

Temporal extensions to Defeasible Logic

Guido Governatori¹ Paolo Terenziani²

¹ School Information Technology and Electrical Engineering,
The University of Queensland, email: guido@itee.uq.edu.au

² Dipartimento di Informatica, Università degli Studi del Piemonte Orientale, email: terenz@di.unito.it

Abstract

In this paper, we extend defeasible logic (a computationally-oriented non-monotonic logic) in order to deal with temporalised rules. In particular, we extend the logic to cope with durative facts, as well as with delays between the antecedent and the consequent of rules. We showed that the extended temporalised framework is suitable to model different types of causal relations which have been identified by the specialised literature. Finally, we also demonstrate that the computational properties of the original logic are still retained by the extended approach.

1 Introduction

Non-monotonic logic has been proposed as an appropriate representation of common-sense reasoning. As a matter of fact, the notion of defeasible conclusion, i.e., a conclusion that can be revised if more evidence is provided, is at heart of commonsense reasoning and therefore of non-monotonic logic. Over the years a plethora of non-monotonic formalisms have been investigated. However, a common issue of most non-monotonic logics is their high computational complexity. To obviate this problem Nute [20] proposed Defeasible logic. Defeasible logic is a rule based non-monotonic formalism that has been designed from beginning to be computationally feasible and easily implementable. Of course there is a trade-off between the expressive power of a logic and its computational complexity. Recently, a number of studies has shown that Defeasible Logic seems to be appropriate to reason in several application areas, ranging from modelling of contracts [12; 6], legal reasoning [21; 11], agent negotiations [7], modelling of agents and agent societies [10; 9], and applications to the Semantic Web [2]. An important finding of these investigations is that time is essential for an accurate representation of real world scenarios. While Defeasible logic proved to be suitable to cope with most of the phenomena specific to the application domains, the treatment of the temporal issues have been by large ignored, and still is an open problem. Recently a few extensions of Defeasible logic with time have been proposed [11; 8].

[11] proposes an extension where each proposition is paired with a timestamp representing the time when the proposition holds. The aim of this work [11] is to study the notion of persistence. On the other hand [8] extends defeasible logic by attaching two temporal dimensions to each rule: the first to tell when a rule is deliberated (i.e., when a rule is created), and the second for the time of validity of a rule (i.e. when a rule can be used). In both cases only instantaneous events are considered and there are no explicit temporal relationships between the antecedent and the conclusion of a rule. On the other hand such issues are important not only to model many application domains but also to represent theoretical notions such as, for instance, causal relationships.

In this paper we overcome such limitations by proposing an extension of defeasible Logic to cope with durative events and with delays between antecedents (causes) and conclusions (effects) of rules in the logic.

In this paper we will first recall the basics of defeasible logic (Section 2), then we introduce and motivate the extended temporal formalism (Section 3) 4 illustrates the application of the formalism to model some types of causal relations. In Section 5 we formally describe the logic and its inference mechanism. We also show that the proposed extension preserves the computational behaviour of defeasible logic. Finally we sketch conclusions and possible extensions.

2 Defeasible Logic

Over the years Defeasible Logic [20; 3] proved to be a flexible non-monotonic formalism able to capture different and sometimes incompatible facets of non-monotonic reasoning [4], and efficient and powerful implementations have been proposed [17; 12].

Knowledge in Defeasible Logic can be represented in two ways: facts and rules.

Facts are indisputable statements, represented either in form of states of affairs (literal and modal literal) and actions that have been performed. Facts are represented by predicates. For example, “Tweety is a penguin” is represented by *Penguin(Tweety)*.

A *rule*, on the other hand, describes the relationship between a set of literals (premises) and a literal (conclusion), and we can specify how strong the relationship is. As usual rules allow us to derive new conclusions given a set of

premises. As far as the strength of rules is concerned we distinguish between *strict rules*, *defeasible rules* and *defeaters*.

Strict rules, defeasible rules and defeaters are represented, respectively, by expressions of the form $A_1, \dots, A_n \rightarrow B$, $A_1, \dots, A_n \Rightarrow B$ and $A_1, \dots, A_n \rightsquigarrow B$, where A_1, \dots, A_n is a possibly empty set of prerequisites (causes) and B is the conclusion (effect) of the rule. We only consider rules that are essentially propositional. Rules containing free variables are interpreted as the set of their ground instances.

