Towards a Logical Framework for Reasoning about Risk

Matteo Cristani, Erisa Karafili, Luca Viganò

Towards a Logical Framework for Reasoning about Risk

Matteo Cristani, Erisa Karafili, and Luca Viganò

Dipartimento di Informatica, Università degli Studi di Verona, Italy

Abstract. Evaluating the effectiveness of the security measures undertaken to protect a distributed system (e.g., protecting privacy of data in a network or in an information system) is a difficult task that, among other things, requires a risk assessment. We introduce a logical framework that allows one to reason about risk by means of operators that formalize causes, effects, preconditions, prevention and mitigation of events that may occur in the system. This is work in progress and we describe a number of interesting variants that could be considered.

1 Introduction

Evaluating the effectiveness of the security measures undertaken to protect a system, such as protecting privacy of data in a network or in a distributed system, is a difficult task that, among other things, requires one to carry out a *risk assessment*. To illustrate this, let us consider a real-life scenario in which privacy problems arise and security measures should be evaluated in terms of their effectiveness for risk reduction: the information system of a hospital should manage the process of hospitalizing patients coming directly from family doctors or from other wards or hospitals, where rules of access to patient data are employed to guarantee a satisfactory level of privacy within the system. We may thus look at these rules as risk reduction measures, and represent the relationships among different situations (or processes) in which the measures are taken or not.

For concreteness, let us consider a case study taken from the project *eFA* for personal health information management in hospitals [1]. All the records of a patient are stored in a *logical file* that contains a set of the *patient's medical records*, which are of three kinds: *administrative records* (AR) that contain the personal data of the patient, *normal records* (NR) that contain all the information that the doctors and nurses that attend to the patient should know, and *restricted records* (RR) that contain particularly sensitive information (like a record of a treatment for depression or some infectious disease).

To ensure the patient's health and simultaneously protect her privacy, the system must thus control the access to this information and enforce measures for risk reduction. The decisions about which measures of risk reduction should be adopted are based on relations among the events involved in the analyzed process; in particular, we adopt here the commonly accepted view point that assessing risk consists in managing causes and effects of a family of specific events

(often also called *threats*). To enable such a decision procedure, in this paper we introduce a logical framework that allows one to represent the flow of time, and thereby capture the *temporal relationships* between events, and to formalize *causes*, *effects* and *preconditions* of events. Moreover, we also formalize event *prevention* (diminishing the number of occurrences of a threat) and *mitigation* (diminishing the impact of the effects of a threat).

This is work in progress and we describe a number of interesting variants that could be considered. To our knowledge, this is the first attempt at a general logical framework that accommodates causal relationships between events, as well as their prevention and mitigation. In contrast to quantitative approaches, where risk is assessed in terms of probabilities, we adopt here a *qualitative* approach in which we reason symbolically about the occurrence of events and associated risks (leaving an extension with probabilities for future work).

Risk assessment and risk reduction have been considered quite often in information security, environmental security (in the engineering context), portfolio management, and medicine. For instance, the algebraic framework RT^R given in [5] extends RT_0 [12] to provide a formal approach to risk evaluation in distributed authorization. Similarly, [2] proposes a tool for assessing risks related to policy overwrite and privacy leak; the approach is quite practical but it considers, as further work, the interesting possibility of extending the calculus of overwrite permissions "without expert intervention". A more mature investigation of risk assessment is given in [15], where, in particular, the authors consider the design of suitable experiments and analyze the effects of security enhancements within an organization, recommending methods for deciding the correct mixture of security measures to be chosen automatically. Although quite different in nature, all these investigations aim at tackling the problem of the automation of the selection process of security measures to reduce risks. The approach we propose will eventually lead to a system where such automation will be possible.

In [10, 11], Lewis developed an approach to the representation of causes based on two distinct concepts: *causal dependency* and *counterfactuals*. Fundamentally, Lewis's theory describes two different causal relationships: the *precondition relation* (as developed in many temporal theories of AI, e.g., [14]) and *direct causation* (see, e.g., [4, 16]). As further illustrated in [9, 17], these notions can be interpreted by means of systems of possible worlds as in the Kripke semantics for modal logics. We follow a similar semantic approach to base our framework on *labeled deduction* [6, 8, 18]. In Section 2, we give syntax and semantics of our language and a set of tableau rules, discussing the different variants one may consider. In Section 3, we show, proof-of-concept, our framework at work on the case study. In Section 4, we conclude and discuss future work.

2 The Framework

2.1 Syntax

We consider a language structured in three layers. Given a set Π of *propositional variables* $p, q, r, \ldots$, the set of *well-formed formulas* (or, simply, *formulas*) ϕ of

the first layer is defined by the grammar

$$\phi ::= \ p \mid \neg\phi \mid \phi \rightarrow \phi.$$

Other connectives (e.g., $\wedge$ and $\vee$) can be defined as usual. We refer to formulas ϕ also as *events* or *basic formulas* to stress that they are used by the formulas of the other two layers (we write E to denote the set of such formulas).

Formulas σ of the second layer are built from events, the two standard modal operators $\Box$ and $\Diamond$[1], and the two causal operators $\mathcal{C}$ and $\mathcal{P}$:

$$\sigma ::= \phi \mid \Box\phi \mid \Diamond\phi \mid \phi\,\mathcal{C}\,\phi \mid \phi\,\mathcal{P}\,\phi.$$

$\phi_1\,\mathcal{C}\,\phi_2$ and $\phi_1\,\mathcal{P}\,\phi_2$ are called *causal formulas*: ϕ_1 is the *cause event* and ϕ_2 is the *effect event*. $\mathcal{C}$ denotes *causation*: $\phi_1\,\mathcal{C}\,\phi_2$ denotes, intuitively, that if the cause ϕ_1 (e.g., winning a presidential election) occurs then the effect ϕ_2 (e.g., becoming the nation's president) will occur in the future. $\mathcal{P}$ denotes *precondition*: $\phi_1\,\mathcal{P}\,\phi_2$ denotes, intuitively, that if the effect ϕ_2 (e.g., suffering from a viral disease) occurs, then the cause ϕ_1 (e.g., being infected by a virus) occurred in its past.

The third layer extends the first one (so, the second and third layer are actually "side by side") by introducing formulas τ for *prevention* (or *block*) and *mitigation* of events:

$$\tau ::= \phi \mid \phi\,\mathcal{B}\,\phi \mid \phi\,\mathcal{M}\,\phi.$$

The prevention/block $\phi_1\,\mathcal{B}\,\phi_2$ denotes that the event ϕ_1 (e.g., a 100% effective prophylactic vaccine—if such a thing existed) prevents the event ϕ_2 (e.g., the illness being vaccinated against), i.e., if we have ϕ_1 then for sure we will not have ϕ_2 in the future. The mitigation $\phi_1\,\mathcal{M}\,\phi_2$ denotes that the event ϕ_1 (e.g., taking a medicine against the flu) prevents the effects of the event ϕ_2 (e.g., fever), i.e., if we have ϕ_1 then for sure we will not have in the future the effects of ϕ_2.

Note that some of the these operators can be defined in terms of standard modal operators, e.g., $\phi_1\,\mathcal{C}\,\phi_2$ could be defined as $[\Box](\phi_1 \rightarrow \Diamond\phi_2)$ and $\phi_1\,\mathcal{B}\,\phi_2$ as $[\Box](\phi_1 \rightarrow \Box\neg\phi_2)$, where $[\Box]$ denotes that this $\Box$ is optional depending on how one defines the semantics. Similarly, $\phi_1\,\mathcal{P}\,\phi_2$ could be defined as $[\Box](\phi_2 \rightarrow \blacklozenge\phi_1)$, where $\blacklozenge$ is the symbol for "$\Diamond$ in the past", which we could easily add, and $\phi_1\,\mathcal{M}\,\phi_2$ could be defined as $[\Box](\phi_1 \wedge [\Box]\forall\phi_3.\,(\phi_2\,\mathcal{C}\,\phi_3 \rightarrow \Box\neg\phi_3))$, which highlights the second-order nature of the mitigation operator. We, however, prefer to keep these operators as primitive operators in order to stress their relevance and, most importantly, to consider a number of interesting variants (e.g., there are indeed various kinds of vaccines, depending on their effectiveness or on their effects). In fact, we believe that the framework that we introduce here can be quite easily extended to encompass the several interesting variants of cause, precondition, block and mitigation that could be considered. For instance, in the following

[1] Note that we have here a kind of positive modal logic, in which $\Box$ and $\Diamond$ are not duals (since we have no negation in front of formulas of the second layer). This is not a problem, as positive modal logics have been well studied (see, e.g., [7, 18], where different accessibility relations are considered for $\Box$ and $\Diamond$), but of course a full modal logic (in which $\Diamond A = \neg\Box A$) could be considered as well.

we will actually consider a more "refined" variant of block where the event ϕ_1 implies that the number of occurrences of the event ϕ_2 decreases with respect to the case when ϕ_1 did not hold. These choices depend on what exactly one aims to capture and thus we will now define the semantics of these formulas (actually, of their extension to labeled formulas) discussing the different options and variants that one could consider. We will also define tableaux rules for the different operators to formalize reasoning in our framework.

