \mathsf{b})$ following the textual description given above, the formula gets intepreted as "whenever a occurs, then b is expected to occur at some later instant". This formulation intuitively hints at the fact that the occurrence of a *activates* the RESPONSE(a, b) constraint, requiring the *target* b to occur. In turn, we get that a trace not containing any occurrence of a is *less interesting* than a trace containing occurrences of a, each followed by one or more occurrences of b: even though both traces satisfy RESPONSE(a, b), the first trace never "interacts" with RESPONSE(a, b), while the second does. This relates to the notion of vacuous satisfaction in LTL [56] and that of interestingness of satisfaction in LTL_f [42].

The point is, all such considerations are not captured by the formula $\square(\mathsf{a} \to \Diamond\, \mathsf{b})$, but are related to pragmatic interpretation of how it relates to traces. To see this aspect, let us consider that we can equivalently express the formula above as $\square \neg\mathsf{a} \vee \Diamond(\mathsf{b} \wedge \square\, \neg\mathsf{a})$, which now reads as follows: "Either a never happens at all, or there is some occurrence of b after which a never happens". This equivalent reformulation does not put into evidence the activation or the target.

This problem can be tackled in two possible ways. One option is to attempt at an automated approach where activation, target, and interesting satisfaction are semantically, implicitly characterized once and for all at the logical level; this is the route followed in [42]. The main drawback of this approach is that the user cannot intervene at all in deciding how to fine-tune the activation and target conditions. An alternative possibility is instead to ask the user to explicitly indicate, together with the LTL_f formula φ of the template, also two related LTL_f formulae expressing activation and target conditions for φ. This latter approach, implicitly adopted in [75] and then explicitly formalized in [22], gives more control to the user on how to *pragmatically* interpret constraints. We follow this latter approach.

Intuitively, the *activation* of a constraint is a triggering condition that, once made true, expects that the *target* condition is satisfied by the process execution. Contrariwise, if the constraint is not activated, the satisfaction of the target is not enforced. All in all, to properly constitute an activation-target pair for an LTL_f formula φ, we need them to satisfy the condition that whenever the current instant is such that the activation is satisfied, φ must behave equivalently to the target (thus requiring its satisfaction). This is formally captured as follows.

Definition 10 (Activation and target of a constraint). *The*   activation *and* target *of a constraint* K *over activities* Act *are two* LTL_f *formulae* $_\square\mathrm{K}$ *and* $\mathrm{K}_\blacktriangleright$ *such that for every trace* $t \in \mathrm{Act}^*$ *we have that:*

$$t \vDash \varphi_{\mathrm{K}} \quad \textit{iff} \quad t \vDash \square\left(_\square\mathrm{K} \to (\mathrm{K}_\blacktriangleright)\right)$$

(a) ALT.RESP.(r, v)   (b) CHN.RESP.($, p)   (c) PREC.(u, e)   (d) ATMOSTONE(p)
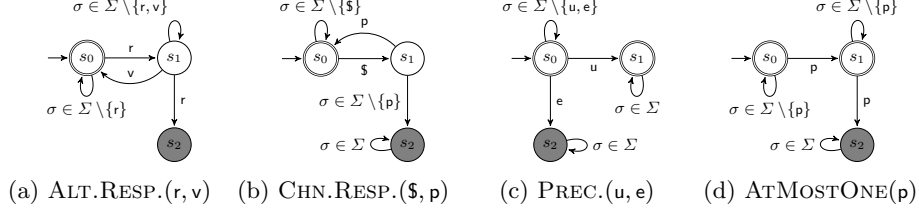
Fig. 5: Example FSAs of DECLARE constraints.

Table 2 shows activations and targets for each constraint, inspired by the work of Cecconi et al. [22]. In the next example, we explain the rationale behind some of the constraint formulations in the table.

Example 7. Consider CHAINRESPONSE($, p), dictating that whenever $ occurs, then p is the activity occurring next. We have $\varphi_{\text{CHAINRESPONSE}(\$,p)} = \Box(\$ \rightarrow \bigcirc p)$. Then, by Definition 10, we can directly fix $_\Box\text{CHAINRESPONSE}(\$, p) = \$$, and $\text{CHAINRESPONSE}(\$, p)_\blacktriangleright = \bigcirc p$, respectively witnessing that every occurrence of $ triggers the constraint, with a target requiring the consequent execution of p in the next instant. Similarly, for PRECEDENCE($, p) we have $\varphi_{\text{PRECEDENCE}(\$,p)} = \Box(p \rightarrow \diamondsuit \$)$, and in turn, by Definition 10, $_\Box\varphi_{\text{PRECEDENCE}(\$,p)} = p$ and $\varphi_{\text{PRECEDENCE}(\$,p)_\blacktriangleright} = \diamondsuit \$$. The case of ATMOSTONE(p) is also similar. In this case, $\varphi_{\text{ATMOSTONE}(p)}$ formalizes that p cannot occur twice, which in LTL_f can be directly captured by $\neg \diamondsuit(p \wedge \bigcirc \diamondsuit p)$. This is logically equivalent to $\Box(p \rightarrow \neg \bigcirc \diamondsuit p)$, which directly yields $_\Box\text{ATMOSTONE}(p) = p$ and $\text{ATMOSTONE}(p)_\blacktriangleright = \neg \bigcirc \diamondsuit p$.

A quite different situation holds instead for the other *existence* constraints. Take, for example, ATLEASTONE(a), requiring that a occurs at least once in the execution. This can be directly encoded in LTL_f as $\diamondsuit a$. This formulation, however, does not help to individuate the activation and target of the constraint. Intuitively, we may disambiguate this by capturing that since the constraint requires the presence of a from the very beginning of the execution, the constraint is indeed activated at the beginning, i.e., when **start** holds, imposing the satisfaction of the target $\diamondsuit a$. This intuition is backed up by Definition 10, using the semantics of **start** and noticing the following logical equivalences:

$$\diamondsuit a = \textbf{start} \rightarrow \diamondsuit a = \Box(\textbf{start} \rightarrow \diamondsuit a)$$

This explains why the latter formulation is employed in Table 2.   ◁

Declarative Constraints as FSAs. Crucial for our techniques is that every LTL_f formula φ can be encoded into a corresponding FSA (in the sense of Definition 4) A_φ that recognizes all and only those traces that satisfy the formula. This can be done through different algorithmic techniques. A direct approach that transforms an input formula into a non-deterministic FSAs is presented in [33,32]; notice that the so-obtained FSAs can then be determinized and minimized using standard techniques [55,100]. A fortiori, given a DECLARE specification $\text{DS} = (\text{REP}, \text{Act}, K)$, we proceed as follows:

- We pair each constraint $\kappa \in K$ to a corresponding, so-called *local automaton* A_κ. This automaton is the FSA A_{φ_κ} of the constraint formula φ_κ, and is used to characterize all and only those traces that satisfy κ;
- We pair the whole specification to a so-called *global automaton* A_{DS}, that is, the FSA $A_{\varphi_{\mathrm{DS}}}$ of the constraint formula φ_{DS}. It thus recognizes all and only the model traces of DS. Recall that, as introduced in Definition 9, φ_{DS} is the conjunction of the formulae of the constraints in K, and thus the language $\mathscr{L}(A_{\mathrm{DS}})$ corresponds to $\bigcap_{\kappa \in K} \mathscr{L}(A_\kappa)$. By definition of automata product, this means that $\mathscr{L}(A_{\mathrm{DS}})$ can be obtained by computing the product of the local automata of the constraints in K.

Figure 5 shows four local automata for constraints taken from our running example: ALTERNATERESPONSE(r, v), CHAINRESPONSE($, p), PRECEDENCE(u, e) and ATMOSTONE(p). Examples of global automata are instead given in Fig. 6.

In the remainder of this chapter, we will extensively use local and global automata for reasoning, discovery, and monitoring. Though out of scope for this chapter, it is also worth mentioning that the automata-based approach has also been used for simulation of DECLARE models and thereby the production of event logs from declarative specifications [40], and also to define enactment engines for DECLARE specifications [83,98].

4.2 Reasoning on Declare Specifications

Reasoning on a DECLARE specification is necessary to understand which model traces are supported and, in turn, to ascertain its correctness. Reasoning is also key to unveil how constraints interact with each other, and check whether activations and targets are properly defined. As we will see, this is instrumental not only to analyze specifications, but it is also an integral part of declarative process mining.

In general, reasoning on declarative specifications is of particular importance: while they enjoy flexibility, they typically do not explicitly indicate how execution has to be controlled. We have seen how this phenomenon concretely manifests itself in the context of DECLARE: traces conforming to the specification (that is, *model traces*) are only implicitly described as those that satisfy all the given constraints. Constraints, in turn, may be quite diverse from each other (e.g., indicating what *is expected* to occur, but also what should *not* happen) and, even more importantly, may affect each other in subtle, difficult to detect ways. This phenomenon is known, in the literature that studies the cognitive impact of languages and notations, under the name of *hidden dependencies* [52]. Hidden dependencies in DECLARE have been studied in [76,35], and their impact on understandability and interpretability of declarative process models has spawned a dedicated line of research, started in [53].

We detail next key reasoning tasks in the context of DECLARE, substantiating how hidden dependencies enter into the picture. We show that all such reasoning tasks can be homogeneously tackled by a single check on the global automaton of the specification under study.
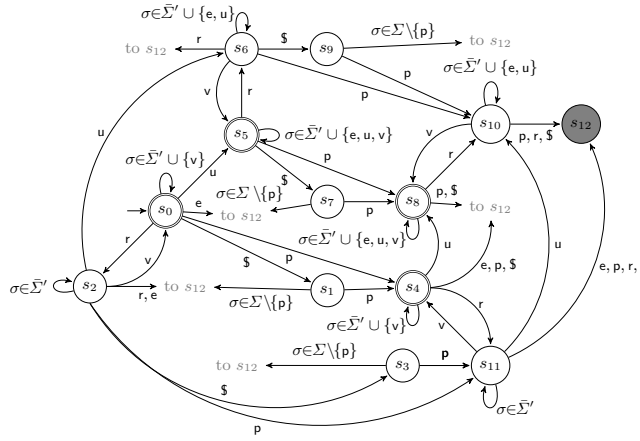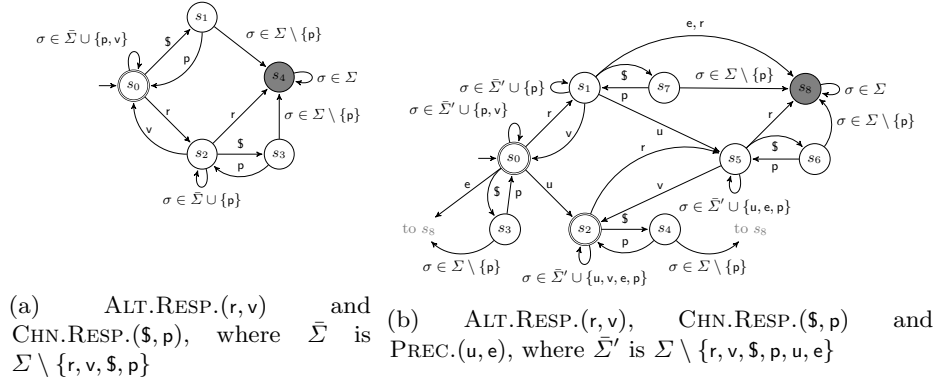
(a)   ALT.RESP.(r, v) and
CHN.RESP.$(\$, p)$,   where $\bar{\Sigma}$ is
$\Sigma \setminus \{r, v, \$, p\}$

(b)   ALT.RESP.(r, v),   CHN.RESP.$(\$, p)$ and
PREC.(u, e), where $\bar{\Sigma}'$ is $\Sigma \setminus \{r, v, \$, p, u, e\}$

(c) ALT.RESP.(r, v), CHN.RESP.$(\$, p)$, PREC.$(u, e)$ and ATMOSTONE(p), where $\bar{\Sigma}'$ is
$\Sigma \setminus \{r, v, \$, p, u, e\}$ (for the sake of readability, a few transitions to s_{12} are omitted)-

Fig. 6: Global automata for the interplay of DECLARE constraints.

Specification consistency. This is the most fundamental task, defined as follows.

Definition 11 (Consistent specification). *A* DECLARE *specification* DS *is* consistent *if there exists at least one model trace for* DS. ◁

Example 8. Consider the DECLARE specification in Fig. 7(a). The specification is inconsistent. This is not due to conflicting constraints insisting on the same activity, but due to hidden dependencies arising from the interplay of multiple constraints. To see why the specification is inconsistent, we can try to construct a trace that satisfies some of the constraints in the model, until we reach a contradiction (i.e., the "trace pattern" constructed so far violates a constraint of the specification). This is graphically shown next:
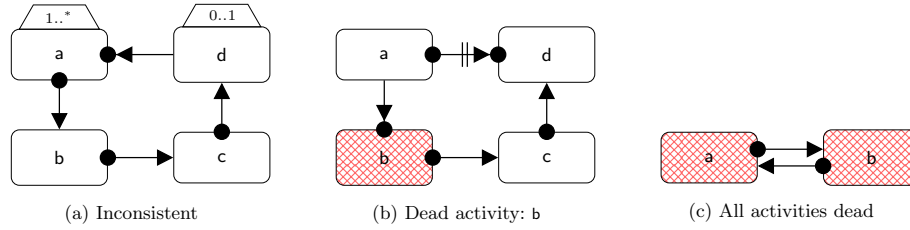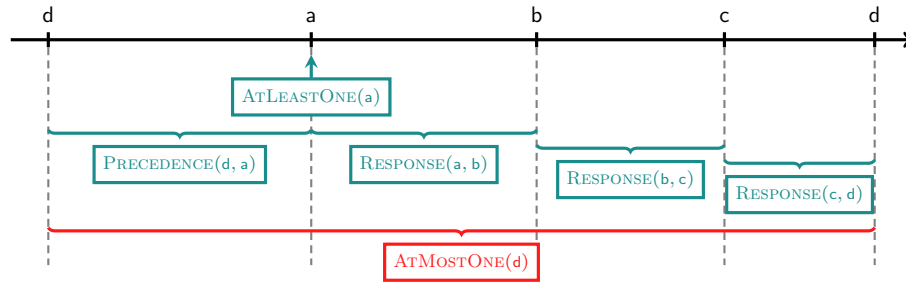
Fig. 7: Examples of incorrect DECLARE specifications.

