LTLf and LDLf Monitoring
A Technical Report

Giuseppe De Giacomo1, Riccardo De Masellis1, Marco Grasso1, Fabrizio Maria Maggi2, and Marco Montali3

1 Sapienza Università di Roma, Via Ariosto, 25, 00185 Rome, Italy
(degiacomo|demasellis)@dis.uniroma1.it
2 University of Tartu, J. Liivi 2, 50409 Tartu, Estonia
f.m.maggi@ut.ee
3 Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy
montali@inf.unibz.it

Abstract. Runtime monitoring is one of the central tasks to provide operational decision support to running business processes, and check on-the-fly whether they comply with constraints and rules. We study runtime monitoring of properties expressed in LTL on finite traces (LTLf) and in its extension LDLf. LDLf is a powerful logic that captures all monadic second order logic on finite traces, which is obtained by combining regular expressions and LTLf, adopting the syntax of propositional dynamic logic (PDL). Interestingly, in spite of its greater expressivity, LDLf has exactly the same computational complexity of LTLf. We show that LDLf is able to capture, in the logic itself, not only the constraints to be monitored, but also the de-facto standard RV-LTL monitors. This makes it possible to declaratively capture monitoring metaconstraints, and check them by relying on usual logical services instead of ad-hoc algorithms. This, in turn, enables to flexibly monitor constraints depending on the monitoring state of other constraints, e.g., “compensation” constraints that are only checked when others are detected to be violated. In addition, we devise a direct translation of LDLf formulas into nondeterministic automata, avoiding to detour to Büchi automata or alternating automata, and we use it to implement a monitoring plug-in for the ProM suite.

1 Introduction

Runtime monitoring is one of the central tasks to provide operational decision support to running business processes, and check on-the-fly whether they comply with constraints and rules. In order to provide well-founded and provably correct runtime monitoring techniques, this area is usually rooted into that of verification, the branch of formal analysis aiming at checking whether a system meets some property of interest. Being the system dynamic, properties are usually expressed by making use of modal operators accounting for the time.

Among all the temporal logics used in verification, Linear-time Temporal Logic (LTL) is particularly suited for monitoring, as an actual system execution is indeed linear. However, the LTL semantics is given in terms of infinite traces, hence monitoring must check whether the current trace is a prefix of an infinite trace, that will never be completed [2]. In several context, and in particular often in BPM, we can assume that
the trace of the system is if face finite [17]. For this reason, finite-trace variant of the LTL have been introduced. Here we use the logic $\text{LTL}_f$ (LTL on finite traces), investigated in detail in [4], and at the base of one of the main declarative process modeling approaches: DECLARE [17][15][9]. Following [9], monitoring in $\text{LTL}_f$ amounts to checking whether the current execution belongs to the set of admissible prefixes for the traces of a given $\text{LTL}_f$ formula $\varphi$. To achieve such a task, $\varphi$ is usually first translated into a finite-state automaton for $\varphi$, which recognizes all those finite executions that satisfy $\varphi$.

Despite the presence of previous operational decision support techniques to monitoring $\text{LTL}_f$ constraints over finite traces [9][10], two main challenges have not yet been tackled in a systematic way. First of all, several alternative semantics have been proposed to make LTL suitable for runtime verification (such as the de-facto standard RV monitor conditions [2]), but no comprehensive technique based on finite-state automata is available to accommodate them. On the one hand, runtime verification for such logics typically considers finite partial traces whose continuation is however infinite [2], with the consequence that the corresponding techniques detour to Büchi automata for building the monitors. On the other hand, the incorporation of such semantics in the BPM setting (where also continuations are finite) has only been tackled so far with effective but ad-hoc techniques (cf. the “coloring” of automata in [9] to support the RV conditions), without a corresponding formal underpinning.

A second, key challenge is the incorporation of advanced forms of monitoring, where some constraints become of interest only in specific, critical circumstances (such as the violation of other constraints). This is the basis for supporting monitoring of compensation constraints and so-called contrary-to-duty obligations [19], i.e., obligations that are put in place only when other obligations have not been fulfilled. While this feature is considered to be a fundamental compliance monitoring functionality [8], it is still an open challenge, without any systematic approach able to support it at the level of the constraint specification language.

In this paper, we attack these two challenges by studying runtime monitoring of properties expressed in $\text{LTL}_f$ and in its extension $\text{LDL}_f$ [4]. $\text{LDL}_f$ is a powerful logic that captures all monadic second order logic on finite traces, which is obtained by combining regular expressions and $\text{LTL}_f$, adopting the syntax of propositional dynamic logic (PDL). Interestingly, in spite of its greater expressivity, $\text{LDL}_f$ has exactly the same computational complexity of $\text{LTL}_f$. We show that $\text{LDL}_f$ is able to capture, in the logic itself, not only the usual $\text{LDL}_f$ constraints to be monitored, but also the de-facto standard RV monitor conditions. Indeed given an $\text{LDL}_f$ formula $\phi$, we show how to construct the $\text{LDL}_f$ formulas that captures whether prefixes of $\phi$ satisfy the various RV monitor conditions. This, in turn, makes it possible to declaratively capture monitoring metaconstraints, and check them by relying on usual logical services instead of ad-hoc algorithms. Metaconstraints provide a well-founded, declarative basis to specify and monitor constraints depending on the monitoring state of other constraints, such as “compensation” constraints that are only checked when others are violated.

Interestingly, in doing so we devise a direct translation of $\text{LDL}_f$ (and hence of $\text{LTL}_f$) formulas into nondeterministic automata, which avoid the usual detour to Büchi automata. The technique is grounded on alternating automata (AFW), but it actually avoids also their introduction all together, and directly produces a standard nondeterministic finite-state automaton (NFA). Notably, such technique has been concretely
implemented, and embedded into a monitoring plug-in for the PROM, which supports checking LDLₚ constraints and metaconstraints.

2 LTLₚ and LDLₚ

In this paper we will adopt the standard LTL and its variant LDL interpreted over on finite runs.

LTL on finite traces, called LTLₚ [4], has exactly the same syntax as LTL on infinite traces [18]. Namely, given a set of P of propositional symbols, LTLₚ formulas are obtained through the following:

$$\varphi ::= \phi \mid \neg \varphi \mid \varphi_1 \land \varphi_2 \mid \varphi_1 \lor \varphi_2 \mid \varnothing \varphi \mid \bullet \varphi \mid \Diamond \varphi \mid \Box \varphi \mid \varphi_1 \mathcal{U} \varphi_2$$

where $\phi$ is a propositional formula over $P$, $\varnothing$ is the next operator, $\bullet$ is weak next, $\Diamond$ is eventually, $\Box$ is always, $\mathcal{U}$ is until.

It is known that LTLₚ is as expressive as First Order Logic over finite traces, so strictly less expressive than regular expressions which in turn are as expressive as Monadic Second Order logic over finite traces. On the other hand, regular expressions are a too low level formalism for expressing temporal specifications, since, for example, they miss a direct construct for negation and for conjunction [4].