2.2 Time flow, traces and worlds

In the systems that our framework allows us to model, we consider an underlying *time flow* $(T, <)$ where T is a non-empty set of *time instants* and $<$ is a binary relation in T that is irreflexive, transitive, dense (for all $i, j \in T$ such that $i < j$ there exists a $k \in T$ such that $i < k < j$) and linear ($i < j$ or $j < i$ for any two distinct $i, j \in T$). On top of this time flow, we may have more than one course of events, depending on the future that is in fact going to occur. We will thus define a Kripke-style model $\mathfrak{M}$ comprising of traces of possible worlds in which our events occur. Let us consider a non-empty set W of *worlds*, and the binary *adjacence* relation $\lhd \subseteq W \times W$ such that for every $w_i \in W$ there is a $w_j \in W$ for which $w_i \lhd w_j$. A *trace* is then a (possibly infinite) sequence of worlds $\vartheta = w_1 w_2 \cdots w_m w_n \cdots$ such that $w_m \lhd w_n$ for every two adjacent worlds $w_m, w_n \in \vartheta$. The traces of our framework are *discrete*: if two worlds are adjacent then there is no other world between them. We write Θ for the set of all possible traces of the system.

To go back and forth from time instants and worlds, we define the following mappings, which are illustrated in Fig. 1:

- $\mathsf{itw} : (T \times \Theta) \to W$ maps a time instant i and a trace ϑ to a world of that trace $\mathsf{itw}(i, \vartheta) = w \in \vartheta$. To make this mapping possible, we assume that a world may actually span different time instants until a new event occurs that makes us move to the following world.
- $\mathsf{wi} : W \to \wp(T)$ maps a world to a set of time instants.
- $\mathsf{iw} : T \to \wp(W)$ maps an instant of time i, to a set of subsets of W, $\mathsf{iw}(i) = \bigcup_{\vartheta \in \Theta} \mathsf{itw}(i, \vartheta)$. This captures all the worlds in the different traces that occur in the same instants of time.
- $\mathsf{wt} : W \to \wp(\Theta)$ maps a world to its corresponding subset of traces. This expresses that a world may occur in two (or more) traces, e.g., when they intersect or when a trace at some point diverges in two traces.

Given a trace $\vartheta = w_1 w_2 \cdots w_k w_l w_m \cdots$ and a time instant $i \geq 0$ such that $w_l = \mathsf{itw}(i, \vartheta)$, we can define the following *prefix* (sub-)traces $\vartheta|_{\leq i} = \vartheta|_{\leq w_l} = w_1 w_2 \cdots w_k w_l$ and $\vartheta|_{<i} = \vartheta|_{<w_l} = w_1 w_2 \cdots w_k$, and the following *suffix* (sub-)traces $\vartheta|_{\geq i} = \vartheta|_{\geq w_l} = w_l w_m \cdots$ and $\vartheta|_{>i} = \vartheta|_{>w_l} = w_m \cdots$.

2.3 Labeled formulas

To be able to reason in a fine-grained way about the formulas holding in the structures providing our models, we base our framework on *labeled deduction* [6,

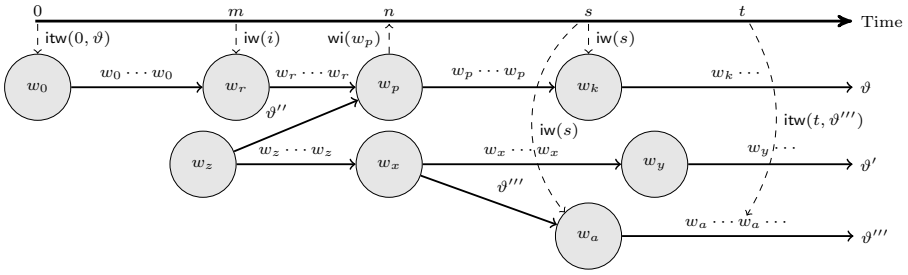

Fig. 1. Time flow, traces of worlds and mappings.

8, 18]. We thus extend the language with a set $\mathcal{L}$ of labels and introduce the notions of labeled formula and relational formula. Our labels represent traces or pairs trace-world, which, slightly abusing notation, we will respectively denote by ϑ and (ϑ, w), possibly subscripted or superscripted.

Definition 1. *Let α be a well-formed formula (ϕ, σ or τ) and $(\vartheta, w), \vartheta \in \mathcal{L}$. Then $(\vartheta, w) : \alpha$ is a* labeled well-formed formula *(*labeled formula *or* lwff *for short), and the set of* relational well-formed formulas *(*relational formulas *or* rwffs *for short) ρ is defined as follows:*

$$\rho ::= (\vartheta, w_i) \lhd (\vartheta, w_j) \mid (\vartheta, w_i) < (\vartheta, w_j) \mid (\vartheta, w_i) \prec (\vartheta, w_j) \mid (\vartheta, w_i) \simeq (\vartheta', w_j) \mid (\vartheta, w_i)\iota(\vartheta', w_j) \mid \vartheta \sim \vartheta'$$

Intuitively, $(\vartheta, w) : \alpha$ means that α is true at the world represented by w in the trace represented by ϑ and

- $(\vartheta, w_i) \lhd (\vartheta, w_j)$ means that the world represented by w_j is the immediate successor of the world represented by w_i in the trace represented by ϑ;
- $(\vartheta, w_i) < (\vartheta, w_j)$ means that the world represented by w_i precedes the world represented by w_j in the trace represented by ϑ, i.e., either $(\vartheta, w_i) \lhd (\vartheta, w_j)$ or there is a w_k such that $(\vartheta, w_i) \lhd (\vartheta, w_k)$ and $(\vartheta, w_k) < (\vartheta, w_j)$;
- $(\vartheta, w_i) \prec (\vartheta, w_j)$ means that the world represented by w_i precedes the world represented by w_j in the trace represented by ϑ and that in that trace there is at least one world w_k in-between w_i and w_j, i.e., there is a w_k such that $(\vartheta, w_i) \lhd (\vartheta, w_k)$ and $(\vartheta, w_k) < (\vartheta, w_j)$;
- $(\vartheta, w_i) \simeq (\vartheta', w_j)$ means that the world represented by w_i in the trace represented by ϑ is equivalent (as defined formally in Section 2.6) to the world represented by w_j in the trace represented by ϑ';
- $(\vartheta, w_i)\iota(\vartheta', w_j)$ means that the world represented by w_i in the trace represented by ϑ occurs in the same instant of time of the world represented by w_j in the trace represented by ϑ'; and
- $\vartheta \sim \vartheta'$ means that the trace represented by ϑ is equivalent to the trace represented by ϑ', i.e., $(\vartheta, w_1) \simeq (\vartheta', w'_1), (\vartheta, w_2) \simeq (\vartheta', w'_2)$ and so on when $\vartheta = w_1, w_2, \cdots$ and $\vartheta' = w'_1, w'_2, \cdots$.

2.4 Semantics of the first two layers

We are now ready to give the semantics for the first two layers of our language, postponing the interpretation of block and mitigation to a later subsection.