The picture depicts that AtLeastOne(a) triggers:

- on the one hand Precedence(d, a), calling for a preceding occurrence of d;
- on the other hand, in cascade, Response(a, b), Response(b, c), and Response(c, d), calling for a later occurrence of d.

Considering the interplay of the involved constraints, d is required to occur in different instants, hence twice, in turn violating AtMostOne(d).   ◁

By definition of model trace, it is immediate to see that DS *is consistent if and only if the* LTL_f *specification formula* φ_{DS} *is satisfiable*. This, in turn, can be algorithmically verified by first constructing the global automaton A_{DS}, and then checking whether such an automaton is empty (i.e., it does not recognize any trace). Specifically, φ_{DS} *is* satisfiable *if and only if* A_{DS} *is non-empty*.

Detection of dead activities. This task amounts to check whether a DECLARE specification is over-constrained, in the sense that it contains an activity that can never be executed (in that case, such an activity is called *dead*).

Definition 12 (Dead activity). *Let* DS = (Rep, Act, K) *be a* DECLARE *specification. An activity* a ∈ Act *is* dead *in* DS *if there is no model trace of* DS *where* a *occurs.*   ◁

Example 9. Consider the DECLARE specification in Fig. 7(b). The specification is consistent; as an example, trace ⟨c, d⟩ is a model trace. However, none of its model traces can foresee the execution of b. This can be seen if one tries to construct a trace containing an occurrence of b. The result is the following:
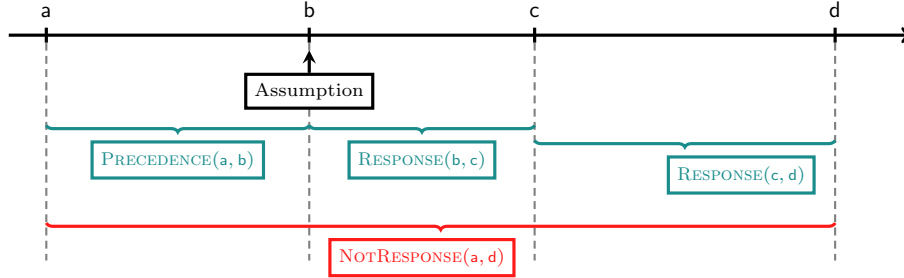
It is apparent that the presence of b requires a previous occurrence of a and, indirectly, a future occurrence of d, violating NOTRESPONSE(a, d). This shows that b is a dead activity.

Consider now the specification in Fig. 7(c). The situation here is trickier. The specification is consistent, as it accepts the empty trace (where no activity is executed, and hence none of the two response constraints present in the specification gets activated). However, none of the two activities a and b present therein can occur. As soon as this happens, the combination of the two response constraints cannot be *finitely* satisfied. In fact, an occurrence of a requires a later occurrence of b, which in turn requires a later occurrence of a, and so on and so forth, indefinitely. In other words, in every instant, one between RESPONSE(a, b) and RESPONSE(b, a) must be active and waiting for a later occurrence of its target, in a future instant. Since every instant must have a next instant, it is not possible to construct a satisfying (finite) trace.   ◁

Dead activity detection can be directly reduced to (in)consistency of a specification. Specifically, activity a is dead in a DECLARE specification DS = (REP, Act, K) if and only if the specification (REP, Act, K ∪ {ATLEASTONE(a)}), obtained from DS by forcing the existence of a is inconsistent (i.e., its specification formula is not satisfiable).

Valid activation and target. To ensure that a DECLARE constraint κ comes with a valid activation $_□κ$ and target $κ_▸$ for its formula $\varphi_κ$, we can directly apply Definition 10 and check whether the LTL$_f$ formula $\varphi_κ \leftrightarrow \Box(_□κ \to κ_▸)$ is valid, that is, whether its negation is not satisfiable.

Checking relations between constraints/specifications. We establish two key relations between constraints/specifications. The first is that of *subsumption* between templates, leveraging the entailment relation between LTL$_f$ formulae to constraints. We formally define it as follows.

Definition 13 (Subsumption). *Let* $κ(x_1, \ldots, x_m), κ'(x_1, \ldots, x_m) \in$ REP *two templates.* $κ(x_1, \ldots, x_m)$ subsumes $κ'(x_1, \ldots, x_m)$ *(in symbols,* $κ(x_1, \ldots, x_m) \sqsubseteq κ'(x_1, \ldots, x_m)$*) if, given any mapping* κ *assigning* $x_1, \ldots, x_m$ *with activities* $a_1, \ldots, a_m \in$ Act, $\varphi_{κ(a_1,\ldots,a_m)} \models \varphi_{κ'(a_1,\ldots,a_m)}$.   ◁

This relation can be checked by verifying that $\varphi_{κ(a_1,\ldots,a_m)} \to \varphi_{κ'(a_1,\ldots,a_m)}$ is valid, that is, the negated formula $\varphi_{κ(a_1,\ldots,a_m)} \land \neg\varphi_{κ'(a_1,\ldots,a_m)}$ is not satisfiable for any

$a_1, \ldots, a_m \in$ Act. For example, $\textsc{Alt.Prec.}(x, y) \sqsubseteq \textsc{Precedence}(x, y)$ as the former requires that y can occur only if preceded by x (just as the latter) *and* y does not recur in between. Therefore, every event that satisfies the former must satisfy the latter too. In the following, we shall lift this notion to constraints too (e.g., we say that $\textsc{AlternatePrecedence}(\mathsf{y}, \mathsf{p})$ subsumes $\textsc{Precedence}(\mathsf{y}, \mathsf{p})$).

By Definition 8 and Definition 9, since both Declare constraints and specifications correspond to LTL_f formulae, we can use subsumption for a twofold purpose:

- Consider two candidate constraints K_1 and K_2. If $\mathsf{K}_1 \sqsubseteq \mathsf{K}_2$, then we know that adding K_1 to a Declare specification will make the addition of K_2 irrelevant, and that adding K_1 or K_2 will determine whether the specification is more or less constraining.
- Consider a candidate constraint K and a target specification DS. If the former logically entails the latter, $\varphi_{\text{DS}} \models \varphi_{\mathsf{K}}$, then K is *redundant* in DS, and it makes no sense to include it in DS.

The second relation characterizes constraints that are the negated version of each other. Let K_1 and K_2 be two Declare constraints, coming with activation formulae $_\square\mathsf{K}_1$ and $_\square\mathsf{K}_2$ and target formulae $\mathsf{K}_{1\blacktriangleright}$ and $\mathsf{K}_{2\blacktriangleright}$, respectively. We say that K_1 and K_2 are the *negated versions* of one another if their activations are logically equivalent, that is $_\square\mathsf{K}_1 \leftrightarrow {}_\square\mathsf{K}_2$, and their targets are incompatible, that is, $\mathsf{K}_{1\blacktriangleright} \wedge \mathsf{K}_{2\blacktriangleright}$ is false. An example is that of Response vs NotResponse.

Consider now the situation where a decision must be taken concerning which of two candidate constraints K_1 and K_2 can be added to a Declare specification. Knowing that K_1 and K_2 are the negated versions of one another indicates that they should not *both* be added to the specification, as including them both would make the specification inconsistent as soon as the two constraints are activated.

As we will see in the next section, these notions become key when dealing with declarative process mining, and in particular the discovery of Declare specifications from event logs. Figure 8 graphically depicts how the main Declare constraint templates relate to each other in terms of subsumption and negated versions.

5   Declarative process mining

Declarative process constraints depict the interplay of every activity in the process with the rest of the activities. As a consequence, the behavioural relationships that hold among activities can be analysed with a local focus on each one [13], as a projection of the whole process behaviour on a single element thereof. The constraints pertaining to a single activity thus be seen as its footprint in the global behaviour of the process. We shall interchangeably interpret Declare constraints as *(i)* behavioural relations between activities in a process specification or *(ii)* rules exerted on the occurrence of events in traces. Notice that the latter is a different approach than the former, typically used for process modelling as originally conceived by the seminal work of Pesic et al. [84]. The former is instead

Cardinality templates	ATLEASTONE(x)		ATMOSTONE(x)

Position templates	INIT(x)	END(x)

(a) Existence templates

Relation templates	Negative relation templates	Relation templates	Negative relation templates

negates ~~~~

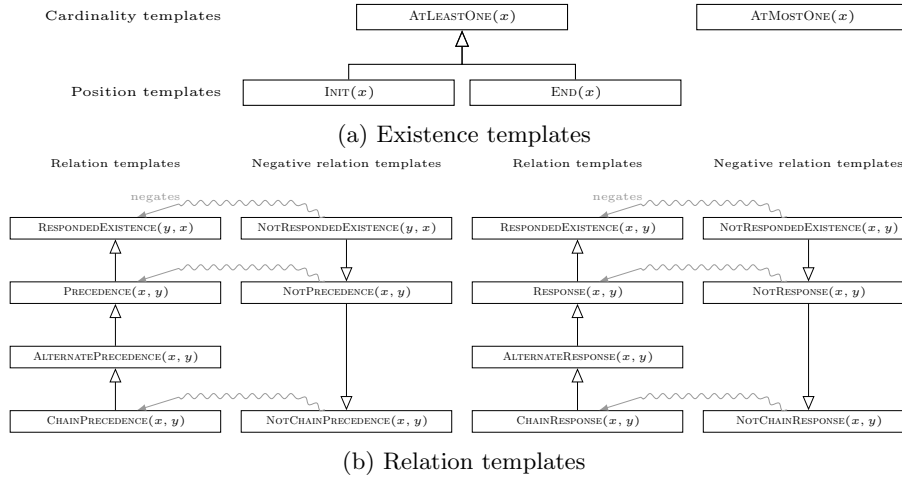| RESPONDEDEXISTENCE(y, x)	NOTRESPONDEDEXISTENCE(y, x)	RESPONDEDEXISTENCE(x, y)	NOTRESPONDEDEXISTENCE(x, y)
PRECEDENCE(x, y)	NOTPRECEDENCE(x, y)	RESPONSE(x, y)	NOTRESPONSE(x, y)
ALTERNATEPRECEDENCE(x, y)		ALTERNATERESPONSE(x, y)	
CHAINPRECEDENCE(x, y)	NOTCHAINPRECEDENCE(x, y)	CHAINRESPONSE(x, y)	NOTCHAINRESPONSE(x, y)

(b) Relation templates

Fig. 8: The subsumption map of DECLARE templates. Templates are indicated with solid boxes. The subsumption relation is depicted as a line starting from the subsumed template and ending in the subsuming one, with an empty triangular arrow recalling the UML IS-A graphical notation. The negative templates are graphically linked to the corresponding relation templates by means of wavy grey arcs.

the basis for declarative process mining. In the following, we describe how process specifications can be discovered and monitored.

5.1 Declarative process discovery

Declarative process discovery refers to the inference of those constraints that significantly rule the behaviour of a process, based upon an input event log. The problem can be framed in two distinct ways:

- A *discriminative discovery problem*, reminiscent of a classification task. This requires to split the input event log in two partitions, one containing "positive" examples and the second containing "negative" examples. Discovery amounts to find a suitable DECLARE specification that correctly reconstructs the classification, that is, accepts all positive examples and reject all negative ones.
- A *standard discovery problem* – also known as *specification mining* in the software engineering literature [58]. This calls for the individuation of which DECLARE constraints best describe the traces in the log, considering all of them as "positive" examples.

The first discovery algorithm for DECLARE treated discovery as a discriminative problem, exploiting inductive logic programming to tackle it [57,24]. In parallel, Goedertier et al. [51] brought forward techniques to generate negative examples

---

Algorithm 1: Overview of the discovery algorithm

Input: $L \in \mathcal{B}(\mathcal{U}_{act}^*)$, the event log to be analyzed;
  REP, a finite set of DECLARE templates to be considered to express the discovered specification;
  Act $\subseteq \mathcal{U}_{act}$, a finite set of activities to be included in the discovered specification;
  conf_t^{min}, supp_t^{min}, conf_e^{min}, supp_e^{min}, the minimum thresholds for trace-based confidence and support, and event-based confidence and support, respectively (default for all four paramenters: 0.0);
Output: DS, a declarative process specification

1 $K \leftarrow \Big\{ \text{K}(a_1, \ldots, a_m) : \text{K} \in \text{REP}, a_1; \ldots, a_m \in \text{Act}, a_i \neq a_j \text{ with } 1 \leq i, j \leq m \Big\}$
                                    /* candidate constraints: templates assigned with any pair of distinct activities */

2 **foreach** K $\in K$ /* compute measures */
3 **do**
4      $c_t \leftarrow \mathrm{conf}_t(\text{K}, L)$; $s_e \leftarrow \mathrm{supp}_t(\text{K}, L)$; $c_e \leftarrow \mathrm{conf}_e(\text{K}, L)$; $s_e \leftarrow \mathrm{supp}_e(\text{K}, L)$
5   **if** $c_t \leq \mathrm{conf}_t^{min}$ *or* $s_t \leq \mathrm{supp}_t^{min}$ *or* $c_e \leq \mathrm{conf}_e^{min}$ *or* $s_e \leq \mathrm{supp}_t^{min}$ **then**
6         $K \leftarrow K \setminus \{\text{K}\}$ /* remove constraints with a measure below the threshold */

7 **foreach** K $\in K$ /* remove constraints as per subsumption hierarchy and negated v. */
8 **do**
9   **foreach** K$'$ $\in K$ *s.t.* K$' \sqsubseteq$ K   /* for every K$'$ that subsumes K in K */
10   **do**
11   **if** $\mathrm{allm}\big(\text{K}', L\big) \leq \mathrm{allm}(\text{K}, L)$ /* if the measures of K$'$ are $\leq$ those of K */
12   **then**
13            $K \leftarrow K \setminus \{\text{K}\}$
14   **else** $K \leftarrow K \setminus \{\text{K}'\}$

15   **foreach** K$'$ $\in$ DS *s.t.* K$'$ *is the negated version of* K **do**
16   **if** $\mathrm{allm}\big(\text{K}', L\big) < \mathrm{allm}(\text{K}, L)$ **then** $K \leftarrow K \setminus \{\text{K}\}$
17   **else** $K \leftarrow K \setminus \{\text{K}'\}$

18 **return** DS $= (\text{REP}, \text{Act}, K)$

---

from positive ones. Interestingly, this line of investigation recently received again the attention of the community [95,23].