Strict rules are rules in the classical sense: whenever the premises are indisputable then so is the conclusion. Thus they can be used for definitional clauses. An example of a strict rule is “Penguins are birds”, formally: $Penguin(X) \rightarrow Bird(X)$.

Defeasible rules are rules that can be defeated by contrary evidence. An example of such a rule is “Birds usually fly”: $Bird(X) \Rightarrow Fly(X)$. The idea is that if we know that X is a bird, then we may conclude that X is entitled to a discount *unless there is other evidence suggesting that she may not fly*.

Defeaters are a special kind of rules. They are used to prevent conclusions not to support them. For example: $Heavy(X) \rightsquigarrow \neg Fly(X)$. This rule states that if something is heavy then it might not fly. This rule can prevent the derivation of a “fly” conclusion. On the other hand it cannot be used to support a “not fly” conclusion.

DL is a “skeptical” non-monotonic logic, meaning that it does not support contradictory conclusions. Instead DL seeks to resolve conflicts. In cases where there is some support for concluding A but also support for concluding $\neg A$, DL does not conclude either of them (thus the name “skeptical”). If the support for A has priority over the support for $\neg A$ then A is concluded. No conclusion can be drawn from conflicting rules in DL unless these rules are prioritised. The *superiority relation* among rules is used to define priorities among rules, that is, where one rule may override the conclusion of another rule. For example, given the defeasible rules

$$r : Bird(X) \Rightarrow Fly(X)$$

$$r' : Penguin(X) \Rightarrow \neg Fly(X)$$

which contradict one another, no conclusive decision can be made about whether a Tweety can fly or not. But if we introduce a superiority relation \succ with $r' \succ r$, then we can indeed conclude that Tweety cannot fly since it is a penguin.

We now give a short informal presentation of how conclusions are drawn in DL. A D be a theory in DL (i.e., a collection of facts, rules and a superiority relation). A *conclusion* of D is a tagged literal and can have one of the following four forms:

- $+\Delta_X q$ meaning that q is definitely provable in D (i.e., using only facts and strict rules).
- $-\Delta_X q$ meaning that we have proved that q is not definitely provable in D .
- $+\partial_X q$ meaning that q is defeasibly provable in D .
- $-\partial_X q$ meaning that we have proved that q is not defeasibly provable in D .

Strict derivations are obtained by forward chaining of strict rules, while a defeasible conclusion p can be derived if there

is a rule whose conclusion is p , whose prerequisites (antecedent) have either already been proved or given in the case at hand (i.e., facts), and any stronger rule whose conclusion is $\neg p$ has prerequisites that fail to be derived. In other words, a conclusion p is derivable when:

- p is a fact; or
- there is an applicable strict or defeasible rule for p , and either
 - all the rules for $\neg p$ are discarded (i.e., are proved to be not applicable) or
 - every applicable rule for $\neg p$ is weaker than an applicable strict¹ or defeasible rule for p .

The formal definitions of derivations and the proof conditions or inference rules are in Section 5, for the extension of the logic with the features we are going to describe in the next section.

3 Modelling temporal rules

In the following, we extend our previous approach in order to introduce the temporal dimension within the rules. At a first stage, we provide an informal high-level representation of the rules. We then refine it in the next section, showing how high-level rules can be actually modelled within DL. In the discussion below, we use the causal metaphora to explain the relationship between the antecedents and the consequent of the rules. Such a metaphora seems to us quite natural, especially in case rules are interpreted within a temporal framework. We will further investigate the relationships between our temporal rules and different definitions of causation in the literature in Section 4.

We represent rules of the general form:

$$a_1 : d_1, \dots, a_n : d_n \Rightarrow^d b : d_b$$

Such rules have the following general intuitive interpretation:

- In the rule, all events are represented as a pair $e : d$, where e is the identifier of the event and $d \in \mathbb{N}, d \geq 1$ is the duration of the event. As special cases, instantaneous event are modelled by $d = 1$, while persistent events by $d = +\infty$.
- a_1, \dots, a_n are the “causes”. They can start at different points in time.
- b is the “effect”.
- “ d ” ($d \in \mathbb{Z}$) is the *exact* delay between “causes” and “effect”. “ d ” cannot be infinite.