To overcome this difficulties, in [4] Linear Dynamic Logic of Finite Traces, or LDLₚ, has been proposed. This logic is as natural as LTLₚ but with the full expressive power of Monadic Second Order logic over finite traces. LDLₚ is obtained by merging LTLₚ with regular expression through the syntax of the well-know logic of programs PDL, Propositional Dynamic Logic, [6,7] but adopting a semantics based on finite traces. This logic is an adaptation of LDL, introduced in [21], which, like LTL, is interpreted over infinite traces.

Formally, LDLₚ formulas are built as follows:

$$\varphi ::= \phi \mid tt \mid ff \mid \neg \varphi \mid \varphi_1 \land \varphi_2 \mid \varphi_1 \lor \varphi_2 \mid \langle \rho \rangle \varphi \mid [\rho] \varphi \mid \rho$$

$$\rho ::= \phi \mid \varnothing \varphi \mid \rho_1 + \rho_2 \mid \rho_1 ; \rho_2 \mid \rho^*$$

where $\phi$ is a propositional formula over $P$; $tt$ and $ff$ denote respectively the true and the false LDLₚ formula (not to be confused with the propositional formula true and false). $\rho$ denotes path expressions, which are REₚ expressions over propositional formulas $\phi$, with the addition of the test construct $\varnothing \varphi$ typical of PDL; and $\varnothing \varphi$ stand for LDLₚ formulas built by applying boolean connectives and the modal connectives $\langle \rho \rangle \varphi$ and $[\rho] \varphi$. In fact $[\rho] \varphi \equiv \neg \langle \rho \rangle \neg \varphi$.

Intuitively, $\langle \rho \rangle \varphi$ states that, from the current step in the trace, there exists an execution satisfying the regular expression $\rho$ such that its last step satisfies $\varphi$. While $[\rho] \varphi$ states that, from the current step, all executions satisfying the regular expression $\rho$ are such that their last step satisfies $\varphi$. Tests are used to insert into the execution path checks for satisfaction of additional LDLₚ formulas.

As for LTLₚ, the semantics of LDLₚ is given in terms of finite traces denoting a finite, possibly empty, sequence of consecutive steps in the trace, i.e., finite words $\pi$ over the alphabet of $2^P$, containing all possible propositional interpretations of the propositional symbols in $P$. We denote by $\pi(i)$ the $i$th step in the trace. If the trace is shorter and does
not include an \(i\)th step \(\pi(i)\) is undefined. We denote by \(\pi(i, j)\) the segment of the trace \(\pi\) starting at \(i\)th step and ending at the \(j\)th step (included). If \(i\) or \(j\) are out of range wrt the trace that \(\pi(i, j)\) is undefined, except \(\pi(i, i) = \epsilon\) (i.e., the empty trace).

The semantics of \(\text{LDL}_f\) is as follows: an \(\text{LDL}_f\) formula \(\varphi\) is true at a step \(i\), in symbols \(\pi, i \models \varphi\), as follows:

- \(\pi, i \models \top\)
- \(\pi, i \not\models \bot\)
- \(\pi, i \models \phi\) iff \(1 \leq i \leq n\) and \(\pi(i) \models \phi\) (\(\phi\) propositional).
- \(\pi, i \models \neg \varphi\) iff \(\pi, i \not\models \varphi\).
- \(\pi, i \models \varphi_1 \land \varphi_2\) iff \(\pi, i \models \varphi_1\) and \(\pi, i \models \varphi_2\).
- \(\pi, i \models \varphi_1 \lor \varphi_2\) iff \(\pi, i \models \varphi_1\) or \(\pi, i \models \varphi_2\).
- \(\pi, i \models \langle \rho \rangle \varphi\) iff for some \(i\) we have \(\pi(i, j) \in \mathbb{L}(\rho)\) and \(\pi, j \models \varphi\).
- \(\pi, i \models [\rho] \varphi\) iff for all \(i\) such that \(\pi(i, j) \in \mathbb{L}(\rho)\) we have \(\pi, j \models \varphi\).

The relation \(\pi(i, j) \in \mathbb{L}(\rho)\) is defined inductively as follows:

- \(\pi(i, j) \in \mathbb{L}(\phi)\) if \(j = i + 1 \leq n\) and \(\pi(i) \models \phi\) (\(\phi\) propositional).
- \(\pi(i, j) \in \mathbb{L}(\varphi')\) if \(j = i\) and \(\pi, i \models \varphi\).
- \(\pi(i, j) \in \mathbb{L}(\rho_1 + \rho_2)\) if \(\pi(i, j) \in \mathbb{L}(\rho_1)\) or \(\pi(i, j) \in \mathbb{L}(\rho_2)\).
- \(\pi(i, j) \in \mathbb{L}(\rho_1; \rho_2)\) if there exists \(k\) s.t. \(\pi(i, k) \in \mathbb{L}(\rho_1)\) and \(\pi(k, j) \in \mathbb{L}(\rho_2)\).
- \(\pi(i, j) \in \mathbb{L}(\rho^*)\) if \(j = i\) or there exists \(k\) s.t. \(\pi(i, k) \in \mathbb{L}(\rho)\) and \(\pi(k, j) \in \mathbb{L}(\rho^*)\).

Observe that for \(i > n\), hence e.g., for \(\pi \equiv \epsilon\) we get:

- \(\pi, i \models \top\)
- \(\pi, i \not\models \bot\)
- \(\pi, \epsilon \models \varphi\) (\(\varphi\) propositional).
- \(\pi, \epsilon \not\models \neg \varphi\)
- \(\pi, \epsilon \models \varphi_1 \land \varphi_2\)
- \(\pi, \epsilon \models \varphi_1 \lor \varphi_2\)
- \(\pi, \epsilon \models \langle \rho \rangle \varphi\)
- \(\pi, \epsilon \models [\rho] \varphi\)

The relation \(\pi(i, i) \in \mathbb{L}(\rho)\) with \(i > n\) is defined inductively as follows:

- \(\pi(i, i) \not\in \mathbb{L}(\phi)\) (\(\phi\) propositional).
- \(\pi(i, i) \in \mathbb{L}(\varphi')\) if \(\pi, i \models \varphi\).
- \(\pi(i, i) \in \mathbb{L}(\rho_1 + \rho_2)\) if \(\pi(i, i) \in \mathbb{L}(\rho_1)\) or \(\pi(i, i) \in \mathbb{L}(\rho_2)\).
- \(\pi(i, i) \in \mathbb{L}(\rho_1; \rho_2)\) if \(\pi(i, i) \in \mathbb{L}(\rho_1)\) and \(\pi(i, i) \in \mathbb{L}(\rho_2)\).
- \(\pi(i, i) \in \mathbb{L}(\rho^*)\)

Notice we have the usual boolean equivalences such as \(\varphi_1 \lor \varphi_2 \equiv \neg \varphi_1 \land \neg \varphi_2\), furthermore we have that: \(\phi \equiv \langle \phi \rangle \top\), and \([\rho] \varphi \equiv \neg \langle \rho \rangle \neg \varphi\). It is also convenient to introduce the following abbreviations:

- \(\text{end} = \langle \text{true}\rangle \top\) that denotes that the traces is been completed (the remaining trace is \(\epsilon\) the empty one)
- \(\text{last} = \langle \text{true}\rangle \text{end}\) which denotes the last step of the trace.

