Ontology-based inference for causal explanation

Ph. Besnard[1], M.-O. Cordier[2], and Y. Moinard[2]

[1] IRIT, CNRS, Université Paul Sabatier
118 route de Narbonne, 31062 Toulouse cedex, France
besnard@irit.fr
[2] IRISA, INRIA, Université Rennes I
Campus de Beaulieu, 35042 Rennes cedex, France
{cordier,moinard}@irisa.fr

Abstract. We define an inference system to capture explanations based on causal statements, using an ontology in the form of an *IS-A* hierarchy. We first introduce a simple logical language which makes it possible to express that a fact causes another fact and that a fact explains another fact. We present a set of formal inference patterns from causal statements to explanation statements. These patterns exhibit ontological premises that are argued to be essential in deducing explanation statements. We provide an inference system that captures the patterns discussed.

1 Introduction

We are aiming at a logical formalization of explanations from causal statements. For example, it is usually admitted that fire is an explanation for smoke, on the grounds that fire causes smoke. In other words, fire causes smoke is a premise from which it can be inferred that fire is an explanation for smoke. In this particular example, concluding from cause to explanation is immediate but such is not always the case, far from it. In general, the reasoning steps leading from cause to explanation are not so trivial:

Example. *We consider two causal statements:*

(i) Any ship that is about to sink causes her crew to launch some red rocket(s)

(ii) In France, July the 14^{th} *causes colourful rockets to be launched (fireworks)*

So, if the place is a coastal city in France, on July the 14^{th}*, then red rockets being launched could be explained either by some ship(s) sinking or by the French national day.*

In this example, it is needed to acknowledge the fact that a red rocket is a kind of (colourful) rocket in order to get the second explanation, which makes sense.

Example (con'd). *Suppose that we now add the following statement:*

(i) Seeing a red rocket being launched triggers a rescue process

Now, a possible explanation for the triggering of the rescue process, as happens in practice, is that we are on July the 14^{th} *in a coastal city in France.*

In this paper, we define a dedicated inference system to capture explanations based on causal statements and stress that the rôle of ontology-based information is essential. In the second section, we introduce the logical language that we propose to use. In the third section, we define the set of patterns dedicated to inferring explanations from causal statements and ontological information. In the fourth section, we give the formal inference system allowing us to derive explanations from a given set of formulas. We conclude by discussing the rôle of ontology and by giving some perspectives.

2 Language

We distinguish various types of statements in our formal system:

C: A theory expressing causal statements. E.g., $On(alarm)$ *causes* $Heard(bell)$.
O: An ontology listing entities, in the form of an IS-A hierarchy which defines classes and sub-classes. E.g., *loud_bell is a bell* and *soft_bell is a bell*.
W: A classical first order theory expressing truths (i.e., incompatible facts, co-occurring facts, ...). E.g., $Heard(soft\_bell) \to \neg Heard(loud\_bell)$.

We assume a logical language whose alphabet consists of constants $a, b, c, \ldots$ and unary predicates such as $P, \ldots$ Intuitively, constants denote objects or entities, including classes and sub-classes to be found in the IS-A hierarchy. The predicates (unary for simplicity, they could be n-ary in theory) are used to express facts or events on these objects or entities as $Heard(soft\_bell)$ or $Own(blue\_car)$. The causal statements express causal relations between facts or events expressed by these predicates as in $On(alarm)$ *causes* $Heard(bell)$.

Moreover, we assume that these unary predicates "inherit upwards" through the IS-A hierarchy in the following sense: If b IS-A c then $P(b)$ entails $P(c)$. Consider owning as an example, together with small car and car. Of course, a small car is a car and "I own a small car" entails "'I own a car". This property happens to be fundamental when designing inference patterns (next section) as it allows us to apply inheritance properties between entities to facts and events on these entities.. It also means that we restrict ourselves to those predicates that "inherit upwards", which precludes for instance the predicate $Dislike$ as $Dislike(small\_car)$ does not imply $Dislike(car)$. In the following, we say that $Heard(soft\_bell)$ is a specialization of $Heard(bell)$ and that $Heard(bell)$ is a generalization of $Heard(soft\_bell)$.

The formal system we introduce below is meant to infer (this inference will be noted $\vdash_C$), from such premises $C \cup O \cup W$, formulas denoting explanations.