The above definition represents a whole general schema of rules, depending on the choice of the event endpoints which are related by the delay relation. One may consider the following alternative interpretations :

- (1) d is the delay between the start of the last cause and the beginning of the effect;

¹ Notice that a strict rule can be defeated only when its antecedent is defeasibly provable.

- (2) d is the delay between the end of the last cause and the beginning of the effect (of course, in such an interpretation the duration of the causes cannot be infinite).

Other alternatives are technically possible, but seem to us not very motivated from the applicative point of view. For instance, considering d the delay between the start (or end) of the first cause and the start of the effect seems to us quite counter-intuitive. In fact, if all a_1, \dots, a_n are explicitly necessary to trigger the rule, it seems quite unnatural that the rules is then triggered as soon as the first of them starts (ends) to hold.

Then to trigger a rule, different conditions have to be satisfied, depending on the chosen interpretation. For instance, let us consider the intuitive triggering conditions for case (2) above (and supposing that all durations are finite).

1. we must be able to prove each a_i for *exactly* d_i consecutive time intervals, i.e., $\forall i \exists t_0, t_1, \dots, t_{d_i}, t_{d_i+1}$ consecutive time intervals such that we can not prove a_i at time t_0 , we can prove a_i at times t_1, \dots, t_{d_i} , and we cannot prove a_i at time t_{d_i+1}
2. Let $t_a^{\max} = \max\{t_{d_j}\} 1 \leq j \leq n$ be the last time when the latest cause can be proven.
3. b can be proven for d_b instants starting from time $t_a^{\max} + d$, i.e., it can be proven in all the time instants between $t_a^{\max} + d - 1$ and $t_a^{\max} + d + d_b + 1$.

4 Causation

Since causation “is common in scientific thinking” and “is dominant in everyday commonsense thinking” [24], it has been widely studied by many AI approaches, mainly aiming to provide computational models of causal reasoning. In particular, many approaches have focused on the relationships between time and causation, starting, e.g., from the pioneering reified logics by McDermott [18] and Allen [1], and by the approaches by Shoham [24], Konolidge [15] and Sandewall [23], that underlined the defeasible character of causal connection and/or stressed the strict relationship between causal reasoning and non-monotonic reasoning (e.g., to deal with change).

Moreover, several approaches focused specifically on the temporal relationships between causes and effects. The problem of determining such temporal constraints is still an open one: for instance, there seems to be no agreement on whether “backward” causation (i.e., causal relations in which the effects can precede the causes) is possible or not (for a collection of accounts on causation by many of the outstanding contemporary philosophers see Sosa [25]).

While domain-dependent approaches in specific application areas have devised restrictive “ad-hoc” definitions (see, e.g., [14] in the medical field), many approaches only identified very general temporal constraint (namely, the constraint that the beginning of the effect cannot precede the beginning of the causes [24; 18; 1]). An interesting proposal has been devised by Reiger and Grinberg [22], who distinguished between different types of causation depending on the different temporal relations they impose between causes and effects. In particular, in a “One-shot” causal relation the presence of

the cause is required only momentarily in order to allow the effect to begin (e.g., to kick a ball causes the ball to move), while in a “Continuous” causal relation the presence of the cause is required in order to begin and sustain the effect (e.g., to run causes the production of lactic acid). Other researchers have followed such a line of research. For instance, within the CYC ontology, Guha and Lenat [13] have identified also “Mutually sustaining” causal relations (e.g., “poverty and illiteracy were mutually sustaining problems plaguing America all during the 1970s”), in which each bit of the cause causes a slightly later bit of the effect, and viceversa (i.e., causes and effects are co-temporal). Analogously, Terenziani and Torasso [26] have introduced also “Culminated Event Causation”, to model causal relationships where the cause is an accomplishment, whose culmination must be reached in order to enforce the effect to start (e.g., switching the light on is an activity with a goal, and its effect –having the light on– is reached with the culmination of such an activity).