It easy to encode \(\text{LTL}_f\) into \(\text{LDL}_f\): it suffice to observe that we can express the various \(\text{LTL}_f\) operators by recursively applying the following translations:

- \(\Diamond \varphi\) translates to \(\langle \text{true} \rangle \varphi\);
- \(\bullet \varphi\) translates to \(\neg \langle \text{true} \rangle \neg \varphi = [\text{true}] \varphi\) (notice that \(\bullet \alpha\) is translated into \([\text{true}] [\neg \alpha] \top\), since \(\alpha\) is equivalent to \(\langle \alpha \rangle \top\);
- \(\Diamond \varphi\) translates to \(\langle \text{true}^* \rangle \varphi\);
- \(\Box \varphi\) translates to \([\text{true}^*] \varphi\) (notice that \(\Box \alpha\) is translated into \([\text{true}^*] [\neg \alpha] \top\);
- \( \varphi_1 \cup \varphi_2 \) translates to \( (\varphi_1; \text{true})^* \varphi_2 \).

It is also easy to encode regular expressions, used as a specification formalism for traces into LDL: \( \rho \) translates to \( (\rho)^\text{end} \).

We say that a trace satisfies an LTL\(_f\) or LDL\(_f\) formula \( \varphi \), written \( \pi \models \varphi \) if \( \pi, 1 \models \varphi \).

(Note that if \( \pi \) is the empty trace, and hence 1 is out of range, still the notion of \( \pi, 1 \models \varphi \) is well defined). Also sometimes we denote by \( \mathcal{L}(\varphi) \) the set of traces that satisfy \( \varphi \): \( \mathcal{L}(\varphi) = \{ \pi \mid \pi \models \varphi \} \).

### 3 LDL\(_f\) Automaton

We can associate with each LDL\(_f\) formula \( \varphi \) an (exponential) NFA \( A_\varphi \) that accepts exactly the traces that make \( \varphi \) true. Here, we provide a simple direct algorithm for computing the NFA corresponding to an LDL\(_f\) formula. The correctness of the algorithm is based on the fact that (i) we can associate with each LDL\(_f\) formula \( \varphi \) a polynomial alternating automaton on words (AFW) \( A_\varphi \) which accepts exactly the traces that make \( \varphi \) true \([4]\), and (ii) every AFW can be transformed into an NFA, see, e.g., \([4]\). However, to formulate the algorithm we do not need these notions, but we can work directly on the LDL\(_f\) formula. In order to proceed with the construction of the AFW \( A_\varphi \), we put LDL\(_f\) formulas \( \varphi \) in negation normal form \( \text{nff}(\varphi) \) by exploiting equivalences and pushing negation inside as much as possible, until is eliminated except in propositional formulas. Note that computing \( \text{nff}(\varphi) \) can be done in linear time. In other words, wlog, we consider as syntax for LDL\(_f\) the one in the previous section but without negation. Then we define an auxiliary function \( \delta \) that takes an LDL\(_f\) formula \( \psi \) (in negation normal form) and a propositional interpretation \( \Pi \) for \( \mathcal{P} \) (including last), or a special symbol \( \epsilon \), returning a positive boolean formula whose atoms are (quoted) \( \psi \) subformulas.

\[
\begin{align*}
\delta("\text{tt}"; \Pi) &= \text{true} \\
\delta("\text{ff}"; \Pi) &= \text{false} \\
\delta("\phi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \models \phi \\
\text{false} & \text{if } \Pi \not\models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\varphi_1 \land \varphi_2"; \Pi) &= \delta("\varphi_1"; \Pi) \land \delta("\varphi_2"; \Pi) \\
\delta("\varphi_1 \lor \varphi_2"; \Pi) &= \delta("\varphi_1"; \Pi) \lor \delta("\varphi_2"; \Pi) \\
\delta("\langle \psi? \rangle \varphi"; \Pi) &= \delta("\psi"; \Pi) \land \delta("\varphi"; \Pi) \\
\delta("\langle \rho_1; \rho_2 \rangle \varphi"; \Pi) &= \delta("\langle \rho_1 \rangle \varphi"; \Pi) \lor \delta("\langle \rho_2 \rangle \varphi"; \Pi) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\delta("\varphi"; \Pi) & \text{if } \rho \text{ is test-only} \\
\delta("\varphi"; \Pi) \lor \delta("\langle \rho \rangle \varphi"; \Pi) & \text{else} 
\end{cases} \\
\delta("\langle \rho^* \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho^* \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho^* \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho^* \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) \\
\delta("\langle \rho \rangle \varphi"; \Pi) &= \begin{cases} 
\text{true} & \text{if } \Pi \not\models \phi \\
\text{false} & \text{if } \Pi \models \phi 
\end{cases} \quad (\phi \text{ propositional}) 
\end{align*}
\]
1: **algorithm LDL\(_f\)2NFA()**
2: **input** LTL\(_f\) formula \(\varphi\)
3: **output** NFA \(A_\varphi = (2^\mathcal{P}, S, \{s_0\}, \varrho, \{s_f\})\)
4: \(s_0 \leftarrow \{"\varphi^n\}\) \hspace{1cm} \(\triangleright\) single initial state
5: \(s_f \leftarrow \emptyset\) \hspace{1cm} \(\triangleright\) single final state
6: \(S \leftarrow \{s_0, s_f\}, \varrho \leftarrow \emptyset\)
7: **while** (\(S\) or \(\varrho\) change) **do**
8: \(\triangleright\) update set of states
9: \(S \leftarrow S \cup \{q'\}\)
10: \(\varrho \leftarrow \varrho \cup \{(q, \Theta, q')\}\) \hspace{1cm} \(\triangleright\) update transition relation