Declarative process discovery framed as a standard discovery problem finds its two main exponents in *Declare Miner* [64] and *MINERful* [43], which have been then extended with an arsenal of techniques to improve the quality and correctness of the discovered specifications. We follow the second thread, summarizing the main ideas exploited therein, though reshaping the core concepts in an attempt to embrace the wider plethora of declarative process discovery techniques and the advancements they brought [12,22,65].

Process discovery in a declarative setting typically consists of the following phases:
1) The initial setup, i.e., the selection of *(i)* the templates to be sought for, *(ii)* the activities to be considered for the candidate constraints instantiating those templates, and *(iii)* the minimum thresholds for constraint interestingness measures to retain a candidate constraint;
2) The computation of interestingness measures for all the constraints that instantiate the given templates;
3) The simplification of the returned specification, through *(i)* the removal of constraints whose measures do not reach the user-specified thresholds, *(ii)* the pruning of the redundant constraints from the set, and *(iii)* the removal of one constraint for every pair of constraints that are the negated version of one another.

Algorithm 1 gives a bird-eye view of the approach in pseudocode. As we can observe, interestingness measures are crucial to determine the degree to which constraints are satisfied in the log. They have been introduced to indicate

the level of reliability and relevance of constraints discovered from event logs, originally devised in the field of association rule mining [7] and adapted to the declarative process discovery context [71,21]. Among them, we recall support and confidence. Intuitively, support is a normalized measure quantifying how often the constraint is satisfied in the event log. Confidence considers the number of satisfactions with respect to the occurrences of the activations. We define them formally as follows.

Definition 14 (Trace-based measures). *Let L be a non-empty simplified event log with at least a non-empty trace, and κ a declarative constraint as per Definition 1. We define the trace-based support* supp_t *and the trace-based confidence* conf_t *as follows:*

$$\mathrm{supp}_\mathrm{t}(\kappa, L) = \frac{\sum\limits_{t \in L : t \models \Diamond(_\square\kappa) \wedge \kappa} L(t)}{\sum\limits_{t \in L} L(t)}; \tag{1}$$

$$\mathrm{conf}_\mathrm{t}(\kappa, L) = \frac{\sum\limits_{t \in L : t \models \Diamond(_\square\kappa) \wedge \kappa} L(t)}{\max\left\{1, \sum\limits_{t \in L : t \models \Diamond(_\square\kappa)} L(t)\right\}}. \tag{2}$$
◁

We remark that the condition at the numerator that the trace has to satisfy not only the constraint κ but also eventually its activation, i.e., $t \models \Diamond(_\square\kappa) \wedge \kappa$, serves the purpose of avoiding to count "vacuous satisfactions" discussed in Section 4.1. For example, while trace $\langle \mathsf{b}, \mathsf{c} \rangle$ satisfies CHAINRESPONSE(a, b), it does so vacuously, in the sense that it never activates the constraint. This intuitively means that CHAINRESPONSE(a, b), albeit satisfied, it cannot be interestingly used to describe the behaviour encoded in the trace. We recall that $L(t)$ denotes the multiplicity of occurrences of t in the log L (see [2]). The max term at the denominator of the formulation of confidence serves the purpose of avoiding a division by zero in case no trace satisfies $\Diamond(_\square\kappa)$.

Declare Miner first introduced the trace-based measures to discover specifications from logs, counting traces that (non-vacuously) satisfy constraints as a whole. MINERful, instead, advocated also the adoption of measures that lie at the level of granularity of events. The similarities and differences between the two measuring schemes and the role of explicit activations and targets to tackle vacuity has been later systematized in [22]. The motivation behind the use of event-based measures is the ability to give a differently weight to traces violating the constraints in more than one instant: with trace-based measures, e.g., both traces $\langle \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{c}, \mathsf{a}, \mathsf{b}, \mathsf{a}, \mathsf{b}, \mathsf{a}, \mathsf{b}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{a}, \mathsf{b}, \mathsf{a}, \mathsf{b}, \mathsf{a}, \mathsf{c} \rangle$ and $\langle \mathsf{b}, \mathsf{a}, \mathsf{c}, \mathsf{a}, \mathsf{c}, \mathsf{a}, \mathsf{a}, \mathsf{a}, \mathsf{a}, \mathsf{a}, \mathsf{a}, \mathsf{c} \rangle$ would count as single violations for CHAINRESPONSE(a, b). However, only the last occurrence of a out of ten leads to violation in the first trace, whereas all eight occurrences of a lead to violation in the second trace. Next, we formally capture the notion of event-based measures.

Definition 15 (Event-based measures). *Let L be a non-empty simplified event log with at least a non-empty trace, and κ a declarative constraint as*

per [Definition 1](). We define the event-based support $\mathrm{supp_e}$ *and the event-based confidence* $\mathrm{conf_e}$ *as follows:*

$$\mathrm{supp_e}(\textsc{k}, L) = \frac{\sum\limits_{t \in L} |\{a_i \in t : a, i \models ({}_\square\textsc{k} \wedge \textsc{k}_{\blacktriangleright})\}| \times L(t)}{\sum\limits_{t \in L} |t| \times L(t)}; \tag{3}$$

$$\mathrm{conf_e}(\textsc{k}, L) = \frac{\sum\limits_{t \in L} |\{a_i \in t : a, i \models ({}_\square\textsc{k} \wedge \textsc{k}_{\blacktriangleright})\}| \times L(t)}{\max\left\{1, \sum\limits_{t \in L} |\{a_i \in t : a, i \models {}_\square\textsc{k}\}| \times L(t)\right\}}. \tag{4}$$
$\triangleleft$

Again, the condition at the numerator that events satisfy both activation and target of the constraint is intended to avoid including vacuous satisfactions in the sum. The max term at the denominator of confidence is intended to avoid a division by zero in case no event satisfies ${}_\square\textsc{k}$.

For the sake of readability, we shall denote with $\mathrm{allm}(\textsc{k}, L)$ the tuple containing all computed measures for a constraint $\textsc{k}$ on the event log L: $\mathrm{allm}(\textsc{k}, L) = (\mathrm{supp_t}(\textsc{k}, L), \mathrm{conf_t}(\textsc{k}, L), \mathrm{supp_e}(\textsc{k}, L), \mathrm{conf_e}(\textsc{k}, L))$. Given two constraints $\textsc{k}_1$ and $\textsc{k}_2$, we write $\mathrm{allm}(\textsc{k}_1, L) \leq \mathrm{allm}(\textsc{k}_2, L)$ if $\mathrm{supp_t}(\textsc{k}_1, L) \leq \mathrm{supp_t}(\textsc{k}_2, L)$, $\mathrm{conf_t}(\textsc{k}_1, L) \leq \mathrm{conf_t}(\textsc{k}_2, L)$, $\mathrm{supp_e}(\textsc{k}_1, L) \leq \mathrm{conf_t}(\textsc{k}_2, L)$, and $\mathrm{conf_e}(\textsc{k}_1, L) \leq \mathrm{conf_t}(\textsc{k}_2, L)$. We write $\mathrm{allm}(\textsc{k}_1, L) \leq \mathrm{allm}(\textsc{k}_2, L)$ if $\mathrm{allm}(\textsc{k}_1, L) \leq \mathrm{allm}(\textsc{k}_2, L)$ and $\mathrm{allm}(\textsc{k}_2, L) \leq \mathrm{allm}(\textsc{k}_1, L)$.

Example 10 (An event log for the specification in [Example 1]()). Let $\mathcal{U}_{act} \doteq \{\mathsf{c}, \mathsf{r}, \mathsf{v}, \mathsf{t}, \mathsf{n}, \mathsf{y}, \$, \mathsf{p}, \mathsf{e}, \mathsf{u}\} \cup \{\mathsf{@}\}$ be an alphabet of activities. We interpret $\mathsf{@}$ as an email exchange, which can occur at any stage during the process. The other activities in $\mathcal{U}_{act}$ are those that were considered in the process specification in [Example 1](). Let the following event log be built on $\mathcal{U}_{act}$: $L = [t_1^{200}, t_2^{100}, t_3^{100}, t_4^{80}, t_5^{80}, t_6^4, t_7^2, t_8^2]$ where

$$t_1 = \langle \mathsf{c}, \mathsf{t}, \mathsf{r}, \mathsf{v}, \mathsf{y}, \$, \mathsf{p}, \mathsf{u}, \mathsf{e} \rangle \qquad t_2 = \langle \mathsf{c}, \mathsf{t}, \mathsf{t}, \mathsf{r}, \mathsf{v}, \mathsf{n}, \mathsf{t}, \mathsf{r}, \mathsf{v}, \mathsf{y}, \$, \mathsf{p}, \mathsf{u}, \mathsf{e} \rangle$$
$$t_3 = \langle \mathsf{c}, \mathsf{t}, \mathsf{r}, \mathsf{t}, \mathsf{v}, \mathsf{y}, \mathsf{u}, \$, \mathsf{p}, \mathsf{e} \rangle \qquad t_4 = \langle \mathsf{c}, \mathsf{t}, \mathsf{@}, \mathsf{t}, \mathsf{r}, \mathsf{v}, \mathsf{n}, \mathsf{@}, \mathsf{r}, \mathsf{v}, \mathsf{n} \rangle$$
$$t_5 = \langle \mathsf{c}, \mathsf{r}, \mathsf{t}, \mathsf{t}, \mathsf{v}, \mathsf{n}, \mathsf{y}, \mathsf{@} \rangle \qquad t_6 = \langle \mathsf{c}, \mathsf{t}, \mathsf{r}, \mathsf{t}, \mathsf{v}, \mathsf{@}, \mathsf{@}, \mathsf{y}, \$, \mathsf{p}, \mathsf{@}, \mathsf{e} \rangle$$
$$t_7 = \langle \mathsf{c}, \mathsf{@}, \mathsf{r}, \mathsf{v}, \mathsf{y}, \$, \mathsf{p}, \mathsf{@}, \mathsf{e} \rangle \qquad t_8 = \langle \mathsf{c}, \mathsf{t}, \mathsf{r}, \mathsf{r}, \mathsf{v}, \mathsf{@}, \mathsf{n} \rangle$$

We observe that the log above does not fully comply with the specification. Indeed, *(i)* trace t_8 violates $\textsc{AlternateResponse}(\mathsf{r}, \mathsf{v})$, as the candidate managed to register twice before evaluation (notice the occurrence of two consecutive r's before v); *(ii)* t_7 violates $\textsc{Precedence}(\mathsf{t}, \mathsf{v})$ and $\textsc{Precedence}(\mathsf{u}, \mathsf{e})$, as the candidate must have sent the admission test score and the necessary enrolment documents via email rather than via the system (see the occurrence of $\mathsf{@}$ in place of t in the second instant and in place of u later in the trace); finally, *(iii)* trace t_6 violates $\textsc{Precedence}(\mathsf{u}, \mathsf{e})$, as the candidate must have submitted the enrolment documents via email in that case too (notice the absence of task u and the presence of $\mathsf{@}$ in its stance). $\triangleleft$

Table 3: Measures computed for the relation constraints of Example 1 from the event log of Example 10.