In the sequel, $\alpha, \beta, \ldots$ denote the so-called sentential atoms (i.e., ground atomic formulas) and $\Phi, \Psi, \ldots$ denote sets thereof.

Atoms

1. *Sentential atoms*: $\alpha, \beta, \ldots$ (Ground atomic formulas)
2. *Causal atoms*: α *causes* β.

3. *Ontological atoms*: $b \rightarrow_{IS-A} c$.
4. *Explanation atoms*: α *explains* β *because_possible* Φ.

An ontological atom reads: b is a c.
An explanation atom reads: α *is an explanation for* β *because* Φ *is possible.*

Notation: α *explains* β *bec_poss* Φ *abbreviates* α *explains* β *because_possible* Φ.

Formulas

1. *Sentential formulas*: Boolean combinations of sentential atoms.
2. *Causal formulas*: Boolean combinations of causal atoms and sentential atoms.

The premises of the inference $\vdash_C$, namely $C \cup O \cup W$, consist of sentential formulas as well as causal formulas and ontological *atoms* (no ontological formula). Notice that explanation atoms cannot occur in the premises.

The properties of causal and ontological formulas we consider are as follows.

1. **Properties of the causal operator**
   (a) *Entailing the conditional*: If α *causes* β, then $\alpha \rightarrow \beta$.
2. **Properties of the ontological operator**
   (a) *Upward inheritance*: If $b \rightarrow_{IS-A} c$, then $\alpha[b] \rightarrow \alpha[c]$[3].
   (b) *Transitivity*: If $a \rightarrow_{IS-A} b$ and $b \rightarrow_{IS-A} c$, then $a \rightarrow_{IS-A} c$.
   (c) *Reflexivity*: $c \rightarrow_{IS-A} c$.

Reflexivity is unconventional a property for an IS-A hierarchy. It is included here because it helps keeping the number of inference schemes low (see later).

W is supposed to include (whether explicitly or via inference) all conditionals induced by the ontology O. For example, if *loud_bell* $\rightarrow_{IS-A}$ *bell* is in O then $Heard(loud\_bell) \rightarrow Heard(bell)$ is in W. Similarly, W is supposed to include all conditionals induced by the causal statements in C. For example, if $On(alarm)$ *causes* $Heard(bell)$ is in C, then $On(alarm) \rightarrow Heard(bell)$ is in W.

3 Patterns for inferring explanations

A set of patterns is proposed to infer explanations from premises $C \cup O \cup W$.

3.1 The base case

A basic idea is that what causes an effect can always be suggested as an explanation when the effect happens to be the case:

$$\text{If} \quad \begin{pmatrix} \alpha \textit{ causes } \beta \\ \text{and} \\ W \not\models \neg\alpha \end{pmatrix} \quad \text{then} \quad \alpha \textit{ explains } \beta \textit{ because\_possible } \{\alpha\}$$

[3] $\alpha[b]$ denotes the atomic formula α and its (only) argument b.

Example. *Consider a causal model such that* $W \not\vdash \neg On(alarm)$ *and* O *is empty whereas*

$$C = \{On(alarm)\ causes\ Heard(bell)\}$$

Then, the atom

$$On(alarm)\ \ explains\ \ Heard(bell)\ \ because\_possible\ \ \{On(alarm)\}$$

is inferred. That is, $On(alarm)$ *is an explanation for* $Heard(bell)$*.*

By the way, "is an explanation" must be understood as provisional. Inferring that $On(alarm)$ is an explanation for $Heard(bell)$ is a tentative conclusion: Should $On(alarm)$ be ruled out, e.g., $\neg On(alarm) \in W$, then $On(alarm)$ is not an explanation for $Heard(bell)$.
Formally, with $Form = On(alarm)\ explains\ Heard(bell)\ bec\_poss\ \{On(alarm)\}$:
$C \cup O \cup W \vdash_C Form$; $\qquad C \cup O \cup W \cup \{\neg On(alarm)\} \not\vdash_C Form$