**Fig. 1.** NFA construction

\[
\delta("\psi?\varphi^n\), \Pi) = \delta("\mathrm{nnf}(\neg\psi)^n\), \Pi) \lor \delta("\varphi^n\), \Pi) \\
\delta("[\rho_1 + \rho_2]\varphi^n\), \Pi) = \delta("[\rho_1]\varphi^n\), \Pi) \land \delta("[\rho_2]\varphi^n\), \Pi) \\
\delta("[\rho_1; \rho_2]\varphi^n\), \Pi) = \delta("[\rho_1]\varphi^n\), \Pi) \land \delta("[\rho_2]\varphi^n\), \Pi) \\
\delta("[\rho]^*\varphi^n\), \Pi) = \begin{cases} 
\delta("\varphi^n\), \Pi) & \text{if } \rho \text{ is test-only} \\
\delta("\varphi^n\), \Pi) \land \delta("[\rho][\rho]^*\varphi^n\), \Pi) & \text{o/w} 
\end{cases}
\]

where \(\delta("\varphi^n\), \epsilon\), i.e., the interpretation of LTL\(_f\) formula in the case the (remaining fragment of the) trace is empty, is defined as follows:

\[
\begin{align*}
\delta("tt^n\), \epsilon) &= true \\
\delta("ff^n\), \epsilon) &= false \hspace{1cm} (\phi \ \text{propositional}) \\
\delta("\phi^n\), \epsilon) &= false \hspace{1cm} (\phi \ \text{propositional}) \\
\delta("\varphi_1 \land \varphi_2^n\), \epsilon) &= \delta("\varphi_1^n\), \epsilon) \land \delta("\varphi_2^n\), \epsilon) \\
\delta("\varphi_1 \lor \varphi_2^n\), \epsilon) &= \delta("\varphi_1^n\), \epsilon) \lor \delta("\varphi_2^n\), \epsilon) \\
\delta("(\phi)\varphi^n\), \epsilon) &= false \hspace{1cm} (\phi \ \text{propositional}) \\
\delta("(\psi?)\varphi^n\), \epsilon) &= \delta("\psi^n\), \epsilon) \land \delta("\varphi^n\), \epsilon) \\
\delta("(\rho_1 + \rho_2)\varphi^n\), \epsilon) &= \delta("(\rho_1)\varphi^n\), \epsilon) \lor \delta("(\rho_2)\varphi^n\), \epsilon) \\
\delta("(\rho_1; \rho_2)\varphi^n\), \epsilon) &= \delta("(\rho_1)\varphi^n\), \epsilon) \land \delta("(\rho_2)\varphi^n\), \epsilon) \\
\delta("(\rho)^*\varphi^n\), \epsilon) &= \delta("\varphi^n\), \epsilon) \\
\delta("[\rho]\psi^n\), \epsilon) &= \delta("\mathrm{nnf}(\neg\psi)^n\), \epsilon) \lor \delta("\varphi^n\), \epsilon) \\
\delta("[\rho_1 + \rho_2]\varphi^n\), \epsilon) &= \delta("[\rho_1]\varphi^n\), \epsilon) \land \delta("[\rho_2]\varphi^n\), \epsilon) \\
\delta("[\rho_1; \rho_2]\varphi^n\), \epsilon) &= \delta("[\rho_1]\varphi^n\), \epsilon) \\
\delta("[\rho]^*\varphi^n\), \epsilon) &= \delta("\varphi^n\), \epsilon)
\end{align*}
\]

Notice also that for \(\phi \ \text{propositional}, \delta("\phi^n\), \Pi) = \delta("(\phi)tt^n\), \Pi)\) and \(\delta("\phi^n\), \epsilon) = \delta("(\phi)tt^n\), \Pi)\), as a consequence of the equivalence \(\phi \equiv (\phi)tt\).

Using the auxiliary function \(\delta\) we can build the NFA \(A_\varphi\) of an LTL\(_f\) formula \(\varphi\) in a forward fashion as described in Figure 1, where: states of \(A_\varphi\) are sets of atoms (recall that each atom is quoted \(\varphi\) subformulas) to be interpreted as a conjunction; the empty conjunction \(\emptyset\) stands for \(true\); \(\Theta\) is either a propositional interpretation
Π over \( P \) or the empty trace \( \epsilon \) (this gives rise to epsilon transition either to true or false) and \( q' \) is a set of quoted subformulas of \( \varphi \) that denotes a minimal interpretation such that \( q' \models \bigwedge_{\psi \in q} \delta(\psi, \emptyset) \). (Note: we do not need to get all \( q \) such that \( q' \models \bigwedge_{\psi \in q} \delta(\psi, \emptyset) \), but only the minimal ones.) Notice that trivially we have \( \langle \emptyset, a, \emptyset \rangle \in \rho \) for every \( a \in \Sigma \).

The algorithm LDLf2NFA terminates in at most exponential number of steps, and generates a set of states \( S \) whose size is at most exponential in the size of \( \varphi \).

**Theorem 1.** Let \( \varphi \) be an LDLf formula and \( A_\varphi \) the NFA constructed as above. Then \( \pi \models \varphi \) iff \( \pi \in L(A_\varphi) \) for every finite trace \( \pi \).

**Proof (sketch).** Given a LDLf formula \( \varphi \), \( \delta \) grounded on the subformulas of \( \varphi \) becomes the transition function of the AFW, with initial state "\( \varphi \)" and no final states, corresponding to \( \varphi \) \(^4\). Then LDLf2NFA essentially transforms the AFW into a NFA.

Notice that above we have assumed to have a special proposition \( \text{last} \in P \). If we want to remove such an assumption, we can easily transform the obtained automaton \( A_\varphi = (2^P, S, \{"\varphi"\}, g, \{\emptyset\}) \) into the new automaton

\[
A'_\varphi = (2^P - \{\text{last}\}, S \cup \{\text{ended}\}, \{"\varphi"\}, g', \{\emptyset, \text{ended}\})
\]

where: \( (q, \Pi', q') \in g' \) iff \( (q, \Pi', q') \in g \) or \( (q, \Pi' \cup \{\text{last}\}, \text{true}) \in g \) and \( q' = \text{ended} \).

It is easy to see that the NFA obtained can be built on-the-fly while checking for nonemptiness, hence we have:

**Theorem 2.** Satisfiability of an LDLf formula can be checked in PSPACE by nonemptiness of \( A_\varphi \) (or \( A'_\varphi \)).

Considering that it is known that satisfiability in LDLf is a PSPACE-complete problem we can conclude that the proposed construction is optimal wrt computational complexity for satisfiability, as well as for validity and logical implication which are linearly reducible to satisfiability in LDLf (see \(^4\) for details).

## 4 Run-time Monitoring

From an high-level perspective, the monitoring problem amounts to observe an evolving system execution and report the violation or satisfaction of properties of interest at the earliest possible time. As the system progresses, its execution trace increases, and at each step the monitor checks whether the trace seen so far conforms to the properties, by considering that the execution can still continue. This evolving aspect has a significant impact on the monitoring output: at each step, indeed, the outcome may have a degree of uncertainty due to the fact that future executions are yet unknown.