Constraint	Event-based		Trace-based	
	Confidence	Support	Confidence	Support
Precedence(c, r)	1	0.129	1	1
AlternatePrecedence(r, v)	1	0.129	1	1
AlternateResponse(r, v)	0.997	0.129	0.996	0.996
Precedence(t, v)	0.997	0.129	0.996	0.996
AlternatePrecedence(v, n)	1	0.059	1	0.461
AlternatePrecedence(v, y)	1	0.084	1	0.856
NotResponse(y, n)	1	0.084	1	0.856
Precedence(y, p)	1	0.07	1	0.715
Precedence($, p)	1	0.07	1	0.715
ChainResponse($, p)	1	0.07	1	0.715
AtMostOne(p)	1	1	1	1
Precedence(p, e)	1	0.07	1	0.715
AtMostOne(e)	1	1	1	1
Precedence(u, e)	0.985	0.069	0.985	0.704

Example 11. With the example above, we have that both the trace-based support and trace-based confidence of Alt.Prec.(r, v), e.g., equate to 1.0: $\mathrm{supp}_t(\mathrm{Precedence}(c, r), L) = \mathrm{conf}_t(\mathrm{Precedence}(c, r), L) = 1.0$. This is because in all traces the activator (i.e., r) occurs and the constraint is not violated in any trace. Instead, $\mathrm{supp}_t(\mathrm{Alt.Prec.}(v, n), L) = \frac{100+80+80+2}{568} \approx 0.461$ and $\mathrm{conf}_t(\mathrm{Alt.Prec.}(v, n), L) = 1.0$. The trace-based support is lower than the trace-based confidence because the activator (n) occurs in 262 traces out of 568 (i.e., in the 100 instances of t_2, the 80 instances of t_4, the 80 instances of t_5, and the 2 instances of t_8). Similarly, $\mathrm{conf}_e(\mathrm{Precedence}(c, r), L) = 1.0$ and $\mathrm{conf}_e(\mathrm{Alt.Prec.}(v, n), L) = 1.0$. The measures do not change for event-based and trace-based confidence because every activation of the two constraints above leads to a satisfaction. In contrast, $\mathrm{supp}_e(\mathrm{Precedence}(c, r), L) = \frac{1 \times 200 + 2 \times 100 + 1 \times 100 + 2 \times 80 + 1 \times 80 + 1 \times 4 + 1 \times 2 + 2 \times 2}{9 \times 200 + 14 \times 100 + 10 \times 100 + 11 \times 80 + 8 \times 80 + 12 \times 4 + 9 \times 2 + 7 \times 2} = \frac{750}{5800} \approx 0.129$. ◁

It is worth noting that discovery approaches such as Declare Miner [64] and Janus [22] adopt (variations of) local constraint automata to count the satisfactions of constraints. MINERful [43] and DisCoveR [12] resort to occurrence statistics of activities gathered from the event log, more closely to the procedural discovery algorithms discussed in [2].

By definition of confidence and support (trace- or event-based), and as exemplified above, we observe that trace-based confidence is an upper bound
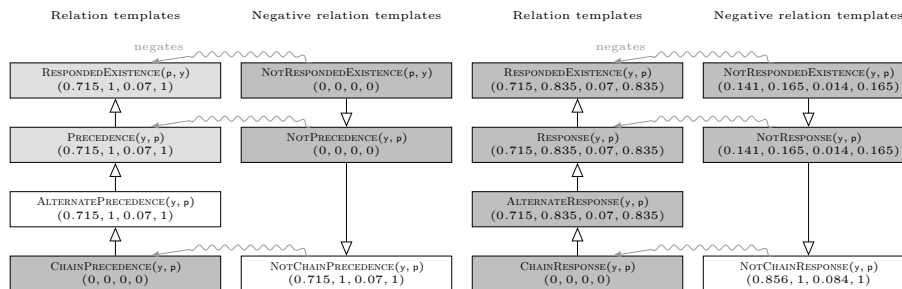
Fig. 9: The subsumption map of relation DECLARE constraints in a discovery context. The graphical notation follows Fig. 8. Gray boxes denote constraints that have measures below the minimum thresholds. Light-gray boxes indicate constraints that are subsumed by others with equivalent measures.

for trace-based support and event-based confidence is an upper bound for event-based support. Next, we illustrate how the discovery algorithm operates with our running example.

Example 12. Table 3 shows the event-based and trace-based measures computed on the basis of our running example for every constraint in the original specification – phase (2) of the discovery procedure described above. They belong to the output of the discovery algorithm running on the event log of Example 10 set at phase (1) to seek for *(i)* all templates from the DECLARE repertoire in Table 2 *(ii)* over activities $\{c, r, v, t, n, y, \$, p, e, u\}$, with *(iii)* minimum event-based confidence of 0.95. We remark that also ALTERNATEPRECEDENCE(y, p), CHAINPRECEDENCE$(\$, p)$, ALTERNATEPRECEDENCE$(p, e)$ and ALTERNATEPRECEDENCE(c, p), NOTCHAINPRECEDENCE(y, p) and NOTCHAINRESPONSE(y, p), among others, fulfil those criteria and thus are part of the returned set. ◁

To increase the information brought by a discovered model, not only we prune the constraints whose measures lie below the given threshold values. Also, we take into account the subsumption hierarchy illustrated in Fig. 8. In addition, we retain in the constraint set only one among pairs that are a negated version of one another. If we kept both, the model would turn the activation in common into a dead activity (see Section 4.2).

Example 13. Figure 9 illustrates the result of the pruning phase (3) based on subsumption and choice of constraints that are the negated version of one another, based on the event log of Example 10. We observe that ALTERNATEPRECEDENCE(y, p) has the same measures as PRECEDENCE(y, p), and we know that PRECEDENCE(y, p) is subsumed by ALTERNATEPRECEDENCE(y, p) (see Section 4.2); as we are interested in more restrictive constraints that reduce the space of possible process runs to more closely define its behaviour, we retain the former and discard the latter. Keeping both would introduce a redundancy,

(a) ALT.RESP.(r, v)   (b) CHN.RESP.($, p)   (c) PREC.(u, e)   (d) ATMOSTONE(p)
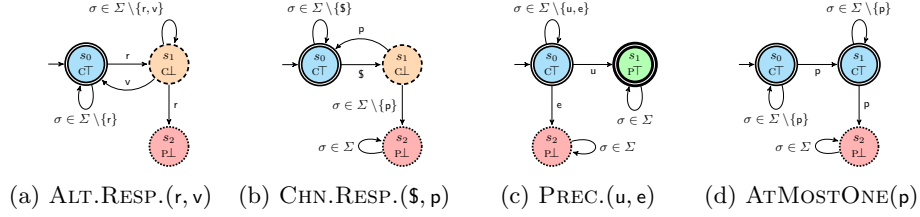
Fig. 10: Example FSAs adapted for the monitoring of constraints. Non-final states indicating current violation (C⊥) are dashed and filled in orange; non-final states indicating permanent violation (P⊥) are dotted and filled in red; final states indicating current satisfaction (C⊤) are thin-solid and filled in blue; final states indicating permanent satisfaction (P⊤) are thick-solid and filled in green.

and retaining only the latter would omit detailed information as not only p must be preceded by y, but also p cannot recur unless y occurs again. By the same line of reasoning, we prefer retaining INIT(c) to ATMOSTONE(c) in the result specification. The same concepts apply with CHAINPRECEDENCE($, p), to be preferred over PRECEDENCE($, p) and ALTERNATEPRECEDENCE(p, e) in place of PRECEDENCE(p, e), among others. Notice that PRECEDENCE(y, p), PRECEDENCE($, p) and PRECEDENCE(p, e) were in the given specification of our running example but, we conclude, are not the most restrictive constraints that could be used in the specification, as the discovery algorithm evidences.   ◁

To conclude, we remark that not all redundancies can be found with the sole subsumption-hierarchy based pruning. The subsumption hierarchy, indeed, checks constraints that are exerted on the same activities – e.g., ALTERNATEPRECEDENCE(y, p) and PRECEDENCE(y, p). Therefore, we need a more powerful redundancy checking mechanism, seeking for constraints that are entailed by the remainder of the specification's constraint set (see Section 4.2).

Example 14. The confidence of ALTERNATEPRECEDENCE(v, p) is 1.0 in the event log of our running example. Yet, it does not add information to the discovered specification as it is redundant, logically entailed by the other constraints – in particular, ALTERNATEPRECEDENCE(r, v), ALTERNATEPRECEDENCE(v, y), PRECEDENCE(y, p) and ATMOSTONE(p).   ◁

   To verify this, we can resort to language inclusion via automata product as in [41]: the language of the product of the four constraint automata is not smaller than the language accepted by the intersection of the second, third and fourth constraint automata. Here, we do not enter the details of the algorithms that detect redundancies at such a deeper level but provide an example of its rationale. The interested reader can find further details in [41,28].

5.2   Declarative process monitoring

(Compliance) process monitoring aims at tracking running process executions to check their conformance to a reference process model, with the purpose of detecting and reporting deviations as soon as possible [63]. It constitutes one of the main tasks of operational decision support [1, Ch. 10], which characterizes process mining applied at runtime to running process executions.

Declarative process monitoring employs a declarative specification (in our case, described using DECLARE) as reference model for monitoring. The central fact in monitoring that process instances are running, that is, their generated traces evolve over time, calls for a finer-grained understanding of the state of constraints and of the whole specification. We illustrate this intuitively in the next example.

Example 15. Consider the excerpt in Fig. 11 of our admission process running example, and an evolving trace that, once completed, corresponds to the following sequence: $\langle \$, p, u, \$, p \rangle$. Let us replay the trace from the beginning.

1. At the beginning, all constraints are *satisfied*, but they are so for sure only *currently*, as events may occur making them violated. For example, a registration without a consequent evaluation would lead to violating ALTERNATERESPONSE(r, v), whereas an enrolment without a prior upload of certificates would lead to a violation of PRECEDENCE(u, e).
2. Upon the occurrence of $, constraint CHAINRESPONSE($, p) becomes pending or, to be more precise, *currently violated*, as paying demands a pre-enrolment occurring immediately after.
3. The execution of p brings CHAINRESPONSE($, p) back to *currently satisfied*, as it does not require the occurrence of further events, but may do so in the future in case of another payment.
4. Upon the occurrence of u, constraint PRECEDENCE(u, e) becomes *permanently satisfied*, as enrolment is now enabled, and there is no way to continue the execution leading to a violation of the constraint.
5. This is indeed what happens with the next occurrence of $, which makes CHAINRESPONSE($, p) *currently violated*.
6. The second pre-enrolment has the effect of bringing CHAINRESPONSE($, p) once again back to *currently satisfied*. However, it has also the effect of *permanently violating* ATMOSTONE(p), as the number of occurrences of p has exceeded the upper bound allowed by ATMOSTONE(p), and there is no way of fixing the violation.   ◁

As witnessed by the example, the state of each constraint can be described in a fine-grained way by considering on the one hand the trace accumulated so far (i.e., the prefix of the whole, still unknown, execution), and by pondering on the other hand about the possible, future continuations. To do so in a formal way, we appeal to the literature on runtime-verification for linear temporal logics, and in particular to the RV-LTL semantics, originally introduced in [15] over infinite traces. This semantics was adopted for the first time in the context of LTL_f over finite traces in [70,72], in order to define an operational technique for
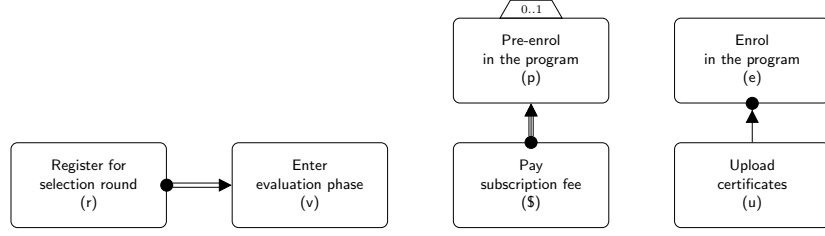
Fig. 11: Excerpt of the DECLARE specification in Fig. 2.

DECLARE monitoring. This led to deeper investigations on the usage of RV-LTLto characterize the relevance of a trace to a declarative specification [42], and to finally obtain a formally grounded, comprehensive framework for monitoring [31,32].

We now define the RV-LTL semantics for LTL_f. In the definition, we denote the concatenation of trace t_1 with t_2 as $t_1 \cdot t_2$.

Definition 16 (RV-LTL states). *Consider an* LTL_f *formula* φ *over* Σ, *and a trace* t *over* Σ^*. *We say that* φ *is in (RV-LTL) state* s *after* t, *written* $[t \models \varphi]_{\mathrm{RV}} = v$, *if:*

(Permanent satisfaction) *(i)* $v = \mathrm{P}\top$, *(ii) the current trace satisfies* φ *($t \models \varphi$), and (iii) every possible suffix keeps* φ *satisfied (for every trace* $t' \in \Sigma^*$, *we have* $t \cdot t' \models \varphi$).

(Permanent violation) *(i)* $v = \mathrm{P}\bot$, *(ii) the current trace violates* φ *($t \not\models \varphi$), and (iii) every possible suffix keeps* φ *violated (for every trace* $t' \in \Sigma^*$, *we have* $t \cdot t' \not\models \varphi$).

(Current satisfaction) *(i)* $v = \mathrm{C}\top$, *(ii) the current trace satisfies* φ *($t \models \varphi$), and (iii) there exists a suffix that leads to violate* φ *(for some trace* $t' \in \Sigma^*$, *we have* $t \cdot t' \not\models \varphi$).

(Current violation) *(i)* $v = \mathrm{C}\bot$, *(ii) the current trace violates* φ *($t \not\models \varphi$), and (iii) there exists a suffix that leads to satisfy* φ *(for some trace* $t' \in \Sigma^*$, *we have* $t \cdot t' \models \varphi$).

We also say that t conforms to φ *if* $[t \models \varphi]_{\mathrm{RV}} = \mathrm{P}\top$ *or* $[t \models \varphi]_{\mathrm{RV}} = \mathrm{C}\top$ *(i.e., stopping the execution in* t *satisfies the formula).*   ◁

By inspecting the definition, we can directly see that monitoring is at least as hard as LTL_f satisfiability/validity checking. To see this, consider what happens at the *beginning of an execution*, where the current trace is empty. By applying Definition 16 to this special case, and by recalling the notion of satisfiability/validity of an LTL_f formula, we in fact get that an LTL_f formula φ is:

- permanently satisfied if φ is valid;
- permanently violated if φ is unsatisfiable;
- currently satisfied if the two formulae $\varphi \wedge \mathbf{end}$ and $\neg\varphi$ are both satisfiable;
- currently violated if the two formulae $\neg\varphi \wedge \mathbf{end}$ and φ are both satisfiable.