The definition of rules we have proposed in Section 3 can model causation, by interpreting the antecedents of the rules as causes and the consequents as effects. Of course, the defeasible non-monotonic character of causation is naturally coped with by the underlying defeasible logic. Moreover, the definition of rules we have provided is very general, so that all the above different types of temporal constraints between causes and effects can be captured (as special cases of the general rule) by suitably restricting the values of delay and durations in the rule. For the sake of simplicity, we exemplify such a claim considering simple rules with just one antecedent (cause), i.e., rules of the form

$$a : t_a \Rightarrow^d b : t_b$$

and considering the interpretation (2) for delays (the case of interpretation (1) is analogous):

- We can model “backward” causation, by considering a negative delay between causes and effects, so that $0 > t_a + d$;
- “one-shot” causation can be modelled with $0 \leq t_a + d$ and $0 < t_b + d$;
- “continuous” causation can be modelled with $0 \leq t_a + d$ and $0 \geq t_b + d$;
- “mutually sustaining” causation can be modelled with $0 = t_a + d$, and $0 = t_b + d$;
- “culminated event causation” can be modelled simply by imposing $0 \leq d$.

5 Temporalised Defeasible Logic

As usual with non-monotonic reasoning, we have to specify 1) how to represent a knowledge base and 2) the inference mechanism used to reason with the knowledge base. The language of temporal Defeasible Logic consists of a (numerable) set of atomic propositions $Prop = \{p, q, \dots\}$, a discrete totally ordered set of instants of time $\mathcal{T} = \{t_1, t_2, \dots\}$ and the negation sign \neg .

We supplement the usual definition of literal (an atomic proposition in $Prop$ or the negation of it), with two notions of temporalised literals, one notion to indicate when a literal

holds (i.e., $p@t$, where $t \in \mathcal{T}$) and the duration of an event (i.e., $p : d$, $d \in \mathbb{N}$). Intuitively $p@t$ means that p holds at time t , and $p : d$ means that p holds for d time instants. Given a (temporalised) literal p , $\sim p$ denotes the complement of p , that is $\sim p = \neg q$ if $p = q$, $q \in Prop$, and $\sim p = q$ if $p = \neg q$ for some $q \in Prop$.

Formally a rule r consists of its *antecedent* (or *body*) $A(r)$ ($A(r)$ may be omitted if it is the empty set) which is a finite set of temporal literals of the form $p : d$, an arrow (we will use \rightarrow for strict rules, \Rightarrow for defeasible rules, and \rightsquigarrow for defeaters, and \hookrightarrow as a variable ranging over any of three types of rule), and its *consequent* (or *head*) $C(r)$ which is a temporal literal of the form $p : d$. The arrow of a rule r is labelled by a natural number d , the delay of r ; we will use $\delta(r)$ to denote the delay of r . Given a set R of rules, we denote the set of strict rules in R by R_s , the set of strict and defeasible rules in R by R_{sd} , the set of defeasible rules in R by R_d , and the set of defeaters in R by R_{df} . $R[q]$ denotes the set of rules in R with consequent $q : d$, for some d .

A temporalised defeasible theory is a structure $(\mathcal{T}, t_0, F, R, \succ)$ where: $\mathcal{T} = \{t_1, t_2, \dots\}$ is a discrete totally ordered set of instants of time; $t_0 \in \mathcal{T}$ is the reference point (or origin) of \mathcal{T} ; F is a finite set of facts, i.e., of temporalised literals of form $p@t$; R is a finite set of rules; \succ , the superiority relation, is a binary relation over R such that the transitive closure of it is acyclic.

A *conclusion* of D is a tagged literal and can have one of the following four forms:

- $+\Delta q@t$ meaning that q is definitely provable, at time t , in D (i.e., using only facts and strict rules).
- $-\Delta q@t$ meaning that we have proved that q is not definitely provable, at time t , in D .
- $+\partial q@t$ meaning that q is defeasibly provable, at time t , in D .
- $-\partial q@t$ meaning that we have proved that q is not defeasibly provable, at time t , in D .