3.2 Wandering the IS-A hierarchy: Going upward

What causes an effect can be suggested as an explanation for any consistent ontological generalization of the effect:

$$\text{If} \quad \begin{pmatrix} \alpha\ causes\ \beta[b] \\ \text{and} \\ b \rightarrow_{IS-A} c \\ \text{and} \\ W \not\models \neg\alpha \end{pmatrix} \quad \text{then} \quad \alpha\ \ explains\ \ \beta[c]\ \ because\_possible\ \ \{\alpha\}$$

Example. $C = \{On(alarm)\ causes\ Heard(bell)\}$ *and* $O = \{bell \rightarrow_{IS-A} noise\}$*.* W *contains no statement apart from those induced by* C *and* O*, that is:*

$$W = \{On(alarm) \rightarrow Heard(bell), Heard(bell) \rightarrow Heard(noise)\}$$

Inasmuch as noise could be bell, $On(alarm)$ *then counts as an explanation for* $Heard(noise)$*.*

$$C \cup O \cup W \vdash_C On(alarm)\ explains\ Heard(noise)\ bec\_poss\ \{On(alarm)\}$$

Again, it would take $On(alarm)$ *to be ruled out for the inference to be prevented.*

Example. $C = \{On(alarm)\ causes\ Heard(bell)\}$

$$O = \{bell \rightarrow_{IS-A} noise, hooter \rightarrow_{IS-A} noise\}$$

W states that a hooter is heard (and that $Heard(bell)$ is not $Heard(hooter)$) and additionally expresses the conditionals induced by C and O, that is:

$$W = \left\{ \begin{array}{c} Heard(hooter) \\ \neg(Heard(bell) \leftrightarrow Heard(hooter)) \\ On(alarm) \rightarrow Heard(bell) \\ Heard(hooter) \rightarrow Heard(noise) \\ Heard(bell) \rightarrow Heard(noise) \end{array} \right\}$$

Even taking into account the fact that bell is an instance of noise, it cannot be inferred that $On(alarm)$ is an explanation for $Heard(noise)$. The inference fails because it would need noise to be of the bell kind (which is false, cf hooter). Technically, the inference fails because $W \vdash \neg On(alarm)$.

The next example illustrates why resorting to ontological information is essential when attempting to infer explanations: the patterns in sections 3.2-3.3 extend the base case for explanations to *ontology-based* consequences, not any consequences.

Example. *Rain makes me growl. Trivially, I growl only if I am alive. However, rain cannot be taken as an explanation for the fact that I am alive.*

$$C = \{Rain\ causes\ I\_growl\}$$
$$W = \{I\_growl \rightarrow I\_am\_alive\}$$

$C \cup O \cup W \not\vdash_C$ *Rain explains I_am_alive bec_poss* $\{Rain\}$

3.3 Wandering the IS-A hierarchy: Going downward

What causes an effect can presumably be suggested as an explanation when the effect takes place in one of its specialized forms:

$$\text{If} \quad \begin{pmatrix} \alpha\ causes\ \beta[c] \\ \text{and} \\ b \rightarrow_{IS-A} c \\ \text{and} \\ W \not\models \neg(\alpha \wedge \beta[b]) \end{pmatrix} \quad \text{then} \quad \alpha\ \ explains\ \ \beta[b]\ \ because\_possible\ \ \{\alpha, \beta[b]\}$$

Example. *Consider a causal model with C and O as follows:*

$$C = \{On(alarm)\ causes\ Heard(bell)\} \quad and \quad O = \left\{ \begin{array}{l} loud\_bell \rightarrow_{IS-A} bell \\ soft\_bell \rightarrow_{IS-A} bell \end{array} \right\}$$

O means that loud_bell is more precise than bell. Since $On(alarm)$ is an explanation for $Heard(bell)$, it also is an explanation for $Heard(loud\_bell)$ and similarly $Heard(soft\_bell)$. This holds inasmuch as there is no statement to the contrary:

The latter inference would not be drawn if for instance $\neg Heard(soft\_bell)$ *or* $\neg(Heard(soft\_bell) \rightarrow On(alarm))$ *were in* W*. Formally, with* $Form(loud) =$ $On(alarm)$ *explains* $Heard(loud\_bell)$ *bec_poss* $\{On(alarm), Heard(loud\_bell)\}$ *and* $Form(soft) =$
$On(alarm)$ *explains* $Heard(soft\_bell)$ *bec_poss* $\{On(alarm), Heard(soft\_bell)\}$*:*
$C \cup O \cup W \vdash_C Form(loud)$, $C \cup O \cup W \vdash_C Form(soft)$*, and*
$C \cup O \cup W \cup \{\neg(Heard(soft\_bell) \rightarrow On(alarm))\} \not\vdash_C Form(soft)$

3.4 Transitivity of explanations

We make no assumption as to whether the causal operator is transitive (from α *causes* β and β *causes* γ does α *causes* γ follow?). However, we do regard inference of explanations as transitive which, in the simplest case, means that if α explains β and β explains γ then α explains γ.