To perform monitoring according to the RV-LTL states from Definition 16, we can once again exploit the automata-theoretic characterization of LTL_f. In particular, given an LTL_f formula φ, we construct its FSA A_φ, and *color* the automaton states according to the RV-LTL semantics. As introduced in [70] and then formally verified in [32], this can be simply done as follows. Consider a state s in of A_φ. We label it by:

- $\text{P}\top$, if s is final and all the states reachable from s in A_φ are final as well; if A_φ is minimized, this means that s only reaches itself.
- $\text{P}\bot$, if s is non-final and all the states reachable from s in A_φ are non-final as well; if A_φ is minimized, this means that s only reaches itself.
- $\text{C}\top$, if s is final and can reach a non-final state in A_φ.
- $\text{C}\bot$, if s is non-final and can reach a final state in A_φ.

Figure 10 shows some examples of colored constraint automata, obtained by considering the constraint formulae of some DECLARE constraints from our running example. To monitor the state evolution of a constraint, one has simply to dynamically play the evolving trace on its colored local automaton, returning the updated RV-LTL label as soon as a new event is processed. Doing so on the local automata in Fig. 10 for trace $\langle \$, \mathsf{p}, \mathsf{u}, \$, \mathsf{p} \rangle$ formally reconstructs what discussed in Example 15.

However, this is not enough to *promptly detect violations* as soon as they manifest in the traces. This has been extensively discussed in [72,32], and is at the very core of the power of temporal logic-based techniques for monitoring. We use again Example 15 to illustrate the problem.

Example 16. Consider Example 15 and the following question: is step 6 the earliest at which a violation can be detected? Clearly, if we focus on each constraint in isolation, the answer is affirmative. To see this formally, we play trace $\langle \$, \mathsf{p}, \mathsf{u}, \$, \mathsf{p} \rangle$ on the four colored local automata of Fig. 10, obtaining the following runs:

- For ALTERNATERESPONSE(r, v), we have $s_0 \xrightarrow{\$} s_0 \xrightarrow{\mathsf{p}} s_0 \xrightarrow{\mathsf{u}} s_0 \xrightarrow{\$} s_0 \xrightarrow{\mathsf{p}} s_0$; no violation is encountered.
- For CHAINRESPONSE$(\$, \mathsf{p})$, we have $s_0 \xrightarrow{\$} s_1 \xrightarrow{\mathsf{p}} s_0 \xrightarrow{\mathsf{u}} s_0 \xrightarrow{\$} s_1 \xrightarrow{\mathsf{p}} s_0$; no violation is encountered.
- For PRECEDENCE(u, e), we have $s_0 \xrightarrow{\$} s_0 \xrightarrow{\mathsf{p}} s_0 \xrightarrow{\mathsf{u}} s_1 \xrightarrow{\$} s_1 \xrightarrow{\mathsf{p}} s_1$; no violation is encountered.
- For ATMOSTONE(p), we have $s_0 \xrightarrow{\$} s_0 \xrightarrow{\mathsf{p}} s_1 \xrightarrow{\mathsf{u}} s_1 \xrightarrow{\$} s_1 \xrightarrow{\mathsf{p}} s_2$; a violation is encountered in the last reached state.

The answer changes if we consider the whole DECLARE specification that contains all such constraints at once. In fact, by taking into account the interplay of constraints, we can detect a violation already at step 5, i.e., after the second occurrence of payment. This is because, after that step, the two constraints CHAINRESPONSE$(\$, \mathsf{p})$ and ATMOSTONE(p) enter into a *conflict*, that is, no continuation of the current trace can lead to satisfy them both. In fact, after trace $\langle \$, \mathsf{p}, \mathsf{u}, \$ \rangle$, constraint CHAINRESPONSE$(\$, \mathsf{p})$ is currently violated, waiting for a consequent occurrence of p; however, constraint ATMOSTONE(p), which is
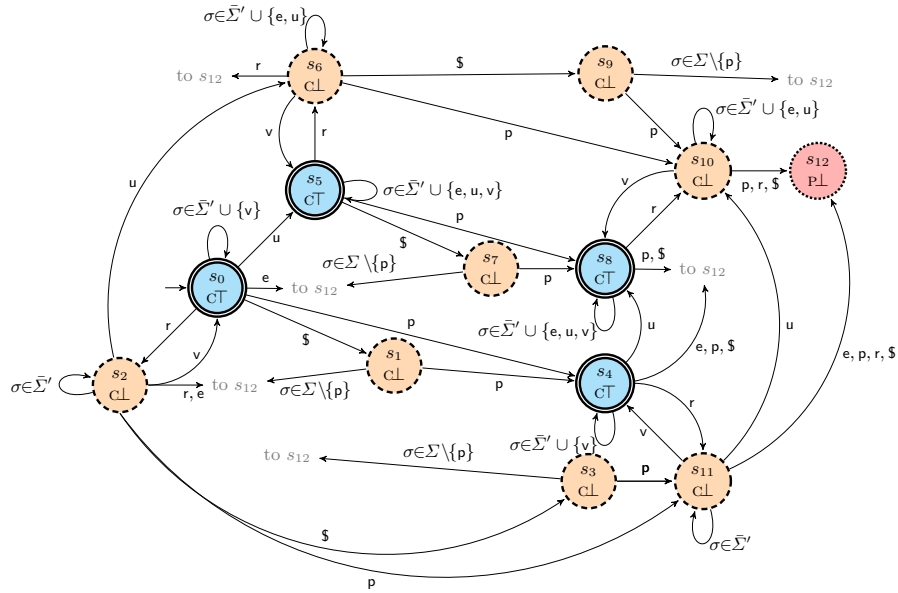
Fig. 12: The colored global automaton automaton obtained as the (colored) cross-product of constraints in Fig. 10 as shown in Fig. 6(c), the states of which are decorated with the four RV-LTL truth values.

currently satisfied, becomes permanently violated upon a further occurrence of p.   ◁

As we have seen, the early detection of violations cannot always be caught by considering the colored local automata of constraints in isolation. However, it can be systematically detected by taking into account the colored global automaton of the whole specification.

Example 17. Figure 12 shows the colored global automaton of the DECLARE specification in Fig. 11. By playing the trace $\langle \$, p, u, \$, p \rangle$ therein, we obtain the following run: $s_0 \xrightarrow{\$} s_1 \xrightarrow{p} s_4 \xrightarrow{u} s_8 \xrightarrow{\$} s_{12} \xrightarrow{p} s_{12}$. Clearly, the violation state s_{12} is already reached in step 5, i.e., just after the second payment.   ◁

All in all, we can then monitor an evolving trace against a DECLARE specification as follows:
- Each constraint is encoded into the corresponding colored local automaton, used to track the state evolution of the constraint itself.
- The whole specification is encoded into the corresponding colored global automaon, used to track the evolution of the whole specification, and in particular to early-detect violations.
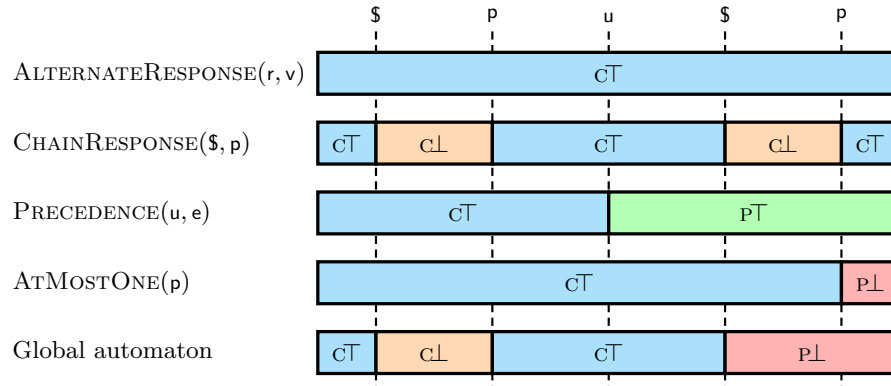
Fig. 13: Monitoring with local and global colored automata, showing a case where the global automaton detects a violation before it actually manifests on a single constraint.

- At runtime, every new event occurrence is delivered in parallel to all the automata, updating each of them by executing the corresponding transition and entering into the next state, at the same time returning the associated RV-LTL label.

Figure 13 shows the result of applying this technique to our running example.

An alternative approach, which is exploited in [70], is to compute, as done before, the global automaton as the cross-product of local automata, remembering, in each global state, the RV-LTL labels of all local states from which such a global state has been produced. In addition, no minimization step is applied on the resulting automaton. Once colored, this non-minimized, global colored automaton combines in a single device the contribution of all local monitors and that of the global monitor.

5.3   A Note on Conformance Checking

In this section, we have focused on monitoring evolving traces against Declare specifications. This can be seen as a form of *online conformance checking*, aiming at detecting deviations at execution time. This technique can be seamlessly lifted to handle the standard conformance checking task, where conformance is evaluated on an event log containing full traces of already completed process executions (cf. [20]). In this setting, the global automaton is not needed anymore, as a-posteriori it is not relevant to compute the earliest moment of a violation, but only to properly detect it at the trace level. The usage of local automata, one per constraint, is enough, and also has the advantage of producing an informative feedback that indicates, trace by trace, how many (and which) constraints are satisfied or violated. Finer-grained feedbacks like those based on the computation of trace alignments have been extensively applied for procedural models (cf. [20]), and can be also recast in the declarative setting, aligning the log traces with

the (closest) model traces accepted by the global automaton of the Declare specification of interest. This is an active line of research, which started from the seminal approach in [60].

6 Recent Advances and Outlook

We close this chapter by reporting about the most recent advances in the field of declarative process mining revolving around Declare, describing the current frontier of research, and highlighting open challenges.

6.1 Beyond Declare Patterns

As we have seen in Section 3, a Declare specification consists of a repertoire of constraint templates grounded on specific activities. At the same time, such templates come with a logic-based semantics given in terms of LTL_f. A natural question is then: can the techniques described in this chapter be used for the *entire* LTL_f logic? This means, more precisely, considering the situation where each constraint corresponds to an arbitrary LTL_f formula while, as usual, the specification formula is constructed by putting in conjunction the LTL_f formulae of all its constituting constraints.

To answer this question, one has to separate the *logical* and *pragmatic* aspects involved in the different tasks we have been introducing. We do so focusing on reasoning, discovery, and monitoring.

Reasoning. As discussed in Section 4.2, all the reasoning tasks we have considered in this chapter can be lifted to the whole LTL_f logic. Indeed, they are reduced to LTL_f satisfiability/validity checking, which in turn can be tackled by checking (non-)emptiness of FSAs. The situation may change if one wants to provide more advanced debugging or diagnosis functionalities – for example, to return the most relevant conflicting set(s) of constraints that are causing inconsistencies or dead activities. While these types of problem can also be attacked at the level of the entire logic [85,29], focusing only on pre-defined patterns becomes necessary if one wants to involve humans in the loop or define preferences over constraints in the case where multiple explanations exist [29]. Considering specific patterns is also relevant when studying the computational complexity of reasoning on pattern combinations [97,49,50], or the scalability and effectiveness of reasoning tools [77,98,49,50].

Discovery. As pointed out in Section 5.1, two distinct process discovery problems are typically tackled in a declarative setting: discriminative discovery and specification mining.

The case of discriminative discovery is tightly related to classification and machine learning, allowing one to rely on general learning algorithms for declarative process mining. Such algorithms tackle general logical frameworks, such as Horn clauses in inductive logic programming or full temporal logics in model learning,

and can thus go far beyond a pre-defined set of templates, either targeting full LTL_f [19,88] or enriching the discoverable Declare templates with further key dimensions, such as metric temporal constraints, event attributes, and data conditions [25,27].

As shown in Section 5.1, standard discovery stands as a radically different problem, since the input event log provides a uniform set of (positive) examples, while no negative example is given. This calls for suitable metrics to measure *how well* a set of constraints characterizes the behaviour contained in the log. In the approach described in this chapter, such metrics are defined starting from the notions of constraint activation and target, which are template-specific. Attempts have been conducted to lift some of these notions (in particular that of activation and "relevant" satisfaction [42]) to full LTL_f, but further research is needed to target the discovery of arbitrary LTL_f formulae from event logs. Notice that while full LTL_f discovery would enrich the expressiveness of the discovered specifications, it would on the other hand pose the issue of *understandability*: end users may struggle when confronted with arbitrary temporal formulae, while they are facilitated when pre-defined templates are used.

Monitoring. As we have discussed in Section 5.2, Declare monitoring is tackled using automata, and consequently seamlessly work for arbitrary LTL_f formulae. As for advanced debugging techniques, the same considerations done for reasoning also hold for monitoring. For example, the detection of minimal conflicting sets of constraints in the case of early detection of violations caused by the interplay of multiple constraints can be tamed at the level of the full logic [72], but would require to focus on patterns if one wants to formulate preferences or incorporate human feedback [29].

Remarkably, working with FSAs allows us to define monitors for temporal formulae that go even beyond LTL_f. In fact, LTL_f is as expressive as star-free regular expressions, while automata are able to capture full regular expressions and, in turn, finite-trace temporal logics incorporating in a single formalism LTL_f and regular expressions, such as Linear Dynamic Logic over finite traces (LDL_f) [34]. Working with LDL_f in our setting has the specific advantage that we can express and monitor *metaconstraints*, that is, constraints that predicate on the RV-LTL truth values of other constraints [31,32].

6.2 Dealing with Uncertainty

In the conventional definition of a Declare specification, constraints are interpreted as being *certain*: every model trace is expected to satisfy all constraints contained in the specification. Such an interpretation is too restrictive in scenarios where the specification should accommodate:

- constraints describing common behaviours, expected to hold in the majority, but not all cases;
- constraints describing exceptional, outlier behaviours that rarely occurs but should be not judged as violating the specification.