Several variant of monitoring semantics have been proposed (see \(^2\) for a survey). In this paper we adopt the semantics in \(^9\), which is basically the finite-trace variant of the RV semantics in \(^2\): Given a LTLf or LDLf formula \( \varphi \), each time the system evolves, the monitor returns one among the following truth values:

- \( [\varphi]_{RV} = \text{temp\_true} \), meaning that the current execution trace temporarily satisfies \( \varphi \), i.e., it is currently compliant with \( \varphi \), but there is a possible system future prosecution which may lead to falsify \( \varphi \);
\[
\begin{align*}
\phi_{RV} &= \text{temp}\_false, \text{ meaning that the current trace temporarily falsify } \phi, \text{ i.e., } \phi \text{ is not current compliant with } \phi, \text{ but there is a possible system future prosecution which may lead to satisfy } \phi; \\
\phi_{RV} &= \text{true}, \text{ meaning that the current trace satisfies } \phi \text{ and it will always do, no matter how it proceeds;} \\
\phi_{RV} &= \text{false}, \text{ meaning that the current trace falsifies } \phi \text{ and it will always do, no matter how it proceeds.}
\end{align*}
\]

The first two conditions are unstable because they may change into any other value as the system progresses. This reflects the general unpredictability of system possible executions. Conversely, the other two truth values are stable since, once outputted, they will not change anymore. Observe that a stable truth value can be reached in two different situations: (i) when the system execution terminates; (ii) when the formula that is being monitored can be fully evaluated by observing a partial trace only. The first case is indeed trivial, as when the execution ends, there are no possible future evolutions and hence it is enough to evaluate the finite (and now complete) trace seen so far according to the \text{LDL}_f semantics. In the second case, instead, it is irrelevant whether the system continues its execution or not, since some \text{LDL}_f properties, such as eventualities or safety properties, can be fully evaluated as soon as something happens, e.g., when the eventuality is verified or the safety requirement is violated. Notice also that when a stable value is outputted, the monitoring analysis can be stopped.

From a more theoretical viewpoint, given an \text{LDL}_f property \( \phi \), the monitor looks at the trace seen so far, assesses if it is a prefix of a complete trace not yet completed, and categorizes it according to its potential for satisfying or violating \( \phi \) in the future. We call a prefix \text{possibly good} for an \text{LDL}_f formula \( \phi \) if there exists an extension of it which satisfies \( \phi \). More precisely, given an \text{LDL}_f formula \( \phi \), we define the set of \text{possibly good} prefixes for \( \text{LDL}_f(\phi) \) as the set
\[
\mathcal{L}_{\text{poss, good}}(\phi) = \{ \pi | \exists \pi'. \pi \pi' \in \mathcal{L}(\phi) \}.
\]

Prefixes for which every possible extension satisfies \( \phi \) are instead called \text{necessarily good}. More precisely, given an \text{LDL}_f formula \( \phi \), we define the set of \text{necessarily good} prefixes for \( \text{LDL}_f(\phi) \) as the set
\[
\mathcal{L}_{\text{nec, good}}(\phi) = \{ \pi | \forall \pi'. \pi \pi' \in \mathcal{L}(\phi) \}.
\]

The set of \text{necessarily bad} prefixes \( \mathcal{L}_{\text{nec, bad}}(\phi) \) can be defined analogously as
\[
\mathcal{L}_{\text{nec, bad}}(\phi) = \{ \pi | \forall \pi'. \pi \pi' \not\in \mathcal{L}(\phi) \}.
\]

Observe that the necessarily bad prefixes for \( \phi \) are the necessarily good prefixes for \( \neg \phi \), i.e., \( \mathcal{L}_{\text{nec, bad}}(\phi) = \mathcal{L}_{\text{nec, good}}(\neg \phi) \).

Using this language theoretic notions, we can provide a precise characterization of the semantics four standard monitoring evaluation functions [9].

**Proposition 1.** Let \( \phi \) be an \text{LDL}_f formula and \( \pi \) a trace. Then:
\[
\begin{align*}
\pi &|= \phi_{RV} = \text{temp}\_true \iff \pi \in \mathcal{L}(\phi) \setminus \mathcal{L}_{\text{nec, good}}(\phi); \\
\pi &|= \phi_{RV} = \text{temp}\_false \iff \pi \in \mathcal{L}(\neg \phi) \setminus \mathcal{L}_{\text{nec, bad}}(\phi); \\
\pi &|= \phi_{RV} = \text{true} \iff \pi \in \mathcal{L}_{\text{nec, good}}(\phi); \\
\pi &|= \phi_{RV} = \text{false} \iff \pi \in \mathcal{L}_{\text{nec, bad}}(\phi).
\end{align*}
\]
Proof (sketch). Immediate from the definitions in \cite{9} and the language theoretic definitions above. $\Box$

We close this section by exploiting the language theoretic notions to better understand the relationships between the various kinds of prefixes. We start by observing that, the set of all finite words over the alphabet $2^P$ is the union of the language of $\varphi$ and its complement $L(\varphi) \cup L(\neg \varphi) = (2^P)^*$. Also, any language and its complement are disjoint $L(\varphi) \cap L(\neg \varphi) = \emptyset$.

Since from the definition of possibly good prefixes we have $L(\varphi) \subseteq L_{\text{poss\_good}}(\varphi)$ and $L(\neg \varphi) \subseteq L_{\text{poss\_good}}(\neg \varphi)$, we also have that $L_{\text{poss\_good}}(\varphi) \cup L_{\text{poss\_good}}(\neg \varphi) = (2^P)^*$. Also from the definition it is easy to see that $L_{\text{poss\_good}}(\varphi) \cap L_{\text{poss\_good}}(\neg \varphi) = \{ \pi \mid \exists \pi'. \pi \pi' \in L(\varphi) \wedge \exists \pi''. \pi \pi'' \in L(\neg \varphi) \}$ meaning that the set of possibly good prefixes for $\varphi$ and the set of possibly good prefixes for $\neg \varphi$ do intersect, and in such an intersection are paths that can be extended to satisfy $\varphi$ but can also be extended to satisfy $\neg \varphi$. It is also easy to see that $L(\varphi) = L_{\text{poss\_good}}(\varphi) \setminus L(\neg \varphi)$.

Turning to necessarily good prefixes and necessarily bad prefixes, it is easy to see that $L_{\text{necc\_good}}(\varphi) = L_{\text{possgood}}(\varphi) \setminus L_{\text{possgood}}(\neg \varphi)$, that $L_{\text{necc\_bad}}(\varphi) = L_{\text{possgood}}(\neg \varphi) \setminus L_{\text{possgood}}(\varphi)$, and also that $\subseteq L(\varphi)$ and $L_{\text{necc\_good}}(\varphi) \not\subseteq L(\neg \varphi)$.