Definition 2. *Models for the first two layers are tuples of the form* $\mathfrak{M} = (W, \Theta, \mathcal{I}, \mathfrak{R}_{\lhd}, \mathfrak{R}_{<}, \mathfrak{R}_{\prec}, V)$*, where* W *is the set of worlds,* Θ *is the set of traces;*

- $\mathcal{I} : \mathcal{L} \to (\Theta, W)$ *is a (overloaded) function that maps every label* (ϑ, w) *to a pair trace-world, i.e.,* $\mathcal{I}((\vartheta, w)) = (\mathcal{I}(\vartheta), \mathcal{I}(w)) = (\vartheta, w)$*, and every label* ϑ *to a single trace, i.e.,* $\mathcal{I}(\vartheta) = \vartheta$*;*[2]
- $\mathfrak{R}_{<}$ *is a relation that holds true for any two worlds* w_i *and* w_j *that are in the same trace and such that* w_i *occurs before* w_j*;*
- $\mathfrak{R}_{\lhd}$ *is a relation that holds true for any two worlds* w_i *and* w_j *that are in the same trace and such that* w_i *is the immediate predecessor of* w_j*, i.e.,* $(w_i, w_j) \in \mathfrak{R}_{\lhd}$ *iff* $(w_i, w_j) \in \mathfrak{R}_{<}$ *and there is no* w_k *such that* $(w_i, w_k) \in \mathfrak{R}_{<}$ *and* $(w_k, w_j) \in \mathfrak{R}_{<}$*;*
- $\mathfrak{R}_{\prec}$ *is a relation that holds true for any two worlds* w_i *and* w_j *that are in the same trace and such that* w_i *precedes* w_j *and is not its immediate predecessor, i.e.,* $(w_i, w_j) \in \mathfrak{R}_{\prec}$ *iff there exists a* w_k *such that* $(w_i, w_k) \in \mathfrak{R}_{\lhd}$ *and* $(w_k, w_j) \in \mathfrak{R}_{<}$*;*
- $V : \mathcal{P} \times W \to \{\top, \bot\}$ *is a* valuation function *that assigns a truth value to a propositional variable* p *with respect to a given world* w*.*

Truth for a rwff or lwff in a model $\mathfrak{M}$ *is the smallest relation* $\models^{\mathfrak{M}}$ *satisfying:*

$$
\begin{array}{lll}
\models^{\mathfrak{M}} (\vartheta, w_i) \bullet (\vartheta, w_j) & \textit{iff} & (\mathcal{I}((\vartheta, w_i)), \mathcal{I}((\vartheta, w_j))) \in \mathfrak{R}_{\bullet} \textit{ for } \bullet \in \{\mathfrak{R}_{\lhd}, \mathfrak{R}_{<}, \mathfrak{R}_{\prec}\} \\
\models^{\mathfrak{M}} (\vartheta, w) : p & \textit{iff} & V(p, \mathcal{I}(w)) = \top \\
\models^{\mathfrak{M}} (\vartheta, w) : \neg\phi & \textit{iff} & \not\models^{\mathfrak{M}} (\vartheta, w) : \phi \\
\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \to \phi_2 & \textit{iff} & \models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \textit{ implies } \models^{\mathfrak{M}} (\vartheta, w) : \phi_2 \\
\models^{\mathfrak{M}} (\vartheta, w) : \Box\phi & \textit{iff} & \textit{for all } (\vartheta, w_i) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_i) \textit{ implies } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi \\
\models^{\mathfrak{M}} (\vartheta, w) : \Diamond\phi & \textit{iff} & \textit{exists } (\vartheta, w_i) \in \mathcal{L} \textit{ s.t. } \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_i) \textit{ and } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi \\
\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \,\mathcal{C}\, \phi_2 & \textit{iff} & \textit{for all } (\vartheta, w_i) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_i) \textit{ and } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1 \\
 & & \textit{imply } \models^{\mathfrak{M}} (\vartheta, w_i) : \Diamond\phi_2 \\
\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \,\mathcal{P}\, \phi_2 & \textit{iff} & \textit{for all } (\vartheta, w_j) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta, w) \prec (\vartheta, w_j) \textit{ and } \models^{\mathfrak{M}} (\vartheta, w_j) : \phi_2 \\
 & & \textit{imply that there exists } (\vartheta, w_i) \in \mathcal{L} \textit{ s.t. } \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_i) \\
 & & \textit{and } \models^{\mathfrak{M}} (\vartheta, w_i) < (\vartheta, w_j) \textit{ and } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1
\end{array}
$$

2.5 Tableau rules for the first two layers

In order to formalize reasoning about risk, we give a set of tableau rules. We assume that the reader is familiar with the standard tableau terminology and notation, e.g., [6]. As usual, a branch of a (possibly infinite) tableau is: *exhausted* if no more rules are applicable, *closed* if it contains the special judgment "Closed" and *open* if it is exhausted but not closed. A tableau is *closed* if all of its branches are closed. A rule that infers Closed is called a *closing rule*.

[2] Hence, for $\vartheta = w_1 w_2 \cdots$, we have $\mathcal{I}(\vartheta) = \mathcal{I}(\vartheta, w_1)\mathcal{I}(\vartheta, w_2) \cdots = w_1 w_2 \cdots$. Strictly speaking, we should use different symbols for labels in the syntax and traces/worlds in the semantics, but for simplicity we abuse notation and use the same symbols, and use labels and worlds/traces as synonyms.

The tableau rules for the first layer, which are shown in Fig. 2, are straightforward: they are just the labeled version of the standard rules (where ABS stands for absurdity). Rules for other connectives, such as $\wedge$ and $\vee$, can be given in the usual way. The tableau rules for the operators of the second layer are shown in Fig. 3. The positive rules ($\Box$) and ($\Diamond$) mimic the semantics, whereas the closure rules ($\Box$ABS) and ($\Diamond$ABS) tell us when we have a contradiction, but note that these two rules are actually derivable and could thus safely be omitted. The closure rules ($\mathcal{C}$ABS) and ($\mathcal{P}$ABS) are also derivable but we show them as their use simplifies the inferences.

The positive rule for ($\mathcal{C}$) follows the semantics, where we force the existence of the world w_k where the effect holds by requiring w_k to be fresh, i.e., different from all the worlds already present in the tableau up to that point.

Similarly, the positive rule of ($\mathcal{P}$) makes use of a fresh w_j in-between w_i and w_k, which we know must exist by the properties of $\prec$. Alternatively, we could have dispensed with the relation $\prec$ and forced our traces to be dense, but we have chosen not to do so as it would have complicated the formalization of the block and mitigation operators. Note also that we have defined ($\mathcal{P}$) to require that in-between the world where $\phi_1 \mathcal{P} \phi_2$ holds and that where ϕ_2 holds, the trace contains at least one world (where ϕ_1 holds). This basically means that if $\phi_1 \mathcal{P} \phi_2$ holds at w_i and ϕ_2 holds at its immediate successor w_j (i.e., $w_i \lhd w_j$), then $\phi_1 \mathcal{P} \phi_2$ has no "control" over ϕ_2. That is, we are modeling the situation where $\phi_1 \mathcal{P} \phi_2$ has been uttered too late in the process to have an influence on the ϕ_2 at w_j; rather, $(\vartheta, w_i) : \phi_1 \mathcal{P} \phi_2$ will require that for other occurrences of ϕ_2 holding at some future w_k, there is at least one world in-between w_i and w_k at which ϕ_1 holds.