To deal with this form of *uncertainty*, DECLARE has been recently extended with *probabilistic constraints* [69]. In this framework, every probabilistic constraint comes with:

- a constraint formula φ (specified, as in the standard case, using LTL_f);
- a comparison operator $\odot \in \{=, \neq, <, \leq, >, \geq\}$;
- a number $p \in [0, 1]$.

The interpretation of this constraint is that φ holds in a random trace generated by the process with a probability that is $\odot p$. In frequentist terms, this can be in turn interpreted as follows: given a log of the process, the ratio of traces satisfying φ must be $\odot p$.

Since a DECLARE specification contains multiple constraints, one has to consider how different probabilistic constraints interact with each other. In particular, n probabilistic constraints yield up to 2^n possible so-called *scenarios*, each highlighting which probabilistic constraints hold and which are violated. Reasoning over such scenarios has to be conducted by suitably mixing their temporal and probabilistic dimensions. The former handles which combinations of constraints and their violations (i.e., which scenarios) are consistent, while the latter lifts the probability conditions attached of single constraints to discrete probability distributions over the possible scenarios.

To carry out this form of combined reasoning, probabilistic constraints are formalized in a well-behaved fragment of the logic introduced in [68]. As it turns out, logical and probabilistic reasoning are loosely coupled in this fragment, and can be carried out resorting to standard finite-state automata and systems of linear inequalities. This approach has been used as the basis for defining a new family of *probabilistic declarative process mining* techniques [10].

6.3   Mixed-Paradigm Models

In Fig. 1, we have intuitively contrasted declarative specifications and imperative models. The distinction of these two approaches is in reality not so crisp. In fact, a single process may contain parts that are more suitably captured using imperative languages, and parts that can be better described as declarative specifications. Take, for instance, a clinical guideline mixing administrative and therapeutic subprocesses [79].

To capture such *hybrid* processes, one needs a multi-paradigm approach that can combine imperative and declarative constructs in a single process model. One of the first proposals doing so is [91], where an imperative process can contain activities that are internally structured using so-called *pockets of flexibility* specified using declarative temporal constraints over a given set of tasks.

This layered approach has been further developed in [96], which brings forward a hierarchical model where each sub-process can be specified either as an imperative or declarative component. Discovery of hierarchical hybrid process models has been subsequently tackled in [93].

Multi-paradigm approaches providing a tighter integration between imperative and declarative components have also been studied. In [36], process models

combining Petri nets and DECLARE constraints at the same modelling level are introduced and studied, singling out methodologies and techniques to handle the intertwined state space emerging from their interaction. Conformance checking for these mixed-paradigm models is extensively assessed in [44]. A different approach is brought forward in [9], where a DECLARE specification is used to express global constraints that "glue together" multiple imperative processes concurrently executed over the same instances. Automata-based techniques extending those illustrated in Section 5.2 are introduced to provide integrated monitoring functionalities dealing at once with the local processes and the global constraints.

At the current stage, further research is needed along the illustrated lines towards a solid theory and corresponding algorithmic techniques for *hybrid, mixed-paradigm process mining*.

6.4   Multi-Perspective Declare Specifications

Throughout the chapter, we have considered pure control-flow specifications, where a process is captured solely in terms of its constitutive activities and of behavioural constraints separating legal from undesired executions. While the control-flow provides the main process backbone, other equally important perspectives should also be taken into account as suggested already in [3]:

- The *resource* perspective deals with the actors that are responsible for executing tasks within the process.
- The *time* perspective focusses on quantitative temporal conditions on when tasks can/must be scheduled and executed, and on their expected durations.
- The *data* perspective captures how data objects and their attributes influence and are manipulated during the process execution.

Several works have investigated the extension of DECLARE with additional perspectives. From the formal point of view, this requires to extend the logic-based formalization of DECLARE with features that can capture resources, metric time, data, and conditions thereof, in turn resorting to variants of metric and/or first-order formalisms over finite traces [80,75,14,18]. It is important to stress that such features may be blurred, considering that data support (if equipped with suitable datatypes and conditions) may be used to predicate over resources and time as well.

Such multi-perspective features have been extensively embedded into DE-CLARE or related approaches (see, for example, [99,75,17] for constraints with metric time and [47] for constraints with metric time and resources). Next, we focus in more detail on the data dimension.

When it comes to data, two main lines of research can be identified. The first one deals with standard "case-centric" processes extended with event and case data. The second one focuses instead on "multi-case" processes, wherein constraints are expressed over multiple objects and their mutual relations. We briefly discuss each line separately.

Declarative process specifications with event/case data. Within a process, activities may be equipped with data attributes that, at execution time, are grounded to actual data values by the involved resources. This means that events witnessing the occurrence of task instances come with a data payload. In addition, each process instance may evolve its own case data in response to the execution of activities.[5] Such case data may be stored in different ways, e.g., as key-value pairs or a full-fledged relational database. In this setting, it becomes crucial to extend DECLARE with so-called *data-aware constraints*, that is, constraints enriched with data-aware conditions over activities. The simple but illustrative example described next motivates why this is needed.

Example 18. We focus on a process where payments are issued by customers through a pay activity, which comes with an attribute indicating the paid amount, in Euros. Two consequent activities check and emit are executed to respectively inspect a payment and emit a receipt.

Let a log for this process contain multiple repetitions of the following traces:

$$t_1 = \langle \mathsf{pay(amount=50), emit} \rangle \qquad t_2 = \langle \mathsf{pay(amount=300), check, emit} \rangle$$

$$t_3 = \langle \mathsf{pay(amount=20)} \rangle \qquad t_4 = \langle \mathsf{pay(amount=100), emit, check} \rangle$$

$$t_5 = \langle \mathsf{pay(amount=90), emit} \rangle \qquad t_6 = \langle \mathsf{pay(amount=800), check} \rangle$$

One may wonder whether RESPONSE(pay, check) is a suitable constraint to explain (part of) the behaviour contained in the log. If considered unrestrictedly, this is not the case, as there are many traces where payment is not followed by any inspection. The situation changes completely if one restricts the scope of the constraint activation only to those payments that involve an amount of 100 or more. ◁

A number of works has brought forward combined techniques to discover DECLARE constraints equipped with various forms of data conditions [66,92,59], to check conformance for data-aware constraints [17,16], and to handle their monitoring [75,9]. This passage has to be carried out with extreme care, as combining event data and time quickly leads to undecidability of reasoning [38,37,18]. Therefore, such techniques have to operate in a limited fashion or suitably controlling the expressiveness of data conditions and the way they interact with time.

Object-centric declarative process specifications. So far, we have discussed the extension of DECLARE with event or case data. In a more general setting, data may refer to more complex networks of objects and their mutual relations, simultaneously co-evolved by one or multiple processes. In this type of processes, known under the umbrella term of *object-centric processes*, there is no single, pre-defined notion of case, and process executions cannot consequently be represented

---

[5]For conciseness of presentation, we will not distinguish between event and case data in our discussion, but technically they pose different, albeit tightly related, requirements.

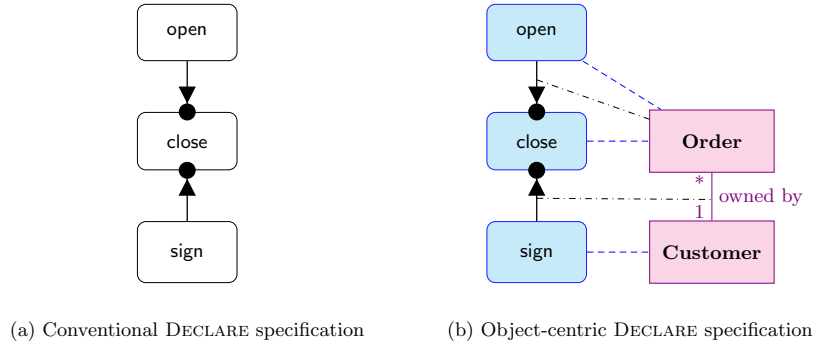(a) Conventional DECLARE specification   (b) Object-centric DECLARE specification

Fig. 14: Comparison of conventional vs object-centric DECLARE.

as flat traces, but call for richer representations (cf. [48]). The following example illustrates why DECLARE, in its conventional version, cannot be used to capture *object-centric* processes.

Example 19. Consider the fragment of an order-to-cash process, containing three activities: sign (indicating the signature of a GDPR form by the customer), open (the opening of an order), and close (the closing of an order). Two constraints apply to close, defining under which conditions it becomes executable:
• An order can be closed only if *that order* has been opened before.
• An order can be closed only if *its owner* has signed the consent before.

   Figure 14(a) shows how these two constraints can be captured in conventional DECLARE. This specification is satisfactory only in the case where each trace refers to a single customer and a single order by that customer. For example, consider the following two traces, respectively referring to an order o_1 by Anne, and an order o_2 by Bob:

$$t_1 = \langle \mathsf{sign}, \mathsf{open}, \mathsf{close} \rangle \qquad\qquad t_2 = \langle \mathsf{open}, \mathsf{close}, \mathsf{sign} \rangle$$

Clearly, t_1 is a model trace, while t_2 is not, as the latter violates PRECEDENCE(sign, close).

   However, one may need to consider multiple orders owned by the same or distinct customers, in the common situation where distinct orders may be later bundled together to handle their shipment. In our example, assuming that o_1 and o_2 are later bundled together in a shipment, this would require to combine t_1 and t_2 in a single object-centric trace, suitably extending each event with a reference to the object(s) it operates on. Suppose this would result into:

$$t = \left\langle \begin{array}{l} \mathsf{sign}(\mathsf{customer{=}Anne}), \mathsf{open}(\mathsf{order{=}o2}), \mathsf{open}(\mathsf{order{=}o1}), \\ \mathsf{close}(\mathsf{order{=}o1}), \mathsf{close}(\mathsf{order{=}o2}), \mathsf{sign}(\mathsf{customer{=}Bob}) \end{array} \right\rangle$$

The DECLARE specification of Fig. 14(a) becomes now inadequate. In fact, it cannot distinguish which events actually *co-refer* to one another and which do

not, so it cannot identify that the first signature by Anne refers to the first occurrence of close, but not to the second one. Hence, it wrongly uses the first occurrence of sign to satisfy PRECEDENCE(sign, close) for both orders.   ◁

Fixing the issue described in Example 19 requires the explicitly extension of DECLARE with the ability of expressing how events relate to objects, how objects relate to each other, and in turn to *scope* the application of constraints, expressing that they must be enforced over events that suitably co-refer to each other – either because they operate on the same object, or because they operate on related objects. In our running example, this would call for the following actions:

- introduce the classes of **Order** and **Customer**;
- capture that there is a many-to-one **owned by** association linking orders to customers;
- indicate that sign refers to a customer, and that open and close refer to an order;
- scope PRECEDENCE(open, close) by enforcing that the two involved activities must *co-refer to the same order* (i.e., that an event of activity close for order *o* can only occur if an event of activity open has previously occurred *for the same order*);
- scope PRECEDENCE(sign, close) by enforcing that the two involved activities must respectively operate with a customer and an order that *co-refer through the **owned by** association* (i.e., that an event of activity close for order *o* can only occur if an event of activity sign has previously occurred *for the customer who owns o*).

Object-centric behavioral constraints (OCBC) [4] have been brought forward to handle this type of scoping through the integration of DECLARE specifications and UML class diagrams. Figure 14(b) shows the OCBC specification correctly capturing the constraints of Example 19. The approach is still at its infancy: some first seminal works have been conducted to handle discovery of OCBC specifications from object-centric event logs recording full database transactions [61], and to formalize and reason upon OCBC specifications through temporal description logics [11]. Further research is being carried out to improve the performance of discovery and frame it in the context of object-centric event logs of the form of [2], and to tackle conformance checking and monitoring. This is particularly challenging, as integrating temporal constraints with data models quickly leads to undecidability [11].

7   Conclusion

Throughout this chapter, we have thoroughly reviewed the declarative approach to process specification and mining. The declarative approach aims at limiting the process behavior by defining the boundaries within which its executions can unfold, yet leaving process executors free to explore at runtime which specific executions are generated. This is in contrast with the imperative approach, where process models compactly depict all and only those traces that are admissible.

In fact, notice that different (imperative) process models can comply with the same declarative specification, just like different dynamic systems can model ($\models$) a set of temporal rules. In the chapter, we have grounded our discussion on the Declare language, but the introduced concepts are broad enough to be seamlessly applicable to other related approaches.

Specifically, we have first discussed how declarative process specifications can be formalized using Linear Temporal Logic on Finite Traces (LTL_f), and in turn operationally characterized in terms finite state automata (FSAs) for their execution semantics. On this solid formal ground, we have examined the core reasoning tasks that relate to declarative specifications and then delved deeper into the discovery and monitoring of processes according to the declarative paradigm. Interestingly, we have observed that the reasoning tasks are pervasive in all stages of declarative process mining, such as within discovery to avoid producing redundant or inconsistent outputs, and within monitoring to speculatively consider the possible future continuations of the monitored execution. In the last part of the chapter, we have provided a summary of the most recent advances in declarative process mining, focusing in particular on: *(i)* the applicability of declarative process mining techniques and concepts to full temporal logics, going beyond pre-defined patterns; *(ii)* the incorporation of uncertainty within constraints; *(iii)* the analysis of hybrid models integrating imperative and declarative fragments; *(iv)* multi-perspective constraints incorporating additional dimensions beyond the control-flow, and supporting the declarative specification of object-centric (multi-case) processes. This bird-eye view provides a fair account of the open research challenges in declarative process mining.