The proof tags tell us whether it is possible to prove a conclusion at a specific time and the strength or confidence of the conclusion.

Provability is based on the concept of a *derivation* (or proof) in D . A derivation is a finite sequence $P = (P(1), \dots, P(n))$ of tagged literals satisfying the proof conditions (which correspond to inference rules for each of the kinds of conclusion). $P(1..n)$ denotes the initial part of the sequence P of length n .

Before introducing the proof conditions for the proof tags relevant to this paper we provide some auxiliary notions. From now on we will use $\#$ as a variable ranging over $\{\Delta, \partial\}$. All the remaining definitions refer to a derivation P .

Definition 1 A temporalised literal $p : d$ is $\#$ -active at t iff 1) $-\#p@t-1 \in P(1..n)$, 2) $\forall t', t-1 < t' \leq t+d$, $+\#p@t' \in P(1..n)$, and 3) $-\#p@t+d+1 \in P(1..n)$. If $d = +\infty$ then the last condition does not apply.

Intuitively that a literal $a : d$ is $\#$ -active at t means that the event a started at t and lasted for d instants. Based on the notion of $\#$ -active time we can define the activation time of

a rule. The idea is that at the activation time a rule becomes active and trigger its own conclusion.

Definition 2 Let r be a rule. A set $AT(r) \subset \mathcal{T}$ is an $\#$ -activation set of r iff

1. there is a surjective mapping ρ from $A(r)$ to $AT(r)$, and
2. $\forall a_i : d_i \in A(r)$, $a_i : d_i$ is $\#$ -active at $\rho(a_i : d_i)$.

t is a $\#$ -activation time for r iff there is a $\#$ -activation set $AT(r)$ such that $t = \max(AT(r))$.

Since there it is possible that a temporal literal $a : d$ is active at different times, then there are many activation times for a rule. The activation time for a rule is used to determine when the rule fires and when the effect of the rule (its consequent) will be produced. The formal relationship between the application time, the delay and the start of the effect is given in the following definition.

Definition 3 A rule r is $\#$ -applicable at t iff $t - \delta(r)$ is a $\#$ -application time for r .

Finally we give a definition about the potential duration of the effect of a rule.

Definition 4 A rule $r \in R_{sd}[p : d_r]$ is (t, t') -effective iff r is ∂ -applicable at t and $t' - t < d_r$.

Let us concentrate on the conditions that prevent a rules to be triggered. Again we start from temporal literals.

Definition 5 A temporal literal $a : d$ is $\#$ -passive at time t iff $\exists t' \in t \leq t < t + d + 1$ such that $-\#a@t$.

This means that a temporal literal is passive at time t if there is an instant in the interval $t, t + d$ where the literal does not hold.

Definition 6 A rule r is $\#$ -discarded at time t iff $\exists a_i : d_i \in A(r)$ such that $\forall t' < t - \delta(r)$, $a_i : d_i$ is $\#$ -passive at t .

We define now a condition to determine whether an applicable rule does not produce its effect since the derivation of the effect is blocked by a stronger rules preventing the effect.

Definition 7 A rule $r \in R_{sd}[p : d_r]$ defeats a rule $s \in R[\sim p : d_s]$ at time t iff 1) s is ∂ -applicable at t , and 2) r is (t', t) -effective, and 3) $r \succ s$.

The effect of a rule, i.e., its consequent, has a duration, so the effect can persist. The next definition captures the idea of persistence in an interval.

Definition 8 A rule $r \in R_{sd}[p : d_r]$ persists at time t iff

- r is ∂ -applicable at t' , $t - d_r < t' < t$ and $+\partial p@t' \in P(1..n)$ and
- $\forall t'', t' < t'' \leq t$, $\forall s \in R[\sim p]$ either
 - s is ∂ -discarded at t'' or
 - if s is (t''', t'') -effective, then $\exists v \in R_{sd}[p]$ such that v defeats s at t''' .

We are now ready to give the proof conditions for $+\Delta$, $-\Delta$, $+\partial$ and $-\partial$.