The general pattern for transitivity of explanations takes two causal statements, α *causes* β and β' *causes* γ where β and β' are ontologically related, as premises in order to infer that α is an explanation for γ.

In the first form of transitivity, β' is inherited from β by going upward in the *IS-A* hierarchy.

$$\text{If} \quad \begin{pmatrix} \alpha \textit{ causes } \beta[b], \\ \beta[c] \textit{ causes } \gamma, \\ b \rightarrow_{IS-A} c, \\ \text{and} \\ W \not\models \neg\alpha \end{pmatrix} \quad \text{then} \quad \alpha \textit{ explains } \gamma \textit{ because\_possible } \{\alpha\}$$

Example. *Sunshine makes me happy. Being happy is why I sing. Therefore, sunshine is a plausible explanation for the case that I am singing.*

$$C = \left\{ \begin{array}{c} Sunshine \textit{ causes } I\_am\_happy \\ I\_am\_happy \textit{ causes } I\_am\_singing \end{array} \right\}$$

$$W = \left\{ \begin{array}{c} Sunshine \rightarrow I\_am\_happy \\ I\_am\_happy \rightarrow I\_am\_singing \end{array} \right\}$$

So, we get:

$$C \cup O \cup W \vdash_C Sunshine \textit{ explains } I\_am\_singing \textit{ bec\_poss } \{Sunshine\}.$$

The above example exhibits transitivity of explanations for the simplest case that $\beta = \beta'$ in the pattern α *causes* β and β' *causes* γ entail α *causes* γ (trivially, if $\beta = \beta'$ then β and β' are ontologically related). Technically, $\beta = \beta'$ is obtained by applying reflexivity in the ontology. This is one illustration that using reflexivity in the ontology relieves us from the burden of tailoring definitions to capture formal degenerate cases.

Example. *Let* $O = \{bell \rightarrow_{IS-A} noise\}$ *and*

$$C = \left\{ \begin{array}{c} On(alarm)\ causes\ Heard(bell) \\ Heard(noise)\ causes\ Disturbance \end{array} \right\}$$

W *states the facts induced by* C *and* O*, that is:*

$$W = \left\{ \begin{array}{c} On(alarm) \rightarrow Heard(bell) \\ Heard(noise) \rightarrow Disturbance \\ Heard(bell) \rightarrow Heard(noise) \end{array} \right\}$$

So, we get:

$$C \cup O \cup W \vdash_C On(alarm)\ explains\ Disturbance\ bec\_poss\ \{On(alarm)\}.$$

In the second form of transitivity, β' is inherited from β by going downward in the IS-A hierarchy.

$$\text{If} \quad \left(\begin{array}{l} \alpha\ causes\ \beta[c], \\ \beta[b]\ causes\ \gamma, \\ b \rightarrow_{IS-A} c, \\ \text{and} \\ W \not\models \neg(\alpha \wedge \beta[b]) \end{array} \right) \quad \text{then} \quad \alpha\ \ explains\ \ \gamma\ \ because\_possible\ \ \{\alpha, \beta[b]\}$$

Example. $O = \{loud\_bell \rightarrow_{IS-A} bell\}$

$$C = \left\{ \begin{array}{c} On(alarm)\ causes\ Heard(bell) \\ Heard(loud\_bell)\ causes\ Deafening \end{array} \right\}$$

$$W = \left\{ \begin{array}{c} Heard(loud\_bell) \rightarrow Heard(bell) \\ On(alarm) \rightarrow Heard(bell) \\ Heard(loud\_bell) \rightarrow Deafening \end{array} \right\}$$

$On(alarm)$ *does not cause* $Heard(loud\_bell)$ *(neither does it cause* $Deafening$*), but it is an explanation for* $Heard(loud\_bell)$ *by virtue of the upward scheme. Due to the base case,* $Heard(loud\_bell)$ *is in turn an explanation for* $Deafening$*. In fact,* $On(alarm)$ *is an explanation for* $Deafening$ *by virtue of transitivity.*

Considering a causal operator which is transitive would give the same explanations but is obviously more restrictive as we may not want to endorse an account of causality which is transitive. Moreover, transitivity for explanations not only seems right in itself but it also means that our model of explanations can be plugged with any causal system whether transitive or not.