References

1. van der Aalst, W.M.P.: Process Mining - Data Science in Action, Second Edition. Springer (2016). https://doi.org/10.1007/978-3-662-49851-4
2. van der Aalst, W.M.P.: Foundations of process discovery. In: van der Aalst and Carmona [6], pp. 37–75. https://doi.org/10.1007/978-3-031-08848-3_2
3. van der Aalst, W.M.P.: Process mining: A 360 degree overview. In: van der Aalst and Carmona [6], pp. 3–34. https://doi.org/10.1007/978-3-031-08848-3_1

4. van der Aalst, W.M.P., Artale, A., Montali, M., Tritini, S.: Object-centric behavioral constraints: Integrating data and declarative process modelling. In: Artale, A., Glimm, B., Kontchakov, R. (eds.) DL. CEUR Workshop Proceedings, vol. 1879. CEUR-WS.org (2017), http://ceur-ws.org/Vol-1879/paper51.pdf

5. van der Aalst, W.M.P., Pesic, M.: DecSerFlow: Towards a truly declarative service flow language. In: Bravetti, M., Núñez, M., Zavattaro, G. (eds.) WS-FM. Lecture Notes in Computer Science, vol. 4184, pp. 1–23. Springer (2006). https://doi.org/10.1007/11841197_1

6. van der Aalst, W.M., Carmona, J. (eds.): Process Mining Handbook, Lecture Notes in Business Information Processing, vol. 448. Springer (2022)

7. Adamo, J.: Data mining for association rules and sequential patterns - sequential and parallel algorithms. Springer New York (2001). https://doi.org/10.1007/978-1-4613-0085-4

8. Alman, A., Di Ciccio, C., Maggi, F.M., Montali, M., van der Aa, H.: Rum: Declarative process mining, distilled. In: Business Process Management - 19th International Conference, BPM 2021, Rome, Italy, September 06-10, 2021, Proceedings. LNCS, vol. 12875, pp. 23–29. Springer (2021)

9. Alman, A., Maggi, F.M., Montali, M., Patrizi, F., Rivkin, A.: Multi-model monitoring framework for hybrid process specifications. In: Franch, X., Poels, G. (eds.) Proceedings of the 34th International Conference on Advanced Information Systems Engineering (CAiSE 2022). Lecture Notes in Computer Science (2022), to appear

10. Alman, A., Maggi, F.M., Montali, M., Peñaloza, R.: Probabilistic declarative process mining. Inf. Syst. (2012), to appear

11. Artale, A., Kovtunova, A., Montali, M., van der Aalst, W.M.P.: Modeling and reasoning over declarative data-aware processes with object-centric behavioral constraints. In: Hildebrandt, T.T., van Dongen, B.F., Röglinger, M., Mendling, J. (eds.) Proceedings of the 17th International Conference on Business Process Management (BPM 2019). Lecture Notes in Computer Science, vol. 11675, pp. 139–156. Springer (2019)

12. Back, C.O., Slaats, T., Hildebrandt, T.T., Marquard, M.: Discover: Accurate & efficient discovery of declarative process models. CoRR abs/2005.10085 (2020), https://arxiv.org/abs/2005.10085

13. Baier, T., Di Ciccio, C., Mendling, J., Weske, M.: Matching events and activities by integrating behavioral aspects and label analysis. Software & Systems Modeling 17(2), 573–598 (May 2018). https://doi.org/10.1007/s10270-017-0603-z, open access

14. Basin, D.A., Klaedtke, F., Müller, S., Zalinescu, E.: Monitoring metric first-order temporal properties. J. ACM 62(2), 15:1–15:45 (2015)

15. Bauer, A., Leucker, M., Schallhart, C.: Runtime verification for LTL and TLTL. ACM Trans. Softw. Eng. Methodol. 20(4), 14:1–14:64 (2011). https://doi.org/10.1145/2000799.2000800

16. Bergami, G., Maggi, F.M., Marrella, A., Montali, M.: Aligning data-aware declarative process models and event logs. In: Polyvyanyy, A., Wynn, M.T., Looy, A.V., Reichert, M. (eds.) Proceedings of the 19th International Conference on Business Process Management (BPM 2021). Lecture Notes in Computer Science, vol. 12875, pp. 235–251. Springer (2021)

17. Burattin, A., Maggi, F.M., Sperduti, A.: Conformance checking based on multi-perspective declarative process models. Expert Syst. Appl. 65, 194–211 (2016). https://doi.org/10.1016/j.eswa.2016.08.040

18. Calvanese, D., De Giacomo, G., Montali, M., Patrizi, F.: Verification and monitoring for first-order ltl with persistence-preserving quantification over finite and infinite traces. In: De Raedt, L. (ed.) Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI 2022). ijcai.org (2022), to appear

19. Camacho, A., McIlraith, S.A.: Learning interpretable models expressed in linear temporal logic. In: Benton, J., Lipovetzky, N., Onaindia, E., Smith, D.E., Srivastava, S. (eds.) Proceedings of the Twenty-Ninth International Conference on Automated Planning and Scheduling (ICAPS 2018). pp. 621–630. AAAI Press (2019)

20. Carmona, J., van Dongen, B.F., Weidlich, M.: Conformance checking: Foundations, milestones and challenges. In: van der Aalst and Carmona [6], pp. 155–190. https://doi.org/10.1007/978-3-031-08848-3_5

21. Cecconi, A., De Giacomo, G., Di Ciccio, C., Mendling, J.: A temporal logic-based measurement framework for process mining. In: van Dongen et al. [45]. https://doi.org/10.1109/ICPM49681.2020.00026

22. Cecconi, A., Di Ciccio, C., De Giacomo, G., Mendling, J.: Interestingness of traces in declarative process mining: The Janus LTLpf approach. In: Weske, M., Montali, M., Weber, I., vom Brocke, J. (eds.) BPM. Lecture Notes in Computer Science, vol. 11080, pp. 121–138. Springer (Sep 2018). https://doi.org/10.1007/978-3-319-98648-7_8

23. Chesani, F., Francescomarino, C.D., Ghidini, C., Loreti, D., Maggi, F.M., Mello, P., Montali, M., Tessaris, S.: Process discovery on deviant traces and other stranger things. CoRR **abs/2109.14883** (2021), https://arxiv.org/abs/2109.14883

24. Chesani, F., Lamma, E., Mello, P., Montali, M., Riguzzi, F., Storari, S.: Exploiting inductive logic programming techniques for declarative process mining. T. Petri Nets and Other Models of Concurrency **2**, 278–295 (2009). https://doi.org/10.1007/978-3-642-00899-3_16, http://dx.doi.org/10.1007/978-3-642-00899-3

25. Chesani, F., Lamma, E., Mello, P., Montali, M., Riguzzi, F., Storari, S.: Exploiting inductive logic programming techniques for declarative process mining. Trans. Petri Nets Other Model. Concurr. **2**, 278–295 (2009)

26. Chomsky, N., Miller, G.A.: Finite state languages. Information and Control **1**(2), 91–112 (1958)

27. Corea, C., Deisen, M., Delfmann, P.: Resolving inconsistencies in declarative process models based on culpability measurement. In: Ludwig, T., Pipek, V. (eds.) WI. pp. 139–153. University of Siegen, Germany / AISeL (2019), https://aisel.aisnet.org/wi2019/track03/papers/2

28. Corea, C., Delfmann, P.: Quasi-inconsistency in declarative process models. In: Hildebrandt, T.T., van Dongen, B.F., Röglinger, M., Mendling, J. (eds.) BPM Forum. Lecture Notes in Business Information Processing, vol. 360, pp. 20–35. Springer (2019). https://doi.org/10.1007/978-3-030-26643-1_2, https://doi.org/10.1007/978-3-030-26643-1

29. Corea, C., Nagel, S., Mendling, J., Delfmann, P.: Interactive and minimal repair of declarative process models. In: Polyvyanyy, A., Wynn, M.T., Looy, A.V., Reichert, M. (eds.) Proceedings of the Business Process Management Forum - BPM Forum 2021. Lecture Notes in Business Information Processing, vol. 427, pp. 3–19. Springer (2021)

30. Davulcu, H., Kifer, M., Ramakrishnan, C.R., Ramakrishnan, I.V.: Logic based modeling and analysis of workflows. In: PODS. pp. 25–33. ACM (1998)

31. De Giacomo, G., De Masellis, R., Grasso, M., Maggi, F.M., Montali, M.: Monitoring business metaconstraints based on LTL & LDL for finite traces. In: Sadiq, S.W.,

Soffer, P., Völzer, H. (eds.) BPM. Lecture Notes in Computer Science, vol. 8659, pp. 1–17. Springer (2014). https://doi.org/10.1007/978-3-319-10172-9_1, http://dx.doi.org/10.1007/978-3-319-10172-9

32. De Giacomo, G., De Masellis, R., Maggi, F.M., Montali, M.: Monitoring constraints and metaconstraints with temporal logics on finite traces. ACM Trans. Softw. Eng. Methodol. (2022), to appear

33. De Giacomo, G., De Masellis, R., Montali, M.: Reasoning on LTL on finite traces: Insensitivity to infiniteness. In: Brodley, C.E., Stone, P. (eds.) AAAI. pp. 1027–1033. AAAI Press (2014), http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8575

34. De Giacomo, G., Vardi, M.Y.: Linear temporal logic and linear dynamic logic on finite traces. In: Rossi, F. (ed.) IJCAI. pp. 854–860. IJCAI/AAAI (2013), http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6997

35. De Smedt, J., De Weerdt, J., Serral, E., Vanthienen, J.: Discovering hidden dependencies in constraint-based declarative process models for improving understandability. Inf. Syst. **74**(Part 1), 40–52 (2018)

36. De Smedt, J., De Weerdt, J., Vanthienen, J., Poels, G.: Mixed-paradigm process modeling with intertwined state spaces. Business & Information Systems Engineering **58**(1), 19–29 (2016). https://doi.org/10.1007/s12599-015-0416-y, http://dx.doi.org/10.1007/s12599-015-0416-y

37. Demri, S., Lazic, R.: LTL with the freeze quantifier and register automata. ACM Trans. Comput. Log. **10**(3), 16:1–16:30 (2009)

38. Demri, S., Lazic, R., Nowak, D.: On the freeze quantifier in constraint LTL: decidability and complexity. Inf. Comput. **205**(1), 2–24 (2007)

39. Di Ciccio, C.: On the Mining of Artful Processes. Ph.D. thesis, SAPIENZA, University of Rome (Oct 2013), http://hdl.handle.net/11573/966501

40. Di Ciccio, C., Bernardi, M.L., Cimitile, M., Maggi, F.M.: Generating event logs through the simulation of Declare models. In: Barjis, J., Pergl, R., Babkin, E. (eds.) EOMAS. Lecture Notes in Business Information Processing, vol. 231, pp. 20–36. Springer (Jun 2015). https://doi.org/10.1007/978-3-319-24626-0_2

41. Di Ciccio, C., Maggi, F.M., Montali, M., Mendling, J.: Resolving inconsistencies and redundancies in declarative process models. Information Systems **64**, 425–446 (Mar 2017). https://doi.org/10.1016/j.is.2016.09.005

42. Di Ciccio, C., Maggi, F.M., Montali, M., Mendling, J.: On the relevance of a business constraint to an event log. Information Systems **78**, 144–161 (Nov 2018). https://doi.org/10.1016/j.is.2018.01.011

43. Di Ciccio, C., Mecella, M.: On the discovery of declarative control flows for artful processes. ACM Trans. Manage. Inf. Syst. **5**(4), 24:1–24:37 (Jan 2015). https://doi.org/10.1145/2629447

44. van Dongen, B.F., De Smedt, J., Di Ciccio, C., Mendling, J.: Conformance checking of mixed-paradigm process models. Inf. Syst. **102**, 101685 (2021)

45. van Dongen, B.F., Montali, M., Wynn, M.T. (eds.): 2nd International Conference on Process Mining, ICPM 2020, Padua, Italy, October 4-9, 2020. IEEE (2020)

46. Dwyer, M.B., Avrunin, G.S., Corbett, J.C.: Patterns in property specifications for finite-state verification. In: Boehm, B.W., Garlan, D., Kramer, J. (eds.) ICSE. pp. 411–420. ACM (1999). https://doi.org/10.1145/302405.302672

47. Elgammal, A., Türetken, O., van den Heuvel, W., Papazoglou, M.P.: Formalizing and appling compliance patterns for business process compliance. Softw. Syst. Model. **15**(1), 119–146 (2016)