Proof condition for $+\Delta$

If $+\Delta p@t = P(n+1)$ then

- (1) $p@t \in F$ or
- (2) $\exists r \in R_s[p]$, r is Δ -applicable at t , or
- (3) $\exists r \in R_s[p]$, r is Δ -applicable at t' and $t - \delta(r) < t'$.

The above condition is the normal condition for definite (positive) proofs in defeasible logic, that is, monotonic derivations using forward chaining, or modus ponens. The only things to notice are that: (1) each literal $a_i : d_i$ in the antecedent must be definitely proved for exactly d_i times; (2) clause (3) is a persistence clause and allows us to derive p at time t if we have previously proved p at time t' and the distance between t and t' is less than the duration of a strict rule for p applicable at t' .

Proof condition for $-\Delta$

If $-\Delta p@t = P(n+1)$ then

- (1) $p@t \notin F$ and
- (2) $\forall r \in R_s[p]$, r is Δ -discarded at t and
- (3) $\forall r \in R_s[p]$, if r is Δ -applicable at t' , then $t' + \delta(r) < t$.

To prove that a definite conclusion is not possible we have to show that all attempts to give a definite proof of the conclusion fail. In particular we have to check that all strict rules for $\sim p$ will not produce an effect at t .

Proof condition for $+\partial$

If $+\partial p@t = P(n+1)$ then

- (1) $+\Delta p@t \in P(1..n)$ or
- (2) $(1) -\Delta \sim p@t \in P(1..n)$ and
 - (2) $\exists r \in R_{sd}[p]$ such that r either r persists at t or r is ∂ -applicable at t , and
- (3) $\forall s \in R[\sim p]$ either
 - (1) s is ∂ -discarded at t or
 - (2) if s is (t', t) -effective, then $\exists v \in R_{sd}[p]$, such that v defeats s at t' .

Clauses (1) and (2.1) are the usual clauses of standard defeasible logic only relativised to a time instant. In particular clause (1) allows us to weaken a definite conclusion to a defeasible conclusions. Clause (2.1) ensures consistency of the logic, i.e., a stronger proof for the opposite of what we want to prove is not possible. Clause (2.2) requires that at least one rule for p can indeed produce p at time t . Thus we have to subtract the delay from t to see whether the last cause of the rule started at $t - d$. Clause (3.1) is intended to check that each rule for the negation of the conclusion we want to prove is either discarded at t . i.e., one of the causes of this rule has been proved not to hold for the right duration. Otherwise, an applicable rule for $\sim p$ producing the effect at t must be blocked by a stronger rule for p . The stronger rule should block the beginning of the effect $\sim p$.

Let us consider a theory where $F = \{a@0, b@5, c@5\}$ and with the rules

$$r_1 : a : 1 \Rightarrow^{10} d : 10, r_2 : b : 1 \Rightarrow^7 \neg d : 5, r_3 : c : 1 \Rightarrow^8 d : 5$$

where $r_3 \succ r_2$. Based on this theory we can derive $+\partial d@10$. r_1 is ∂ -applicable and r_2 is ∂ -discarded. At time 11 we can derive d since r_1 persists here. At time 12 r_2 is no longer discarded, so r_2 interrupts the effect of r_1 . Since we have

no means to determine which reason prevails we derive both $-\partial d@12$ and $-\partial \neg d@12$. At 13 rule r_3 becomes active and defeats r_2 , thus from 13 to 18 we can derive d . However, r_3 terminates its effect at 18, and r_1 , whose duration would ensure that d persists till 20, cannot be used to derive d since the effect of r_1 was terminated by r_2 at time 12.

Proof condition for $-\partial$

If $-\partial p@t = P(n+1)$ then

- (1) $-\Delta p@t \in P(1..n)$ and
- (2) $(1) +\Delta \sim p@t \notin P(1..n)$ or
 - (2) $\forall r \in R_{sd}[p]$ r does not persist at t and is ∂ -discarded at t , or
- (3) $\exists s \in R[\sim p]$ such that
 - (1) s is (t', t) -effective, and
 - (2) $\forall v \in R_{sd}[p]$, v does not defeat s at t' .

To prove that a defeasible conclusion is not possible we have to show that all attempts to give a defeasible proof of the conclusion fail. In particular we have to check that all rules for $\sim p$ will not produce an effect at t , and if they were to produce it there are some active rules for the negation not weaker than rules for the conclusion and the rules are effective at the appropriate instants.