Interestingly, necessarily good, necessarily bad, possibly good prefixes partition all finite traces. Namely

\begin{proposition}
The set of all traces $(2^P)^*$ can be partitioned into

\[ L_{\text{necc\_good}}(\varphi) \quad L_{\text{possgood}}(\varphi) \cap L_{\text{possgood}}(\neg \varphi) \quad L_{\text{necc\_bad}}(\varphi) \]

such that $L_{\text{necc\_good}}(\varphi) \cup (L_{\text{possgood}}(\varphi) \cap L_{\text{possgood}}(\neg \varphi)) \cup L_{\text{necc\_bad}}(\varphi) = (2^P)^*$

$L_{\text{necc\_good}}(\varphi) \cap (L_{\text{possgood}}(\varphi) \cap L_{\text{possgood}}(\neg \varphi)) \cap L_{\text{necc\_bad}}(\varphi) = \emptyset$.

\end{proposition}

\begin{proof}[Proof (sketch)]
Follows from the definitions of the necessarily good, necessarily bad, possibly good prefixes of $L(\varphi)$ and $L(\neg \varphi)$.
\end{proof}

5 Runtime Monitors in \textit{LDL}_f

As discussed in the previous section the core issue in monitoring is prefix recognition. \textit{LTL}_f is not expressive enough to talk about prefixes of its own formulas. Roughly speaking, given a \textit{LTL}_f formula, the language of its possibly good prefixes for cannot be in general described as an \textit{LTL}_f formula. For such a reason, building a monitor usually requires direct manipulation of the automaton for $\varphi$.

\textit{LDL}_f instead can capture any nondeterministic automata as a formula, and it has the capability of expressing properties on prefixes. We can exploit such an extra expressivity to capture the monitoring condition in a direct and elegant way. We start by showing how to construct formulas representing (the language of) prefixes of other formulas, and then we prove how use them for the monitoring problem.

More precisely, given an \textit{LDL}_f formula $\varphi$, it is possible to express the language $L_{\text{possgood}}(\varphi)$ with an \textit{LDL}_f formula $\varphi'$. Such a formula is obtained in two steps.

\begin{lemma}
Given a \textit{LDL}_f formula $\varphi$, there exists a regular expression $\text{pref}_\varphi$ such that $L(\text{pref}_\varphi) = L_{\text{possgood}}(\varphi)$.
\end{lemma}
Proof (sketch). The proof is constructive. We can build the NFA \( A \) for \( \varphi \) following the procedure in [4]. We then set as final all states of \( A \) from which there exists a non-zero length path to a final state. This new finite state machine \( A_{\text{poss,good}}(\varphi) \) is such that \( L(A_{\text{poss,good}}(\varphi)) = L_{\text{poss,good}}(\varphi) \). Since NFA are exactly as expressive as regular expressions, we can translate \( A_{\text{poss,good}}(\varphi) \) to a regular expression \( \text{pref}_\varphi \).

Given that LDL is as expressive as regular expression (cf. [4]), we can translate \( \text{pref}_\varphi \) into an equivalent LDL formula, as the following states.

\[ \pi \in L_{\text{poss,good}}(\varphi) \text{ iff } \pi \models \langle \text{pref}_\varphi \rangle_{\text{end}} \]
\[ \pi \in L_{\text{nec,good}}(\varphi) \text{ iff } \pi \models \langle \text{pref}_\varphi \rangle_{\text{end}} \land \lnot \langle \text{pref}_{\lnot \varphi} \rangle_{\text{end}} \]

Proof (sketch). Any regular expression \( \rho \), and hence any regular language, can be captured in LDL as \( \langle \rho \rangle_{\text{end}} \). Hence the language \( L_{\text{nec,good}}(\varphi) \) can be captured by \( \langle \text{pref}_\varphi \rangle_{\text{end}} \) and the language \( L_{\text{nec,good}}(\varphi) \) which is equivalent \( L_{\text{poss,good}}(\varphi) \setminus L_{\text{poss,good}}(\lnot \varphi) \) can be captured by captured by \( \langle \text{pref}_\varphi \rangle_{\text{end}} \land \lnot \langle \text{pref}_{\lnot \varphi} \rangle_{\text{end}} \).

In other words, given a LDL formula \( \varphi \), formula \( \varphi' = \langle \text{pref}_\varphi \rangle_{\text{end}} \) is a LDL formula such that \( L(\varphi') = L_{\text{poss,good}}(\varphi) \). Similarly for \( L_{\text{nec,good}}(\varphi) \).

Exploiting this result, and the results in Proposition[1] we reduce RV monitoring to the standard evaluation of LDL formulas over a (partial) trace. Formally:

\[ \pi \mid= [\varphi]_{\text{RV}} = \text{temp,true} \text{ iff } \pi \models \varphi \land \langle \text{pref}_\varphi \rangle_{\text{end}}; \]
\[ \pi \mid= [\varphi]_{\text{RV}} = \text{temp,false} \text{ iff } \pi \models \lnot \varphi \land \langle \text{pref}_\varphi \rangle_{\text{end}}; \]
\[ \pi \mid= [\varphi]_{\text{RV}} = \text{true} \text{ iff } \langle \text{pref}_\varphi \rangle_{\text{end}} \land \lnot \langle \text{pref}_{\lnot \varphi} \rangle_{\text{end}}; \]
\[ \pi \mid= [\varphi]_{\text{RV}} = \text{false} \text{ iff } \langle \text{pref}_{\lnot \varphi} \rangle_{\text{end}} \land \lnot \langle \text{pref}_\varphi \rangle_{\text{end}}. \]


6 Monitoring Declare Constraints and Metaconstraints

We now ground our monitoring approach to the case of DECLARE monitoring. DECLARE is a language and framework for the declarative, constraint-based modelling of processes and services. A thorough treatment of constraint-based processes can be found in [16][12]. As a modelling language, DECLARE takes a complementary approach to that of classical, imperative process modeling, in which all allowed control-flows among tasks must be explicitly represented, and every other execution trace is implicitly considered as forbidden. Instead of this procedural and “closed” approach, DECLARE has a declarative, “open” flavor: the agents responsible for the process execution can freely choose how to perform the involved tasks, provided that the resulting execution trace complies with the modeled business constraints. This is the reason why, alongside traditional control-flow constraints such as sequence (called in DECLARE chain succession), DECLARE supports a plethora of peculiar constraints that do not impose

\footnote{http://www.win.tue.nl/declare/}
specific temporal orderings, or that explicitly account with negative information, i.e., prohibition of task execution.

Given a set $P$ of tasks, a DECLARE model is a set $C$ of $LTL_f$ (and hence $LDL_f$) constraints over $P$, used to restrict the allowed execution traces. Among all possible $LTL_f$ constraints, some specific patterns have been singled out as particularly meaningful for expressing DECLARE processes, taking inspiration from [5]. Such patterns are grouped into four families: (i) existence (unary) constraints, stating that the target task must/cannot be executed (a certain amount of times); (ii) choice (binary) constraints, modeling choice of execution; (iii) relation (binary) constraints, modeling that whenever the source task is executed, then the target task must also be executed (possibly with additional requirements); (iv) negation (binary) constraints, modeling that whenever the source task is executed, then the target task is prohibited (possibly with additional restrictions). Table 1 reports some of these patterns.