48. Fahland, D.: Process mining over multiple behavioral dimensions with event knowledge graphs. In: van der Aalst and Carmona [6], pp. 274–319. https://doi.org/10.1007/978-3-031-08848-3_9
49. Fionda, V., Greco, G.: LTL on finite and process traces: Complexity results and a practical reasoner. J. Artif. Intell. Res. **63**, 557–623 (2018)
50. Fionda, V., Guzzo, A.: Control-flow modeling with declare: Behavioral properties, computational complexity, and tools. IEEE Trans. Knowl. Data Eng. **32**(5), 898–911 (2020)
51. Goedertier, S., Martens, D., Vanthienen, J., Baesens, B.: Robust process discovery with artificial negative events. Journal of Machine Learning Research **10**, 1305–1340 (2009). https://doi.org/10.1145/1577069.1577113
52. Green, T.R.G., Petre, M.: Usability analysis of visual programming environments: A 'cognitive dimensions' framework. Vis. Comp. and Lang. **7**(2), 131–74 (1996)
53. Haisjackl, C., Barba, I., Zugal, S., Soffer, P., Hadar, I., Reichert, M., Pinggera, J., Weber, B.: Understanding declare models: strategies, pitfalls, empirical results. Softw. Syst. Model. **15**(2), 325–352 (2016)
54. Hildebrandt, T.T., Mukkamala, R.R.: Declarative event-based workflow as distributed dynamic condition response graphs. In: PLACES. EPTCS, vol. 69, pp. 59–73 (2010)
55. Hopcroft, J.E., Motwani, R., Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation (3rd Edition). Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA (2006)
56. Kupferman, O., Vardi, M.Y.: Vacuity detection in temporal model checking. International Journal on Software Tools for Technology Transfer **4**(2), 224–233 (2003). https://doi.org/c
57. Lamma, E., Mello, P., Montali, M., Riguzzi, F., Storari, S.: Inducing declarative logic-based models from labeled traces. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM. Lecture Notes in Computer Science, vol. 4714, pp. 344–359. Springer (2007). https://doi.org/10.1007/978-3-540-75183-0_25
58. Lemieux, C., Park, D., Beschastnikh, I.: General LTL specification mining (T). In: Cohen, M.B., Grunske, L., Whalen, M. (eds.) 30th IEEE/ACM International Conference on Automated Software Engineering, ASE 2015, Lincoln, NE, USA, November 9-13, 2015. pp. 81–92. IEEE Computer Society (2015)
59. Leno, V., Dumas, M., Maggi, F.M., Rosa, M.L., Polyvyanyy, A.: Automated discovery of declarative process models with correlated data conditions. Inf. Syst. **89**, 101482 (2020)
60. de Leoni, M., Maggi, F.M., van der Aalst, W.M.P.: An alignment-based framework to check the conformance of declarative process models and to preprocess event-log data. Information Systems **47**, 258 – 277 (2015). https://doi.org/10.1016/j.is.2013.12.005
61. Li, G., Medeiros de Carvalho, R., van der Aalst, W.M.P.: Automatic discovery of object-centric behavioral constraint models. In: Abramowicz, W. (ed.) Proceedings of the 20th International Conference on Business Information Systems (BIS 201). Lecture Notes in Business Information Processing, vol. 288, pp. 43–58. Springer (2017)
62. Lichtenstein, O., Pnueli, A., Zuck, L.D.: The glory of the past. In: Parikh, R. (ed.) Logics of Programs. Lecture Notes in Computer Science, vol. 193, pp. 196–218. Springer (1985). https://doi.org/10.1007/3-540-15648-8_16
63. Ly, L.T., Maggi, F.M., Montali, M., Rinderle-Ma, S., van der Aalst, W.M.P.: Compliance monitoring in business processes: Functionalities, application, and tool-support. Inf. Syst. **54**, 209–234 (2015). https://doi.org/10.1016/j.is.2015.02.007

64. Maggi, F.M., Bose, R.P.J.C., van der Aalst, W.M.P.: Efficient discovery of under-standable declarative process models from event logs. In: Ralyté, J., Franch, X., Brinkkemper, S., Wrycza, S. (eds.) CAiSE. Lecture Notes in Computer Science, vol. 7328, pp. 270–285. Springer (2012). https://doi.org/10.1007/978-3-642-31095-9_18, http://dx.doi.org/10.1007/978-3-642-31095-9

65. Maggi, F.M., Di Ciccio, C., Di Francescomarino, C., Kala, T.: Parallel algorithms for the automated discovery of declarative process models. Information Systems **74**, 136–152 (May 2018). https://doi.org/10.1016/j.is.2017.12.002

66. Maggi, F.M., Dumas, M., García-Bañuelos, L., Montali, M.: Discovering data-aware declarative process models from event logs. In: Daniel, F., Wang, J., Weber, B. (eds.) BPM. Lecture Notes in Computer Science, vol. 8094, pp. 81–96. Springer (2013). https://doi.org/10.1007/978-3-642-40176-3_8

67. Maggi, F.M., Montali, M., van der Aalst, W.M.P.: An operational decision support framework for monitoring business constraints. In: de Lara, J., Zisman, A. (eds.) FASE. Lecture Notes in Computer Science, vol. 7212, pp. 146–162. Springer (2012). https://doi.org/10.1007/978-3-642-28872-2_11

68. Maggi, F.M., Montali, M., Peñaloza, R.: Temporal logics over finite traces with uncertainty. In: Proceedings of the 34 AAAI Conference on Artificial Intelligence (AAAI 2020). pp. 10218–10225. AAAI Press (2020)

69. Maggi, F.M., Montali, M., Peñaloza, R., Alman, A.: Extending temporal business constraints with uncertainty. In: Fahland, D., Ghidini, C., Becker, J., Dumas, M. (eds.) Proceedings of the 18th International Conference on Business Process Management (BPM 2020). Lecture Notes in Computer Science, vol. 12168, pp. 35–54. Springer (2020)

70. Maggi, F.M., Montali, M., Westergaard, M., van der Aalst, W.M.P.: Monitoring business constraints with Linear Temporal Logic: An approach based on colored automata. In: Rinderle-Ma et al. [90], pp. 132–147. https://doi.org/10.1007/978-3-642-23059-2_13

71. Maggi, F.M., Mooij, A.J., van der Aalst, W.M.P.: User-guided discovery of declarative process models. In: CIDM. pp. 192–199. IEEE (2011). https://doi.org/10.1109/CIDM.2011.5949297

72. Maggi, F.M., Westergaard, M., Montali, M., van der Aalst, W.M.P.: Runtime verification of ltl-based declarative process models. In: Proceedings of the 2nd International Conference on Runtime (RV). LNCS, vol. 7186, pp. 131–146. Springer (2011)

73. Montali, M.: Specification and Verification of Declarative Open Interaction Models - A Logic-based framework. Ph.D. thesis, University of Bologna, Italy (2009), http://amsdottorato.unibo.it/1829/

74. Montali, M.: Specification and Verification of Declarative Open Interaction Models: a Logic-Based Approach, Lecture Notes in Business Information Processing, vol. 56. Springer (2010). https://doi.org/10.1007/978-3-642-14538-4

75. Montali, M., Maggi, F.M., Chesani, F., Mello, P., van der Aalst, W.M.: Monitoring business constraints with the event calculus. ACM TIST **5**(1), 17:1–17:30 (2013). https://doi.org/10.1145/2542182.2542199

76. Montali, M., Pesic, M., van der Aalst, W.M.P., Chesani, F., Mello, P., Storari, S.: Declarative specification and verification of service choreographies. TWEB **4**(1) (2010). https://doi.org/10.1145/1658373.1658376, http://doi.acm.org/10.1145/1658373.1658376

77. Montali, M., Torroni, P., Alberti, M., Chesani, F., Gavanelli, M., Lamma, E., Mello, P.: Verification from declarative specifications using logic programming.

In: de la Banda, M.G., Pontelli, E. (eds.) Proceedings of the 24th International Conference on Logic Programming (ICLP 2008). Lecture Notes in Computer Science, vol. 5366, pp. 440–454. Springer (2008)

78. Mulyar, N., Pesic, M., van der Aalst, W.M.P., Peleg, M.: Declarative and procedural approaches for modelling clinical guidelines: Addressing flexibility issues. In: ter Hofstede, A.H.M., Benatallah, B., Paik, H. (eds.) Business Process Management Workshops, BPM 2007 International Workshops, BPI, BPD, CBP, ProHealth, Ref-Mod, semantics4ws, Brisbane, Australia, September 24, 2007, Revised Selected Papers. LNCS, vol. 4928, pp. 335–346. Springer (2007). https://doi.org/10.1007/978-3-540-78238-4_35, https://doi.org/10.1007/978-3-540-78238-4_35

79. Munoz-Gama, J., Martin, N., et al.: Process mining for healthcare: Characteristics and challenges. Journal of Biomedical Informatics **127**, 103994 (Mar 2022). https://doi.org/10.1016/j.jbi.2022.103994

80. Ouaknine, J., Worrell, J.: On the decidability and complexity of metric temporal logic over finite words. Log. Methods Comput. Sci. **3**(1) (2007)

81. Pesic, M.: Constraint-based Workflow Management Systems: Shifting Control to Users. Ph.D. thesis, Technische Universiteit Eindhoven (10 2008), http://repository.tue.nl/638413

82. Pesic, M., van der Aalst, W.M.P.: A declarative approach for flexible business processes management. In: Eder, J., Dustdar, S. (eds.) Business Process Management Workshops. Lecture Notes in Computer Science, vol. 4103, pp. 169–180. Springer (2006), http://dx.doi.org/10.1007/11837862_18

83. Pesic, M., Schonenberg, H., van der Aalst, W.M.P.: DECLARE: Full support for loosely-structured processes. In: EDOC. pp. 287–300 (2007)

84. Pesic, M., Schonenberg, H., van der Aalst, W.M.P.: DECLARE: Full support for loosely-structured processes. In: EDOC. pp. 287–300. IEEE Computer Society (2007). https://doi.org/10.1109/EDOC.2007.14

85. Pill, I., Quaritsch, T.: Behavioral diagnosis of LTL specifications at operator level. In: Rossi, F. (ed.) Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI 2013). pp. 1053–1059. IJCAI/AAAI (2013)

86. Pnueli, A.: The temporal logic of programs. In: FOCS. pp. 46–57. IEEE (1977). https://doi.org/10.1109/SFCS.1977.32

87. Rabin, M.O., Scott, D.S.: Finite automata and their decision problems. IBM Journal of Research and Development **3**(2), 114–125 (April 1959). https://doi.org/10.1147/rd.32.0114

88. Raha, R., Roy, R., Fijalkow, N., Neider, D.: Scalable anytime algorithms for learning fragments of linear temporal logic. In: Fisman, D., Rosu, G. (eds.) Proceedings of the 28th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2022). Lecture Notes in Computer Science, vol. 13243, pp. 263–280. Springer (2022)

89. Reichert, M., Weber, B.: Enabling Flexibility in Process-Aware Information Systems - Challenges, Methods, Technologies. Springer (2012)

90. Rinderle-Ma, S., Toumani, F., Wolf, K. (eds.): Business Process Management - 9th International Conference, BPM 2011, Clermont-Ferrand, France, August 30 - September 2, 2011. Proceedings, Lecture Notes in Computer Science, vol. 6896. Springer (2011)

91. Sadiq, S.W., Sadiq, W., Orlowska, M.E.: Pockets of flexibility in workflow specification. In: ER. LNCS, vol. 2224, pp. 513–526. Springer (2001)

92. Schönig, S., Di Ciccio, C., Maggi, F.M., Mendling, J.: Discovery of multi-perspective declarative process models. In: Sheng, Q.Z., Stroulia, E., Tata, S., Bhiri, S. (eds.)

ICSOC. Lecture Notes in Computer Science, vol. 9936, pp. 87–103. Springer (Oct 2016). https://doi.org/10.1007/978-3-319-46295-0_6

93. Schunselaar, D.M.M., Slaats, T., Maggi, F.M., Reijers, H.A., van der Aalst, W.M.P.: Mining hybrid business process models: A quest for better precision. In: Abramowicz, W., Paschke, A. (eds.) Proceedings of the 21st International Conference on Business Information Systems (BIS 2018). Lecture Notes in Business Information Processing, vol. 320, pp. 190–205. Springer (2018). https://doi.org/10.1007/978-3-319-93931-5_14, https://doi.org/10.1007/978-3-319-93931-5_14

94. Singh, M.P.: Distributed enactment of multiagent workflows: Temporal logic for web service composition. In: AAMAS. pp. 907–914. ACM (2003)

95. Slaats, T., Debois, S., Olling Back, C.: Weighing the pros and cons: Process discovery with negative examples. In: Proceedings of the 19th International Conference on Business Process Management (BPM). LNCS, vol. 12875, pp. 47–64. Springer (2021)

96. Slaats, T., Schunselaar, D.M.M., Maggi, F.M., Reijers, H.A.: The semantics of hybrid process models. In: Debruyne, C., Panetto, H., Meersman, R., Dillon, T.S., Kühn, E., O'Sullivan, D., Ardagna, C.A. (eds.) Proceedings of the Confederated International Conferences On the Move to Meaningful Internet Systems (OTM 2016). Lecture Notes in Computer Science, vol. 10033, pp. 531–551 (2016)

97. Sun, Y., Su, J.: Conformance for decserflow constraints. In: Franch, X., Ghose, A.K., Lewis, G.A., Bhiri, S. (eds.) Proceedings of the 12th International Conference on Service-Oriented Computing (ICSOC 2014). Lecture Notes in Computer Science, vol. 8831, pp. 139–153. Springer (2014)

98. Westergaard, M.: Better algorithms for analyzing and enacting declarative workflow languages using ltl. In: Rinderle-Ma et al. [90], pp. 83–98, http://dx.doi.org/10.1007/978-3-642-23059-2_10

99. Westergaard, M., Maggi, F.M.: Looking into the future. using timed automata to provide a priori advice about timed declarative process models. In: Meersman, R., Panetto, H., Dillon, T.S., Rinderle-Ma, S., Dadam, P., Zhou, X., Pearson, S., Ferscha, A., Bergamaschi, S., Cruz, I.F. (eds.) On the Move to Meaningful Internet Systems: OTM 2012, Confederated International Conferences: CoopIS, DOA-SVI, and ODBASE 2012, Rome, Italy, September 10-14, 2012. Proceedings, Part I. Lecture Notes in Computer Science, vol. 7565, pp. 250–267. Springer (2012). https://doi.org/10.1007/978-3-642-33606-5_16, http://dx.doi.org/10.1007/978-3-642-33606-5_16

100. Zhu, S., Tabajara, L.M., Pu, G., Vardi, M.Y.: On the power of automata minimization in temporal synthesis. In: Proceedings 12th International Symposium on Games, Automata, Logics, and Formal Verification (GandALF 2021). EPTCS, vol. 346, pp. 117–134 (2021)