Depending on the application, one might want to impose that causal formulas are *monotonic* in the sense that as soon as they become true, they stay true. This can be formalized by the following rules:

$$\frac{(\vartheta, w_1) : \phi_1 \mathcal{C} \phi_2 \quad (\vartheta, w_1) < (\vartheta, w_2)}{(\vartheta, w_2) : \phi_1 \mathcal{C} \phi_2} \ (\mathcal{C}\text{MON}) \qquad \frac{(\vartheta, w_1) : \phi_1 \mathcal{P} \phi_2 \quad (\vartheta, w_1) \prec (\vartheta, w_2)}{(\vartheta, w_2) : \phi_1 \mathcal{P} \phi_2} \ (\mathcal{P}\text{MON})$$

The rules in Fig. 4 formalize the properties of $<$, $\prec$ and $\lhd$; note that we give some rules schemas to save space (e.g., the rule ($\bullet$ABS) actually stands for three rules, one for each relation). We again have a number of options and variants that could be considered for these relational rules, e.g., we could introduce an explicit equality and add relational formulas of the form $(\vartheta, w_i) = (\vartheta, w_j)$ along with rules for the properties of $=$ (reflexivity, symmetry and transitivity) and extend the linearity rule ($<$ LIN) to

$$\frac{(\vartheta, w_1) \quad (\vartheta, w_2)}{(\vartheta, w_1) < (\vartheta, w_2) \ \mid \ (\vartheta, w_2) = (\vartheta, w_1) \ \mid \ (\vartheta, w_2) < (\vartheta, w_1)}$$

It is not difficult to see that these tableau rules are all sound (we omit the proof for space reasons). However, they are incomplete. Giving a complete tableau system is actually not an obvious task. As a simple example of the underlying difficulties, consider the closing rule ($\mathcal{P}$ABS) for the precondition,

$$\frac{(\vartheta, w) : \neg\neg\phi}{(\vartheta, w) : \phi}\ (\neg\neg) \qquad \frac{(\vartheta, w) : \phi \quad (\vartheta, w) : \neg\phi}{\textsc{Closed}}\ (\textsc{Abs})$$

$$\frac{(\vartheta, w) : \phi_1 \to \phi_2}{(\vartheta, w) : \neg\phi_1 \quad | \quad (\vartheta, w) : \phi_2}\ (\to) \qquad \frac{(\vartheta, w) : \neg(\phi_1 \to \phi_2)}{(\vartheta, w) : \phi_1,\ (\vartheta, w) : \neg\phi_2}\ (\neg\to)$$

Fig. 2. Tableau rules for the first layer

$$\frac{(\vartheta, w_i) : \Box\phi \quad (\vartheta, w_i) < (\vartheta, w_j)}{(\vartheta, w_j) : \phi}\ (\Box) \qquad \frac{(\vartheta, w_i) : \Box\phi \quad (\vartheta, w_i) < (\vartheta, w_j) \quad (\vartheta, w_j) : \neg\phi}{\textsc{Closed}}\ (\Box\textsc{Abs})$$

$$\frac{(\vartheta, w_i) : \Diamond\phi}{(\vartheta, w_j) : \phi,\ (\vartheta, w_i) < (\vartheta, w_j)}\ (\Diamond)\ \Big(w_j \text{ fresh}\Big) \qquad \frac{(\vartheta, w) : \Diamond\phi \quad (\vartheta, w) : \Box\neg\phi}{\textsc{Closed}}\ (\Diamond\textsc{Abs})$$

$$\frac{(\vartheta, w_i) : \phi_1 \,\mathcal{C}\, \phi_2 \quad (\vartheta, w_j) : \phi_1 \quad (\vartheta, w_i) < (\vartheta, w_j)}{(\vartheta, w_k) : \phi_2,\ (\vartheta, w_j) < (\vartheta, w_k)}\ (\mathcal{C})\ \Big(w_k \text{ fresh}\Big)$$

$$\frac{(\vartheta, w_i) : \phi_1 \,\mathcal{C}\, \phi_2 \quad (\vartheta, w_i) < (\vartheta, w_j) \quad (\vartheta, w_j) : \phi_1 \quad (\vartheta, w_j) : \Box\neg\phi_2}{\textsc{Closed}}\ (\mathcal{C}\textsc{Abs})$$

$$\frac{(\vartheta, w_i) : \phi_1 \,\mathcal{P}\, \phi_2 \quad (\vartheta, w_i) \prec (\vartheta, w_k) \quad (\vartheta, w_k) : \phi_2}{(\vartheta, w_j) : \phi_1,\ (\vartheta, w_i) < (\vartheta, w_j),\ (\vartheta, w_j) < (\vartheta, w_k)}\ (\mathcal{P})\ \Big(w_j \text{ fresh}\Big)$$

$$\frac{(\vartheta, w_i) : \phi_1 \,\mathcal{P}\, \phi_2 \quad (\vartheta, w_i) \prec (\vartheta, w_k) \quad (\vartheta, w_k) : \phi_2 \quad (\vartheta, w_i) : \Box\neg\phi_1}{\textsc{Closed}}\ (\mathcal{P}\textsc{Abs})$$

Fig. 3. Tableau rules for the second layer

which is not complete as it does not represent all the possible cases for closure (in fact, as we remarked above, this rule may simply be derived from $(\mathcal{P})$ and (Abs)). If $(\vartheta, w_i) : \phi_1 \,\mathcal{P}\, \phi_2$ and $(\vartheta, w_k) : \phi_2$ for $(\vartheta, w_i) \prec (\vartheta, w_k)$, then by the semantics we have a contradiction if $(\vartheta, w_j) : \neg\phi_1$ for all (ϑ, w_j) such that $(\vartheta, w_i) < (\vartheta, w_j) < (\vartheta, w_k)$. The rule $(\mathcal{P}\textsc{Abs})$, however, captures the scenario where $\neg\phi_1$ is true even after the occurrence of ϕ_2, so we are missing the case in which $\neg\phi_1$ is true between (ϑ, w_i) and (ϑ, w_k) but there may be occurrences of ϕ_1 in the future worlds of (ϑ, w_k). This could be easily formalized by means of the temporal operator *until*, denoted by $\mathcal{U}$, which we apply on event pairs:

$$\begin{aligned} \models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \,\mathcal{U}\, \phi_2 \ \text{ iff } \ & \models^{\mathfrak{M}} (\vartheta, w) : \phi_2 \text{ or there exists } (\vartheta, w_j) \in \mathcal{L} \text{ s.t. } \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_j) \\ & \text{and } \models^{\mathfrak{M}} (\vartheta, w_j) : \phi_2, \text{ and } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1 \text{ for all } (\vartheta, w_i) \in \mathcal{L} \text{ s.t.} \\ & \models^{\mathfrak{M}} (\vartheta, w) < (\vartheta, w_i) \text{ and } \models^{\mathfrak{M}} (\vartheta, w_i) < (\vartheta, w_j). \end{aligned}$$

We can then represent all the possible cases of the closing rule for $\mathcal{P}$ as follows:

$$\frac{(\vartheta, w_i) : \phi_1 \,\mathcal{P}\, \phi_2 \quad (\vartheta, w_i) \prec (\vartheta, w_k) \quad (\vartheta, w_k) : \phi_2 \quad (\vartheta, w_i) \lhd (\vartheta, w_{i+1}) \quad (\vartheta, w_{i+1}) : \neg\phi_1 \,\mathcal{U}\, \phi_2}{\textsc{Closed}}\ (\mathcal{P}\textsc{Abs}^{\mathcal{U}})$$

However, this comes at the cost of having to deal with $\mathcal{U}$, which is a notoriously difficult operator, mainly due to its dual nature of being both an existential and a universal operator (in the sense that it contains both kinds of quantification). While labeled inference rules for $\mathcal{U}$ do exist, they require some technical tricks to guarantee completeness, such as the use of Skolem functions to force the

$$\frac{(\vartheta,w_1)\lhd(\vartheta,w_2)}{(\vartheta,w_1)<(\vartheta,w_2)}\ (\lhd<) \qquad \frac{(\vartheta,w_1)\bullet(\vartheta,w_2)\quad(\vartheta,w_2)\bullet(\vartheta,w_1)}{\textsc{Closed}}\ (\bullet\textsc{Abs})\ \left[\bullet\in\{\lhd,<,\prec\}\right]$$

$$\frac{(\vartheta,w_1)\prec(\vartheta,w_2)}{(\vartheta,w_1)<(\vartheta,w_2)}\ (\prec<) \qquad \frac{(\vartheta,w_1)\bullet(\vartheta,w_2)\quad(\vartheta,w_2)\bullet(\vartheta,w_3)}{(\vartheta,w_1)\bullet(\vartheta,w_3)}\ (\bullet\textsc{Trans})\ \left[\bullet\in\{<,\prec\}\right]$$

$$\frac{(\vartheta,w_1)\quad(\vartheta,w_2)}{(\vartheta,w_1)<(\vartheta,w_2)\ \mid\ (\vartheta,w_2)<(\vartheta,w_1)}\ (<\textsc{Lin}) \qquad \frac{(\vartheta,w_1)\prec(\vartheta,w_2)}{(\vartheta,w_1)\lhd(\vartheta,w_3),\ (\vartheta,w_3)<(\vartheta,w_2)}\ (\prec)\ \left[\begin{smallmatrix}w_3\\ \text{fresh}\end{smallmatrix}\right]$$

$$\frac{(\vartheta,w_1)<(\vartheta,w_2)}{(\vartheta,w_1)\lhd(\vartheta,w_2)\ \mid\ (\vartheta,w_1)\lhd(\vartheta,w_3),\ (\vartheta,w_3)<(\vartheta,w_2)}\ (<)\ \left[w_3\text{ fresh}\right]$$