Example 1. Consider a fragment of a purchase order process, where we consider three key business constraints. First, an order can be closed at most once. In DECLARE, this can be tackled with a absence constraint, visually and formally represented as:

$$\varphi_{close} = \neg \diamondclose order$$

Second, an order can be canceled only until it is closed. This can be captured by a negation succession constraint, which states that after the order is closed, it cannot be canceled anymore:

$$\varphi_{canc} = \Box(close order \rightarrow \neg \diamond-cancel order)$$

Finally, after the order is closed, it becomes possible to do supplementary payments, for various reasons (e.g., to speed up the delivery of the order).

$$\varphi_{pay} = (\neg-pay suppl \cup close order) \lor \neg \diamond-close order$$

Beside modeling and enactment of constraint-based processes, previous works have also focused on runtime verification of DECLARE models. A family of DECLARE monitoring approaches rely on the original $LTL_f$ formalization of DECLARE, and employ corresponding automata-based techniques to track running process instances and check whether they satisfy the modeled constraints or not [9,10]. Such techniques have been in particular used for:

- Monitoring single DECLARE constraints so as to provide a fine-grained feedback; this is done by adopting the RV semantics for $LTL_f$, and tracking the evolution each constraint through the four RV truth values.
- Monitoring the global DECLARE model by considering all its constraints together (i.e., constructing a DFA for the conjunction of all constraints); this is important for computing the early detection of violations, i.e., violations that cannot be explicitly found in the execution trace collected so far, but that cannot be avoided in the future.

We now discuss how $LDL_f$ can be adopted for monitoring DECLARE constraints, with a twofold advantage. First, as shown in Section 5, $LDL_f$ is able to encode the RV semantics directly into the logic, without the need of introducing ad-hoc modifications.
Monitoring Declare Constraints with $\text{LDL}_f$. Since $\text{LDL}_f$ includes $\text{LTL}_f$, DECLARE constraints can be directly encoded in $\text{LDL}_f$ using their standard formalization [17,15]. Thanks to the translation into NFAs discussed in Section 3 (and, if needed, their determinization into corresponding DFAs), the obtained automaton can then be used to check whether a (partial) finite trace satisfies this constraint or not. This is not very effective, as the approach does not support the detection of fine-grained truth values as those of RV. By relying on Theorem 4, however, we can reuse the same technique, this time supporting all RV. In fact, by formalizing the good prefixes of each DECLARE pattern, we can immediately construct the four $\text{LDL}_f$ formulas that embed the different RV truth values, and check the current trace over each of the corresponding automata. Table 1 reports the good prefix characterization of some of the DECLARE patterns; it can be seamlessly extended to all other patterns as well.

Example 2. Let us consider the absence 2 constraint $\varphi_{\text{close}}$ in Example 1. Following Table 1, its good prefix characterization is $\text{pref}_{\varphi_{\text{close}}} = \otimes^+ (\otimes^+; \text{close order}; \otimes^+)$, where $\otimes$ is a shortcut for all the tasks involved in the purchase order process but close order. This can be used to construct the four formulas mentioned in Theorem 4, which in turn provide the basis to produce, e.g., the following result:

<table>
<thead>
<tr>
<th>Time</th>
<th>start</th>
<th>“do close order”</th>
<th>“do pay suppl.”</th>
<th>“close order”</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>close order</td>
<td>temp_true</td>
<td>false</td>
<td></td>
</tr>
</tbody>
</table>

Observe that this baseline approach can be extended along a number of directions. For example, as shown in Table 1, the majority of DECLARE patterns does not cover all the four RV truth values. This is the case, e.g., for absence 2, which can never be evaluated to be true (since it is always possible to continue the execution so as to perform a twice), nor to $\text{temp}_\text{false}$ (the only way of violating the constraint is to perform a twice, and in this case it is not possible to “repair” to the violation anymore). This information can be used to restrict the generation of the automata only to those cases that are relevant to the constraint. Furthermore, it is possible to reconstruct exactly the approach in [9], where every state in the DFAs corresponding to the constraints to be monitored, is enriched with a “color” accounting for one of the four RV truth values. To do so, we have simply to combine the four DFAs generated for each constraint. This is possible because such DFAs are generated from formulas built on top of the good prefix characterization of the original formula, and hence they all produce the same automaton, but with different final states. In fact, this observation provides a formal justification to the correctness of the approach in [9].

Metaconstraints. Thanks to the ability of $\text{LDL}_f$ to directly encode into the logic DECLARE constraints but also their RV monitoring states, we can formalize metaconstraints that relate the RV truth values of different constraints. Intuitively, such metaconstraints allow one to capture that we become interested in monitoring some constraint in the corresponding standard logical services. Second, beside being able to reconstruct all the aforementioned monitoring techniques, our approach also provides a declarative, well-founded basis for monitoring metaconstraints, i.e., constraints that involve both the execution of tasks and the monitoring outcome obtained by checking other constraints.
<table>
<thead>
<tr>
<th>NAME</th>
<th>NOTATION</th>
<th>pref</th>
<th>POSSIBLE RV STATES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existence</td>
<td>$a$</td>
<td>$(a + o)^*$</td>
<td>temp_false, true</td>
</tr>
<tr>
<td>Absence</td>
<td>$a$</td>
<td>$a^* + (a^<em>; a^</em>)^*$</td>
<td>temp_false, true</td>
</tr>
<tr>
<td>Choice</td>
<td>$a$</td>
<td>$(a + o)$</td>
<td>temp_false, true</td>
</tr>
<tr>
<td>Exclusive Choice</td>
<td>$a$</td>
<td>$(a + o)^* + (b + o)^*$</td>
<td>temp_false, true, true</td>
</tr>
<tr>
<td>Resp. existence</td>
<td>$a$</td>
<td>$(a + b + o)^*$</td>
<td>temp_true, temp_false, true</td>
</tr>
<tr>
<td>Coexistence</td>
<td>$a$</td>
<td>$(a + b + o)^*$</td>
<td>temp_true, temp_false, true</td>
</tr>
<tr>
<td>Response</td>
<td>$a$</td>
<td>$(a + b + o)^*$</td>
<td>temp_true, temp_false, true</td>
</tr>
<tr>
<td>Precedence</td>
<td>$a$</td>
<td>$a^<em>: (a + b + o)^</em>^*$</td>
<td>temp_true, true, false</td>
</tr>
<tr>
<td>Succession</td>
<td>$a$</td>
<td>$a^<em>: (a + b + o)^</em>^*$</td>
<td>temp_true, true, false</td>
</tr>
<tr>
<td>Not Coexistence</td>
<td>$a$</td>
<td>$(a + o)^* + (b + o)^*$</td>
<td>temp_true, false</td>
</tr>
<tr>
<td>Neg. Succession</td>
<td>$a$</td>
<td>$(b + o)^<em>: (a + o)^</em>$</td>
<td>temp_true, false</td>
</tr>
</tbody>
</table>

Table 1. Some DECLARE constraints, together with their prefix characterization, minimal bad prefix characterization, and possible RV states; for each constraint, $o$ is a shortcut for “other tasks”, i.e., tasks not involved in the constraint itself.

only when other constraints are evaluated to be in a certain RV truth value. This, in turn, provides the basis to declaratively capture two classes of properties that are of central importance in the context of runtime verification:

- **Compensation constraints**, that is, constraints that should be enforced by the agents executing the process in the case other constraints are violated, i.e., are evaluated to be false. Previous works have been tackled this issue through ad-hoc techniques, with no declarative counterpart \[9,10\].