Maher [16] proved that for standard defeasible logic the extension of a defeasible theory can be computed in time linear to the size of the theory (where the size of the theory is in function of the number of symbols and rules in it). [10] extended the result to more expressive defeasible logic. For the temporal extension presented here we can preserve the complexity results for theories without backward causation.

Theorem 1 *Let D be a temporalised defeasible theory without backward causation. Then the extension of D from t_0 and $t \in \mathcal{T}$ (i.e., the set of all consequences based on the signature of D derivable from t_0 and t) can be computed in time linear to the size of theory. i.e., $O(|Prop| \times |R| \times t)$.*

Proof (sketch). The idea is to consider snapshots of the theory at every single instant from t_0 and t . The rules are linearly reduced to the language of [11]. In particular temporal literals in the antecedent are understood as a set of temporal literals with timestamps and literal with infinite duration are dealt with persistent rules. After these steps we can apply the technique of [10] to compute in linear time the extension of the resulting equivalent theory.

It is important to notice that, although the rules contain temporal constraints that are related to STP (Simple Temporal Problem [5]), in the environment we consider in this paper, no “classical” form of temporal reasoning (in the sense of temporal constraint propagation [5]) is required. For example, let us consider the following simple set of rules, involving delays between instantaneous events:

$$\{a : 1 \Rightarrow^{10} b : 1, b : 1 \Rightarrow^{10} c : 1, a : 1 \Rightarrow^{10} c : 1\}$$

In a “classical” temporal constraint environment, such as in the STP framework, the rules in the set would be interpreted as a *conjunction*, so that constraint propagation could be used in order to infer from the first two dependencies that (in case the interpretation (2) is chosen for the delay) the delay between a and c is 21, so that the whole set of constraints is not consistent.

On the other hand, within the DL environment we are working in this paper, the above set is simply interpreted as a set of *alternative rules*. In such a context, one can consistently say that c can be caused by a through the first two rules (where the delay between a and c is 21), or directly through the third rule (and, in such a case, the delay is 10).

6 Conclusion and Future Extensions

In this paper, we have proposed a temporal extension of defeasible logic in order to deal with temporalised durative facts, and with (possible) delays between antecedents and consequences of the rules. The extension we propose

- increases the expressiveness of the logic, as needed to deal with temporal phenomena in several application domains, and, in particular, with different types of causal relationships;
- maintains the computational properties of the original logic. Indeed, the temporal extension has been carefully shaped in order to maintain (in case backward causation rules are not used) the linear complexity of the original logic.

As future work, we aim at analysing the complexity of the extended logic, in case also backward causation rules (i.e., rules in which the consequent can start before the antecedent) are taken into account. The addition of backward rules makes complexity significantly harder, since it requires to navigate the timeline also backward. Our conjecture is that, in such a case, the complexity becomes quadratic on the overall span of time determined by the period of repetition of all the cycles of the rules in the theory. Moreover, we would like to investigate the possibility of further extending the expressiveness of the temporal part of the rules with the introduction of STP-like ([5]) temporal constraints between the propositions in the rules. The core of this extension will be an accurate analysis of the trade-off between expressiveness and computational complexity, since we aim at retaining as much as possible the linear complexity property of the original logic. Finally, we also would like to analyse the possibility of distinguishing in the logic among different classes of propositions (e.g., stative propositions such as being asleep vs propositions denoting events with a culmination, such as building a house) on the basis of their temporal properties, as suggested by many approaches in the linguistic literature (see, e.g., [19; 27] and also, e.g., in [26]).