Fig. 4. Tableau rules for the relations (for the first two layers)

existence of certain worlds [3] or the use of additional operators such as the history operator of [13]. Rather than giving such rules here as well, we observe that to recover completeness for $\mathcal{P}$ (and the other operators) we can alternatively change the labeling discipline by allowing operations that work directly on the labels. For instance, we could then close for $\mathcal{P}$ as follows:

$$\frac{(\vartheta,w_i):\phi_1\,\mathcal{P}\,\phi_2\quad(\vartheta,w_i)<(\vartheta,w_k)\quad(\vartheta,w_k):\phi_2\quad(\vartheta|_{<w_k},w_i):\Box\neg\phi_1}{\textsc{Closed}}\ (\mathcal{P}\textsc{Abs}^{lab})$$

In both these alternatives ($\mathcal{P}\textsc{Abs}^{\mathcal{U}}$) and ($\mathcal{P}\textsc{Abs}^{lab}$), the technical price to pay is quite high so, depending on the application, one might even want to stick to the sound but incomplete system given above or to select the additional rules that are best fit for the concrete example under consideration.

2.6 Semantics of the third layer

Since a single world $w \in W$ can also be seen as the conjunction of all the formulas that are true at it, we can define two worlds to be *equivalent*, in symbols $w_1 \simeq w_2$, iff they make true the same propositional variables. By extension, two traces are equivalent, in symbols $\vartheta_1 \sim \vartheta_2$, iff their corresponding worlds for every instant of time are equivalent, so $\mathsf{itw}(l,\vartheta_1) \simeq \mathsf{itw}(l,\vartheta_2)$ for every $l \in T$.

Definition 3. *Models for the third layer extend those of the first two layers (cf. Definition 2) with three relations $\mathfrak{R}_\iota$, $\mathfrak{R}_\simeq$ and $\mathfrak{R}_\sim$, where*

- *$\mathfrak{R}_\iota$ is a relation that holds true for any two worlds w_i and w_j that are not in the same trace but occur in the same instant of time;*
- *$\mathfrak{R}_\simeq$ is an equivalence relation that holds true for any two worlds w_i and w_j such that $V(p,w_i) = V(p,w_j)$ for all propositional variables $p \in \Pi$;*
- *$\mathfrak{R}_\sim$ is an equivalence relation that holds true for any two traces ϑ' and ϑ'' whose corresponding worlds in every instant of time are equivalent, i.e., $(\mathsf{itw}(i,\vartheta'),\mathsf{itw}(i,\vartheta'')) \in \mathfrak{R}_\simeq$ for every $i \in T$.*

Truth for these rwffs is then defined as:

$$\begin{array}{ll}\models^{\mathfrak{M}} (\vartheta,w)\bullet(\vartheta',w_i) & \textit{iff}\ \ (\mathcal{I}((\vartheta,w)),\mathcal{I}((\vartheta',w_i))) \in \mathfrak{R}_\bullet \textit{ for } \bullet\in\{\mathfrak{R}_\iota,\mathfrak{R}_\simeq\}\\ \models^{\mathfrak{M}} \vartheta\sim\vartheta' & \textit{iff}\ \ (\mathcal{I}(\vartheta),\mathcal{I}(\vartheta')) \in \mathfrak{R}_\sim\end{array}$$

$$\frac{}{\vartheta \sim \vartheta}\ (\sim \textsc{Refl}) \qquad \frac{\vartheta \sim \vartheta'}{\vartheta' \sim \vartheta}\ (\sim \textsc{Sym}) \qquad \frac{\vartheta \sim \vartheta' \quad \vartheta' \sim \vartheta''}{\vartheta \sim \vartheta''}\ (\sim \textsc{Trans})$$

$$\frac{}{(\vartheta, w) \bullet (\vartheta, w)}\ (\bullet\textsc{Refl}) \qquad \frac{(\vartheta, w_1) \bullet (\vartheta, w_2)}{(\vartheta, w_2) \bullet (\vartheta, w_1)}\ (\bullet\textsc{Sym})$$

$$\frac{(\vartheta, w_1) \bullet (\vartheta, w_2) \quad (\vartheta, w_2) \bullet (\vartheta, w_3)}{(\vartheta, w_1) \bullet (\vartheta, w_3)}\ (\bullet\textsc{Trans}) \qquad \frac{(\vartheta, w_1) \simeq (\vartheta', w_2) \quad (\vartheta, w_1) : \phi}{(\vartheta', w_2) : \phi}\ (\simeq \textsc{Mon})$$

Fig. 5. Tableau rules for the relations for the third layer, where $\bullet \in \{\iota, \simeq\}$

The rules in Fig. 5 capture the properties of these relations.

Prevention As we already remarked in Section 2.1, one can consider different forms of prevention $\phi_1 \mathcal{B} \phi_2$, varying in the strength of the blocking event. For the strongest but also less interesting form, where the blocking event ϕ_1 completely prevents the second event ϕ_2, we can define

$\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \mathcal{B} \phi_2$ iff for all $(\vartheta, w_i), (\vartheta, w_j) \in \mathcal{L}$, $(w, w_i) \in \mathfrak{R}_<$ and $(w_i, w_j) \in \mathfrak{R}_<$ and $\models^{\mathfrak{M}} (\vartheta, w_j) : \phi_2$ imply $\not\models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1$

or, alternatively: iff for all $(\vartheta, w_i) \in \mathcal{L}$, $(w, w_i) \in \mathfrak{R}_<$ and $\models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1$ imply that there does not exist $(\vartheta, w_j) \in \mathcal{L}$ s.t. $(w_i, w_j) \in \mathfrak{R}_<$ and $\models^{\mathfrak{M}} (\vartheta, w_j) : \phi_2$

to express that if ϕ_2 occurs then it cannot be that ϕ_1 occurred previously but after the blocking formula (and thus again note that we can consider variants depending on when we actually let the blocking formula and the two events occur), or that if ϕ_1 occurs then ϕ_2 cannot occur in the future. However, one might typically want to consider a more refined definition of prevention, where ϕ_1 reduces the future occurrences of ϕ_2. To that end, instead of considering worlds occurring in the same trace, we need to compare traces where ϕ_1 occurs with those where ϕ_1 does not occur. In the trace where ϕ_1 occurs, so the prevention measure is in act, the occurrences of ϕ_2 are less than the occurrences of it in the trace where ϕ_1 does not occur. More specifically, the definition of prevention that we introduce says that $\phi_1 \mathcal{B} \phi_2$ is true at a given world w of a given trace ϑ iff, for all traces ϑ' equivalent to ϑ differing only for the occurrence of ϕ_1, since in ϑ we don't have ϕ_1 and instead ϕ_1 occurs in ϑ' after $\phi_1 \mathcal{B} \phi_2$, we have that, for all occurrences of ϕ_2 in ϑ', in the same instant of time we have a world in ϑ where ϕ_2 is true, and there are some occurrences of ϕ_2 in ϑ such that in the same instant of time there is a world in ϑ' where ϕ_2 is not true:

$\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \mathcal{B} \phi_2$ iff for all $\vartheta' \in \mathcal{L}$, for all $(\vartheta, w_i) \in \mathcal{L}$, exists $(\vartheta', w_j) \in \mathcal{L}$ s.t.
$((\vartheta, w) < (\vartheta, w_i)$ and $\models^{\mathfrak{M}} (\vartheta, w_i) : \neg\phi_1$ and $(\vartheta, w_i)\iota(\vartheta', w_j)$ and
$(\vartheta, w) < (\vartheta', w_j)$ and $\vartheta|_{<(\vartheta, w_i)} \sim \vartheta'|_{<(\vartheta', w_j)}$ and $\models^{\mathfrak{M}} (\vartheta', w_j) : \phi_1)$
(imply that for all $(\vartheta', w_y) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta', w_y) : \phi_2$ and $(\vartheta', w_j) < (\vartheta', w_y)$
imply that exists $(\vartheta, w_x) \in \mathcal{L}$, s.t. $\models^{\mathfrak{M}} (\vartheta, w_x) : \phi_2$ and
$(\vartheta, w_i) < (\vartheta, w_x)$ and $(\vartheta, w_x)\iota(\vartheta', w_y))$ and (exist $l > 0$ and $(\vartheta, w_{k,i})$,
$(\vartheta', w_{h,i}) \in \mathcal{L}$, where $0 < i \leq l$, s.t. $(\vartheta, w_i) < (\vartheta, w_{k,i})$ and
$\models^{\mathfrak{M}} (\vartheta, w_{k,i}) : \phi_2$ and $(\vartheta', w_j) < (\vartheta', w_{h,i})$ and $\models^{\mathfrak{M}} (\vartheta', w_{h,i}) : \neg\phi_2$ and
$(\vartheta, w_{k,i})\iota(\vartheta', w_{h,i}))$

We can then give the following tableau rules for prevention, which, however, require us to extend the labeling discipline as we described above and allow for relational formulas of the form $\vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0}$:

$$\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{B}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \\ (\vartheta, w) < (\vartheta, w_i) \quad \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta', w_y) : \phi_2 \quad (\vartheta, w_0) < (\vartheta', w_y)\end{array}}{\begin{array}{c}(\vartheta', w_j)\iota(\vartheta, w_i),\ (\vartheta, w_0) \lhd (\vartheta', w_j),\ (\vartheta', w_j) : \phi_1,\ (\vartheta, w) < (\vartheta', w_j),\ (\vartheta, w_x) : \phi_2, \\ (\vartheta', w_j) < (\vartheta', w_y),\ (\vartheta, w_x)\iota(\vartheta', w_y),\ (\vartheta, w_i) < (\vartheta, w_x),\ (\vartheta, w_i) < (\vartheta, w_k), \\ (\vartheta', w_j) < (\vartheta', w_h),\ (\vartheta, w_k)\iota(\vartheta', w_h),\ (\vartheta, w_k) : \phi_2,\ (\vartheta', w_h) : \neg\phi_2\end{array}}\ (\mathcal{B})\ \left[\begin{array}{c} w_j,\ w_x, \\ w_k,\ w_h \\ \text{fresh}\end{array}\right]$$

$$\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{B}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \quad (\vartheta, w) < (\vartheta, w_i) \\ \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta', w_y) : \phi_2 \quad (\vartheta, w_0) < (\vartheta', w_y) \quad (\vartheta', w_j) : \Box\phi_2 \quad (\vartheta, w_0) \lhd (\vartheta', w_j)\end{array}}{\textsc{Closed}}\ (\mathcal{B}\textsc{Abs}_1)$$

$$\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{B}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \quad (\vartheta, w) < (\vartheta, w_i) \\ \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta', w_y) : \phi_2 \quad (\vartheta, w_0) < (\vartheta', w_y) \quad (\vartheta, w_i) : \Box\neg\phi_2\end{array}}{\textsc{Closed}}\ (\mathcal{B}\textsc{Abs}_2)$$

The positive rule of $\mathcal{B}$ follows the semantics, where given the ϑ-world w_i where ϕ_1 is not true, we force the existence of a fresh world w_j where ϕ_1 is true, in every other trace ϑ', equivalent to the given one, until w_j. For every world w_y in ϑ', where ϕ_2 is true, we introduce a fresh ϑ-world w_x where ϕ_2 is also true. We also introduce two fresh worlds, w_k in ϑ and w_h in ϑ', such that ϕ_2 is true at (ϑ, w_k) and it is instead false at (ϑ', w_h).

As before, these rules are sound but the two closing rules, which are used together and not as alternatives, are not complete because they do not represent all the possible cases. The rule ($\mathcal{B}\textsc{Abs}_1$) does not capture the scenario in which there are occurrences of $\neg\phi_2$ in ϑ' and still the number of occurrences of ϕ_2 in ϑ' is bigger than the number of occurrences of ϕ_2 in ϑ. ($\mathcal{B}\textsc{Abs}_2$) does not capture the scenario in which there are occurrences of ϕ_2 in ϑ and still the number of occurrences of ϕ_2 in ϑ' is bigger than the number of occurrences of ϕ_2 in ϑ.

We can force $\mathcal{B}$ to be monotonic over $<$ by adding a rule ($\mathcal{B}\textsc{Mon}$) analogous to ($\mathcal{C}\textsc{Mon}$).

Mitigation As for prevention, we could consider a strong definition of mitigation $\phi_1 \,\mathcal{M}\, \phi_2$ such that ϕ_1 prevents all occurrences of the effects of ϕ_2:

$$\begin{array}{rl}\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \,\mathcal{M}\, \phi_2 \ \text{ iff } & \text{for all } w_i, w_j \in W,\ (w, w_i) \in \mathfrak{R}_< \text{ and } \models^{\mathfrak{M}} (\vartheta, w_i) : \phi_1 \text{ and} \\ & (w_j, w_i) \in \mathfrak{R}_< \text{ and } \models^{\mathfrak{M}} (\vartheta, w_j) : \phi_2 \text{ imply that for all } \phi_3 \in E \text{ and} \\ & \text{for all } w_x \in W \text{ s.t. } \models^{\mathfrak{M}} (\vartheta, w_x) : \phi_2 \,\mathcal{C}\, \phi_3 \text{ and } (w_x, w_j) \in \mathfrak{R}_< \\ & \text{we have that } \models^{\mathfrak{M}} (\vartheta, w_j) : \Box\neg\phi_3\end{array}$$

However, it is more interesting to consider a form of mitigation that does not block completely the occurrences of the effects of ϕ_2 but instead makes them decrease: $\phi_1 \,\mathcal{M}\, \phi_2$ means that ϕ_1 prevents all the effects of the event ϕ_2. Alternatively, we could define that it prevents only some of the effects. In both such cases, we highlight the second-order nature of this operator, which could however be pushed down to the propositional level if one were certain that such effects were finitely many. Compare, for instance, the well-known (and thus finitely enumerable) undesired effects of a commercial medicine with the still uncategorized undesired effects of an experimental treatment.

We could even define a very weak mitigation $\phi_1 \,\mathcal{M}\, \phi_2[\phi_3]$ that prevents only one effect: the event ϕ_1 mitigates the event ϕ_2 by preventing its effect ϕ_3.

As an example, we formalize the definition of mitigation that we consider the most complete: $\phi_1 \,\mathcal{M}\, \phi_2$ means that ϕ_1 prevents all occurrences of the effects of ϕ_2. This definition can easily be used to provide a more sophisticated formulation, in which we also are able to prevent the effects of a given threat that satisfy a given condition. If we write $\phi_1 \,\mathcal{M}\, \phi_2[\psi]$, we mean that ϕ_1 prevents all the effects of ϕ_2 that satisfy ψ. This extension is left for further work.