- **Recovery mechanisms resembling contrary-to-duty obligations** in legal reasoning \[19\], i.e., obligations that are put in place only when other obligations are not met.

Technically, a generic form for metaconstraints is the pattern $\Phi_{pre} \rightarrow \Psi_{exp}$, where:

- $\Phi_{pre}$ is a boolean formula, whose atoms are membership assertions of the involved constraints to the RV truth values;
- $\Psi_{exp}$ is a boolean formula whose atoms are the constraints to be enforced when $\Phi_{pre}$ evaluates to true.

This pattern can be used, for example, to state that whenever constraints $c_1$ and $c_2$ are permanently violated, then either constraint $c_3$ or $c_4$ have to be enforced. Observe that the metaconstraint so constructed is a standard $\text{LDL}_f$ formula. Hence, we can reapply Theorem 4 to it, getting four $\text{LDL}_f$ formulas that can be used to track the evolution of the metaconstraint among the four RV values.

**Example 3.** Consider the DECLARE constraints of Example 1. We want to enhance it with a compensation constraint stating that whenever $\Phi_{canc}$ is violated (i.e., the order is canceled after it has been closed), then a supplement payment must be issued. This can be easily captured in $\text{LDL}_f$ as follows. First of all, we model the compensation constraint, which corresponds, in this case, to a standard existence constraint over the
pay supplement task. Let $\varphi_{dopay}$ denote the LTL$_f$ formalization of such a compensation constraint. Second, we capture the intended compensation behavior by using the following LDL$_f$ metaconstraint:

$$\{[\varphi_{canc}]_{RV} = false\} \rightarrow \varphi_{dopay}$$

which, leveraging Theorem 3, corresponds to the standard LDL$_f$ formula:

$$(\langle pref \neg \varphi_{canc} \rangle_{end} \wedge \neg (pref \varphi_{canc})_{end}) \rightarrow \varphi_{dopay}$$

A limitation of this form of metaconstraint is that the right-hand part $\Psi_{exp}$ is monitored from the beginning of the trace. This is acceptable in many cases. E.g., in Example 3 it is ok if the user already paid a supplement before the order cancelation caused constraint $\varphi_{canc}$ to be violated. In other situations, however, this is not satisfactory, because we would like to enforce the compensating behavior only after $\Phi_{pre}$ evaluates to true, e.g., after the violation of a given constraint has been detected. In general, we can extend the aforementioned metaconstraint pattern as follows: $\Phi_{pre} \rightarrow [\rho]\Psi_{exp}$, where $\rho$ is a regular expression denoting the paths after which $\Psi_{exp}$ is expected to be enforced.

By constructing $\rho$ as the regular expression accounting for the paths that make $\Phi_{pre}$ true, we can then exploit this improved metaconstraint to express that $\Psi_{exp}$ is expected to become true after all prefixes of the current trace that made $\Phi_{pre}$ true.

**Example 4.** We modify the compensation constraint of Example 3, so as to reflect that when a closed order is canceled (i.e., $\varphi_{canc}$ is violated), then a supplement must be paid afterwards. This is captured by the following metaconstraint:

$$\{[\varphi_{canc}]_{RV} = false\} \rightarrow [re([\varphi_{canc}]_{RV} = false)]\varphi_{dopay}$$

where $re([\varphi_{canc}] = false)$ denotes the regular expression for the language $L([\varphi_{canc}] = false) = L([pref \neg \varphi_{canc}]_{end} \wedge \neg (pref \varphi_{canc})_{end})$. This regular expression describes all paths containing a violation for constraint $\varphi_{canc}$.

## 7 Implementation

The entire approach has been implemented as an operational decision support (OS) provider for the PROM 6 process mining framework$^5$. PROM 6 provides a generic OS environment$^{[22]}$ that supports the interaction between an external workflow management systems at runtime (producing events) and ProM. In particular, it provides an OS service that receives a stream of events from the external world, updates and orchestrates the registered OS providers implementing different types of online analysis to be applied on the stream, and reports the produced results back to the external world.

At the back-end of the plug-in, there is a software module specifically dedicated to the construction and manipulation of NFAs from LDL$_f$ formulas, concretely implementing the technique presented in Section 5. To manipulate regular expressions and automata, we used the fast, well-known library dk.brics.automaton$^{[11]}$.

Figure 2 shows a graphical representation of the evolution of constraints described in Example 1 of Section 6 when monitored using our operational support provider. In the LDL$_f$ formulas, the literals $f$, $g$ and $d$ respectively stand for tasks close order, cancel order, and pay supplement.

$^5$[http://www.promtools.org/prom6/](http://www.promtools.org/prom6/)
8 Conclusion

We can see the approach proposed in this paper, as an extension of the declarative process specification approach, at the basis of DECLARE, to (i) more powerful specification logics (LDL\textsubscript{f}, i.e., Monadic Second Order logics over finite traces, instead of LTL\textsubscript{f}, i.e., First Order logic logics over finite traces), (ii) to monitoring constraints. Notably, this declarative approach to monitoring supports seamlessly monitoring metaconstraints, i.e., constraints that do not only predicate about the dynamics of task executions, but also about the truth values of other constraints. We have grounded this approach on DECLARE itself, showing how to declaratively specify compensation constraints.

The next step will be to incorporate recovery mechanisms into the approach, in particular providing a formal underpinning to the ad-hoc recovery mechanisms studied in \cite{9}. Furthermore, we intend to extend our approach to data-aware business constraints \cite{11,13,14}, mixing temporal operators with first-order queries over the data attached to the monitored events. This setting has been studied using the Event Calculus \cite{13,14}, also considering some specific forms of compensation in DECLARE \cite{3}. However, the resulting approach can only query the partial trace accumulated so far, and not reason upon its possible future continuations, as automata-based techniques are able to do. To extend the approach presented here to the case of data-aware business constraints, we will build on recent, interesting decidability results for the static verification of data-aware business processes against sophisticated variants of first-order temporal logics \cite{11}.

Acknowledgments. This research has been partially supported by the EU IP project Optique: Scalable End-user Access to Big Data, grant agreement n. FP7-318338, and by the Sapienza Award 2013 “SPIRITLETS: Spiritlet-based smart spaces”.

Fig. 2. Screenshot of our operational support provider’s output.
References