References

- [1] J.F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23:123–154, 1984.
- [2] Grigoris Antoniou. Nonmonotonic rule system on top of ontology layer. In I. Horrocks and J. Hendler, editors, *ISWC 2002*, number 2432 in LNCS, pages 394–398, Berlin, 2002. Springer.
- [3] Grigoris Antoniou, David Billington, Guido Governatori, and Michael J. Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, 2(2):255–287, 2001.
- [4] Grigoris Antoniou, David Billington, Guido Governatori, Michael J. Maher, and Andrew Rock. A family of defeasible reasoning logics and its implementation. In Werner Horn, editor, *ECAI 2000. Proceedings of the 14th European Conference on Artificial Intelligence*, pages 459–463, Amsterdam, 2000. IOS Press.
- [5] R. Dechter, I. Meiri, and J. Pearl. Temporal constraint networks. *Artificial Intelligence*, 49:61–95, 1991.
- [6] Guido Governatori. Representing business contracts in RuleML. *International Journal of Cooperative Information Systems*, 14(2-3):181–216, 2005.
- [7] Guido Governatori, Marlon Dumas, Arthur H.M. ter Hofstede, and Phillipa Oaks. A formal approach to protocols and strategies for (legal) negotiation. In *Proc. 8th International Conference on Artificial Intelligence and Law (ICAIL01)*, pages 168–177. ACM Press, 2001.
- [8] Guido Governatori, Vineet Padmanabhan, and Rotolo. Antonino. Rule-based agents in temporalised defeasible logic. In Qiang Yang and Geoff Webb, editors, *Ninth Pacific Rim International Conference on Artificial Intelligence*, number 4099 in LNAI, pages 31–40, Berlin, 7–11 August 2006. Springer.
- [9] Guido Governatori and Antonino Rotolo. Defeasible logic: Agency, intention and obligation. In Alessio Lomuscio and Donald Nute, editors, *Deontic Logic in Computer Science*, volume 3065 of LNAI, pages 114–128, Berlin, 2004. Springer.
- [10] Guido Governatori, Antonino Rotolo, and Vineet Padmanabhan. The cost of social agents. In Peter Stone and Gerhard Weiss, editors, *5th International Conference on Autonomous Agents and Multi-Agent Systems*, pages 513–520, New York, 2006. ACM Press.
- [11] Guido Governatori, Antonino Rotolo, and Giovanni Sartor. Temporalised normative positions in defeasible logic. In *10th International Conference on Artificial Intelligence and Law (ICAIL05)*, pages 25–34. ACM Press, 2005.
- [12] Benjamin N. Grosz. Representing e-commerce rules via situated courteous logic programs in RuleML. *Electronic Commerce Research and Applications*, 3(1):2–20, 2004.
- [13] R.V. Guha and D.B. Lenat. *Building Large Knowledge Based Systems*. Addison Wesley, 1991.
- [14] I Hamlet and J. Hunter. Representation of time in medical expert systems. *LNMI*, 33:112–119, 1987.
- [15] K Konolidge. Using default and causal reasoning in diagnosis. *Annals of Mathematics and Artificial Intelligence*, 11:97–135, 1994.
- [16] Michael J. Maher. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming*, 1(6):691–711, 2001.
- [17] Michael J. Maher, Andrew Rock, Grigoris Antoniou, David Billington, and Tristan Miller. Efficient defeasible reasoning systems. *International Journal of Artificial Intelligence Tools*, 10(4):483–501, 2001.

- [18] D. McDermott. A temporal logic for reasoning about processes and plans. *Cognitive Science*, 6:101–105, 1982.
- [19] A. Mourelatos. Events, processes and states. *Linguistic and Philosophy*, 2:415–434, 1978.
- [20] Donald Nute. Defeasible logic. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, pages 353–395. Oxford University Press, 1994.
- [21] Henry Prakken and Giovanni Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4:331–368, 1996.
- [22] C. Rieger and M Grinberg. The declarative representation and simulation of causality in physical mechanisms. In *Proc. IJCAI’77*, pages 250–256, 1977.
- [23] E. Sandewall. *Features and Fluents*. Oxford Scientific Publications, 1994.
- [24] Yoav Shoham. *Reasoning About Change*. MIT Press, Cambridge, MA., 1987.
- [25] E Sosa. *Causation and Conditionals*. Oxford university Press, 1975.
- [26] P. Terenziani and P. Torasso. Time, action-types and causation: an integrated analysis. *Computational Intelligence*, 11(3):529–552, 1995.
- [27] Z Vendler. Verbs and times. *Linguistic in Philosophy*, pages 97–121, 1967.