The definition of mitigation that we formalize says that $\phi_1 \,\mathcal{M}\, \phi_2$ is true in a given world w of a given trace ϑ iff, for all traces ϑ' equivalent to ϑ differing only for the occurrence of ϕ_1, since in ϑ we don't have ϕ_1 and instead ϕ_1 occurs in ϑ' after $\phi_1 \,\mathcal{M}\, \phi_2$ and ϕ_2, we have that, for all the occurrences of the event ϕ_3 in ϑ', such that ϕ_3 is an effect of ϕ_2 that comes after the occurrences of ϕ_2, which itself comes after $\phi_2 \,\mathcal{C}\, \phi_3$, in the same instant of time we have a world in ϑ at which ϕ_3 is true, and there are some occurrences of ϕ_3 in ϑ such that in the same instant of time there is a world in ϑ' at which ϕ_3 is not true.

$$
\begin{aligned}
\models^{\mathfrak{M}} (\vartheta, w) : \phi_1 \,\mathcal{M}\, \phi_2 \;\text{ iff }\; & \text{for all } \vartheta' \in \mathcal{L},\ \text{for all } (\vartheta, w_i) \in \mathcal{L},\ \text{exists } (\vartheta', w_j) \in \mathcal{L}, (\vartheta, w) < (\vartheta, w_i) \\
& \text{and } \models^{\mathfrak{M}} (\vartheta, w_i) : \neg\phi_1 \text{ and } (\vartheta, w_i)\iota(\vartheta', w_j) \text{ and } (\vartheta, w) < (\vartheta', w_j) \text{ and} \\
& \vartheta|_{<(\vartheta, w_i)} \sim \vartheta'|_{<(\vartheta', w_j)} \text{ and } \models^{\mathfrak{M}} (\vartheta', w_j) : \phi_1 \text{ implies for all } \phi_3 \in E, \\
& \text{for all } (\vartheta, w_p), (\vartheta, w_r) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta, w_p) : \phi_2 \,\mathcal{C}\, \phi_3 \text{ and } (\vartheta, w_p) < (\vartheta, w_r) \\
& \text{and } (\vartheta, w_r) < (\vartheta, w_i) \text{ and } \models^{\mathfrak{M}} (\vartheta, w_r) : \phi_2 \text{ implies} \\
& (\text{for all } (\vartheta', w_y) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta', w_y) : \phi_3 \text{ and } (\vartheta', w_j) < (\vartheta', w_y) \text{ implies} \\
& \text{exists } (\vartheta, w_x) \in \mathcal{L}, \models^{\mathfrak{M}} (\vartheta, w_x) : \phi_3 \text{ and } (\vartheta, w_i) < (\vartheta, w_x) \text{ and} \\
& (\vartheta, w_x)\iota(\vartheta', w_y)) \text{ and } (\text{exists } l > 0,\ \text{exist } (\vartheta, w_{k,i}), (\vartheta', w_{h,i}) \in \mathcal{L}, \\
& \text{where } 0 < i \leq l,\ \text{s.t. } (\vartheta, w_i) < (\vartheta, w_{k,i}) \text{ and } \models^{\mathfrak{M}} (\vartheta, w_{k,i}) : \phi_3 \text{ and} \\
& (\vartheta', w_j) < (\vartheta', w_{h,i}) \text{ and } \models^{\mathfrak{M}} (\vartheta', w_{h,i}) : \neg\phi_3 \text{ and } (\vartheta, w_{k,i})\iota(\vartheta', w_{h,i}))
\end{aligned}
$$

We can then give the following tableau rules for mitigation, a positive rule and two closure rules, which, again, are sound but incomplete:

$$
\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{M}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \\ (\vartheta, w) < (\vartheta, w_i) \quad \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta, w_p) : \phi_2 \,\mathcal{C}\, \phi_3 \quad (\vartheta, w_r) : \phi_2 \\ (\vartheta, w_p) < (\vartheta, w_r) \quad (\vartheta, w_r) < (\vartheta, w_i) \quad (\vartheta', w_y) : \phi_3 \quad (\vartheta, w_0) < (\vartheta', w_y)\end{array}}{\begin{array}{c}(\vartheta', w_j)\iota(\vartheta, w_i),\ (\vartheta, w_0) \lhd (\vartheta', w_j),\ (\vartheta', w_j) : \phi_1,\ (\vartheta, w) < (\vartheta', w_j), \\ (\vartheta', w_j) < (\vartheta', w_y),\ (\vartheta, w_x)\iota(\vartheta', w_y),\ (\vartheta, w_i) < (\vartheta, w_x),\ (\vartheta, w_i) < (\vartheta, w_k), \\ (\vartheta, w_x) : \phi_3,\ (\vartheta', w_j) < (\vartheta', w_h),\ (\vartheta, w_k)\iota(\vartheta', w_h),\ (\vartheta, w_k) : \phi_3,\ (\vartheta', w_h) : \neg\phi_3\end{array}} \ (\mathcal{M}) \left\{ \begin{array}{c} w_j,\ w_x, \\ w_k,\ w_h \\ \text{fresh} \end{array} \right.
$$

$$
\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{M}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \quad (\vartheta, w) < (\vartheta, w_i) \\ \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta, w_p) : \phi_2 \,\mathcal{C}\, \phi_3 \quad (\vartheta, w_r) : \phi_2 \quad (\vartheta, w_p) < (\vartheta, w_r) \\ (\vartheta, w_r) < (\vartheta, w_i) \quad (\vartheta', w_y) : \phi_3 \quad (\vartheta, w_0) < (\vartheta', w_y) \quad (\vartheta', w_j) : \Box\phi_3 \quad (\vartheta, w_0) \lhd (\vartheta', w_j)\end{array}}{\textsc{Closed}} \ (\mathcal{M}\textsc{Abs}_1)
$$

$$
\frac{\begin{array}{c}(\vartheta, w) : \phi_1 \,\mathcal{M}\, \phi_2 \quad (\vartheta, w_i) : \neg\phi_1 \quad (\vartheta, w_0) \lhd (\vartheta, w_i) \\ (\vartheta, w) < (\vartheta, w_i) \quad \vartheta|_{\leq w_0} \sim \vartheta'|_{\leq w_0} \quad (\vartheta, w_p) : \phi_2 \,\mathcal{C}\, \phi_3 \quad (\vartheta, w_r) : \phi_2 \\ (\vartheta, w_p) < (\vartheta, w_r) \quad (\vartheta, w_r) < (\vartheta, w_i) \quad (\vartheta', w_y) : \phi_3 \quad (\vartheta, w_0) < (\vartheta', w_y) \quad (\vartheta, w_i) : \Box\neg\phi_3\end{array}}{\textsc{Closed}} \ (\mathcal{M}\textsc{Abs}_2)
$$

The positive rule of $\mathcal{M}$ follows the semantics, where given the ϑ-world w_i, where ϕ_1 is not true, we force the existence of a fresh world w_j where ϕ_1 is true, in every other trace ϑ', equivalent to the given one, until w_j. For all ϑ'-worlds w_y, where ϕ_3 is true, that come after the occurrence ϕ_2, that itself comes after the

occurrence of $\phi_2 \mathcal{C} \phi_3$ and before the occurrence of ϕ_1, we introduce a fresh ϑ-world w_x, where ϕ_3 is also true. We also introduce two fresh worlds w_k in ϑ and w_h in ϑ', with ϕ_3 true in w_k and not true in w_h.

The rule ($\mathcal{M}$ABS$_1$) doesn't capture the case when there are occurrences of $\neg\phi_3$ in ϑ' and still the number of occurrences of ϕ_3 in ϑ' is bigger than the number of occurrences of ϕ_3 in ϑ, whereas the rule ($\mathcal{M}$ABS$_2$) doesn't capture the case when there are occurrences of ϕ_3 in ϑ and still the number of occurrences of ϕ_3 in ϑ' is bigger than the number of occurrences of ϕ_3 in ϑ.

We can force mitigation to be monotonic as for the other operators.

3 A Case Study

To illustrate our framework at work, we return to the case study presented in the introduction. For simplicity, but without loss of generality, we make the standard *closed world* assumption, i.e., in a given world, every formula is false unless it is explicitly not asserted to be true. Also, we adopt a propositional language: the medical staff STAFF $= \{d_1, d_2, \ldots\}$ and the patients PATIENTS $= \{c_1, c_2, \ldots\}$ are finite sets, where d_i and c_j are propositional variables, and thus we employ $\forall$ and $\exists$ simply as abbreviations for finite conjunctions and disjunctions. We write c to denote a generic patient and d to denote a generic doctor or nurse.

When a new patient c is hospitalized, the first step is her registration $Reg(c)$, which *causes* the generation of three records:

$$Reg(c) \mathcal{C} Gen(AR_c) \qquad Reg(c) \mathcal{C} Gen(NR_c) \qquad Reg(c) \mathcal{C} Gen(RR_c).$$

The registration is a *precondition* for the assignment of a doctor to the patient and for the access of the normal records of a patient by the medical staff. The assigned doctor has full access to all the medical data of her patient (of course, the assigned doctor does not need the administrative data of the patient). The accesses made by the members of the STAFF that are not the assigned doctor cause the leak of personal information of the patient, expressed by a *Privacy.Leak* event. All the concepts represented above are given as follows:

$$\begin{array}{cc} Reg(c) \mathcal{P} \exists d. Assigned(d, c) & Reg(c) \mathcal{P} Access(d, NR_c) \\ Assigned(d, c) \mathcal{P} Access(d, NR_c) & Assigned(d, c) \mathcal{P} Access(d, RR_c) \\ \multicolumn{2}{c}{(Assigned(d, c) \wedge \exists x. (Access(x, NR_c) \wedge x \neq d)) \mathcal{C} Privacy.Leak(c).} \end{array}$$

A patient c can be transferred, $Transfer(c)$, to another hospital (which should be a parameter but we omit it for simplicity): after the occurrence of $Transfer(c)$, the accesses to the records of c are reduced as the accesses to the patient records are made just for examination and consultation with the doctors of the new hospital of the patient, or for statistic or research aims. The transfer is made if c is registered, and when it occurs it *prevents* access to the patient's data:

$$\begin{array}{cc} Reg(c) \mathcal{P} Transfer(c) & Transfer(c) \mathcal{B} Access(d, AR_c) \\ Transfer(c) \mathcal{B} Access(d, NR_c) & Transfer(c) \mathcal{B} Access(d, RR_c) \end{array}$$

We call *rule formulas* all of the above 12 formulas. They have to be true before the involved event takes place. If they are true after the occurrence of the

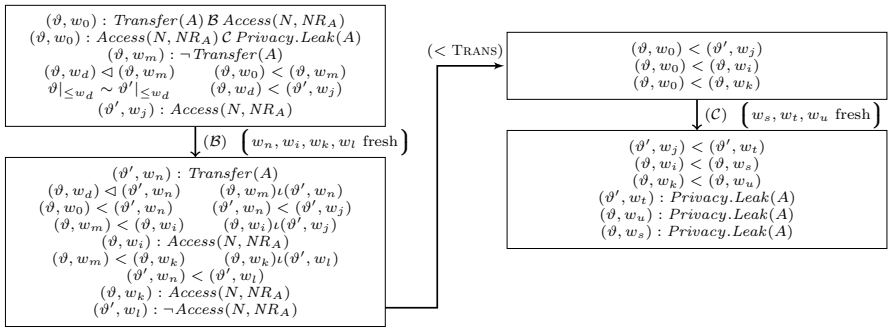

Fig. 6. Tableau fragment for the case study

event, then they can not be applied, but they can be used in the next occurrence of that event, in case they are still true. In the following scenario, we assume all these formulas to be true from the first instant of time, i.e., from the initial world (ϑ, w_0); for the sake of space we are not going to write all of them, but just the formulas we are going to use. Other rule formulas can be added to the system if they are needed.

Assume Alice (A) is a patient of the clinic and registers at time r, so $Reg(A)$ is true at world (ϑ, w) such that $\mathcal{I}(\vartheta, w) = \mathsf{itw}(r, \vartheta)$, and which comes after (ϑ, w_0). Using rule $(\mathcal{C})$, we can deduce that there is a world (ϑ, w_g) in the future of (ϑ, w), where the generation of NR_A takes place, i.e., $(\vartheta, w_g) : Gen(NR_A)$.

At time i, corresponding to (ϑ, w_i), doctor Debbie (D) modifies the restricted records of A, i.e., $(\vartheta, w_i) : Access(D, RR_A)$ and $(\vartheta, w_0) \prec (\vartheta, w_i)$. Then, given $(\vartheta, w_0) : Assigned(D, A)\,\mathcal{P}\,Access(D, RR_A)$, rule $(\mathcal{P})$ yields that there is a new world (ϑ, w_k), between (ϑ, w_0) and (ϑ, w_i), where the assignment of D to A occurs. Assume now that at world (ϑ, w_d), D decides if A needs to be transferred to another hospital. At the immediate successor (ϑ, w_m) of (ϑ, w_d), the formula $Transfer(A)$ is false. There is another branch of the trace, denoted by ϑ', that is equivalent to ϑ up to (and including) (ϑ, w_d). At world (ϑ', w_j), nurse Nancy (N) accesses the data of (A), as shown in Fig. 7. The assertion $Assigned(D, A)$ establishes that D is assigned to A. In this context, we assume that N is not assigned to A, which is anyway guaranteed by the closed world assumption. In Fig. 6, we give a significant fragment of a tableau for this scenario, where rule $(\mathcal{B})$ is used to prevent the accesses to the records of A. If the inference continues with the application of $(<$ Trans$)$ and $(\mathcal{C})$, then we will also see that the occurrences of $Privacy.Leak(A)$ in ϑ are less than those in ϑ.

4 Concluding Remarks

In this paper, we investigated a general framework for reasoning about risks. The approach we have taken consists in designing a flexible system in order to adapt the framework to the different contexts which it may be applied to, and here we

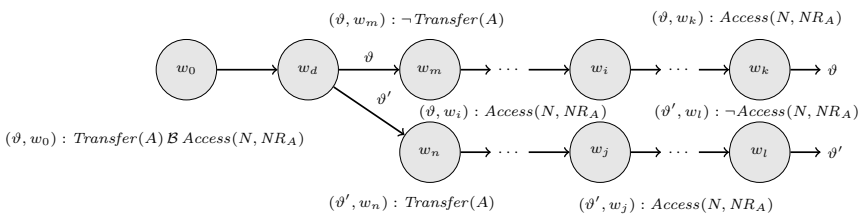

Fig. 7. The scenario for $\mathcal{B}$ in the case study

have over scratched the surface of the landscape of alternative possibilities. As we remarked, there are several ways in which this research can be taken further. We aim, in particular, at devising a complete tableau system and automating the deduction process: in addition to theorem proving, we envision a model checking procedure that would allow us to tackle concrete case studies taken from industrial practice (such as more complex scenarios for the case study discussed above).

References

1. AVANTSSAR. Deliverable 5.1: Problem cases and their trust and security requirements. www.avantssar.eu, 2008.
2. S. Bartsch. A calculus for the qualitative risk assessment of policy override authorization. In *SIN*, pp. 62–70, 2010.
3. D. A. Basin, C. Caleiro, J. Ramos, and L. Viganò. Labelled tableaux for distributed temporal logic. *Journal of Logic and Computation*, 19(6):1245–1279, 2009.
4. J. Bell. A common sense theory of causation. In *CONTEXT*, Springer, 2003.
5. P. C. Chapin, C. Skalka, and X. S. Wang. Risk assessment in distributed authorization. In *FMSE*, pp. 33–42, 2005.
6. M. D'Agostino, D. M. Gabbay, R. Hähnle, and J. Posegga, editors. *Handbook of Tableau Methods*. Kluwer Academic Publishers, 1999.
7. J. M. Dunn. Positive modal logic. *Studia Logica*, 55:301–317, 1995.
8. D. M. Gabbay. *Labelled Deductive Systems*. Clarendon Press, 1996.
9. E. Giunchiglia, J. Lee, V. Lifschitz, N. McCain, and H. Turner. Nonmonotonic causal theories. *Artificial Intelligence*, 153(1-2):49–104, 2004.
10. D. Lewis. Causation. *The Journal of Philosophy*, 70(17):556–567, 1973.
11. D. Lewis. Causation as influence. *The Journal of Philosophy*, 97(4):182–197, 2000.
12. N. Li and J. C. Mitchell. A role-based trust-management framework. In DISCEX-III, pp. 201–212. IEEE Computer Society, 2003.
13. A. Masini, L. Viganò, and M. Volpe. A history of until. *ENTCS* 262:189–204, 2010.
14. G. Shafer, P. R. Gillett, and R. B. Scherl. The logic of events. *Annals of Mathematics and Artificial Intelligence*, 28(1-4):315–389, 2000.
15. A. Singh and D. J. Lilja. Improving risk assessment methodology: a statistical design of experiments approach. In *SIN*, pp. 21–29, 2009.
16. P. Terenziani and P. Torasso. Time, action-types, causation: An integrated analysis. *Computational Intelligence*, 11:529–552, 1995.
17. H. Turner. A logic of universal causation. *AI*, 113(1-2):87–123, 1999.
18. L. Viganò. *Labelled Non-Classical Logics*. Kluwer Academic Publishers, 2000.