3.5 Explanation provisos and their simplifications

Explanation atoms are written α *explains* β *because_possible* Φ

as the definition is intended to make the atom true just in case it is successfully checked that the proviso is possible: An explanation atom is not to be interpreted as a kind of conditional statement. Indeed, we do not write *"if_possible"*. The argument in *"because_possible"* gathers those conditions that must be possible together if α is to explain β (there can be others: α can also be an explanation of β wrt other arguments in *"because_possible"*).

Using $\bigwedge\Phi$ to denote the conjunction of the formulas in the set Φ, the following scheme amounts to simplifying the proviso attached to an explanation atom.

$$\text{If } \begin{pmatrix} W \models \bigwedge\Phi \to \bigvee_{i=1}^{n} \bigwedge\Phi_i \quad \text{and} \\ \text{for all } i \in \{1,\cdots,n\} \alpha \textit{ explains } \beta \textit{ because\_possible } (\Phi_i \cup \Phi) \end{pmatrix}$$
$$\text{then } \quad \alpha \textit{ explains } \beta \textit{ because\_possible } \Phi$$

4 A formal system for inferring explanations

The above ideas are embedded in a short proof system extending classical logic:

1. ***Causal formulas***
   (a) $(\alpha \textit{ causes } \beta) \to (\alpha \to \beta)$

2. ***Ontological atoms***
   (a) If $b \to_{IS-A} c$ then $\alpha[b] \to \alpha[c]$
   (b) If $a \to_{IS-A} b$ and $b \to_{IS-A} c$ then $a \to_{IS-A} c$
   (c) $c \to_{IS-A} c$

3. ***Explanation atoms***
   (a) $\text{If } \begin{pmatrix} \alpha \textit{ causes } \beta[b], \\ a \to_{IS-A} b, \qquad a \to_{IS-A} c \\ W \not\models \neg(\alpha \wedge \beta[a]) \end{pmatrix}$
   then α *explains* $\beta[c]$ *because_possible* $\{\alpha, \beta[a]\}$

   (b) $\text{If } \begin{pmatrix} \alpha \textit{ explains } \beta \textit{ because\_possible } \Phi \\ \beta \textit{ explains } \gamma \textit{ because\_possible } \Psi \end{pmatrix}$
   then α *explains* γ *because_possible* $(\Phi \cup \Psi)$

   (c) $\text{If } \begin{pmatrix} W \models \bigwedge\Phi \to \bigvee_{i=1}^{n} \bigwedge\Phi_i \qquad \text{and} \\ \text{for all } i \in \{1,\cdots,n\}\ \alpha \textit{ explains } \beta \textit{ because\_possible } (\Phi_i \cup \Phi) \end{pmatrix}$
   then α *explains* β *because_possible* Φ

These schemes allow us to obtain the inference patterns described in the previous section. E.g., the base case for explanation is obtained by combining (2c) with (3a) (yielding another illustration of reflexivity in the ontology relieving us from the burden of introducing further formal material) prior to simplifying by means of (3c). Analogously, the upward case is obtained by applying (2c) upon (3a) before using (3c).

A more substantial application is:

$$C = \left\{ \begin{array}{l} \alpha \textit{ causes } \beta[b] \\ \beta[c] \textit{ causes } \gamma \end{array} \right\}$$

$$O = \{b \rightarrow_{IS-A} c\}$$

$$W = \left\{ \begin{array}{c} \alpha \rightarrow \beta[b] \\ \beta[b] \rightarrow \beta[c] \\ \beta[c] \rightarrow \gamma \end{array} \right\}$$

The first form of transitivity in Section 3.4 requires that we infer:

$$\alpha \textit{ explains } \gamma \textit{ because\_possible } \{\alpha\}$$

Let us proceed step by step:

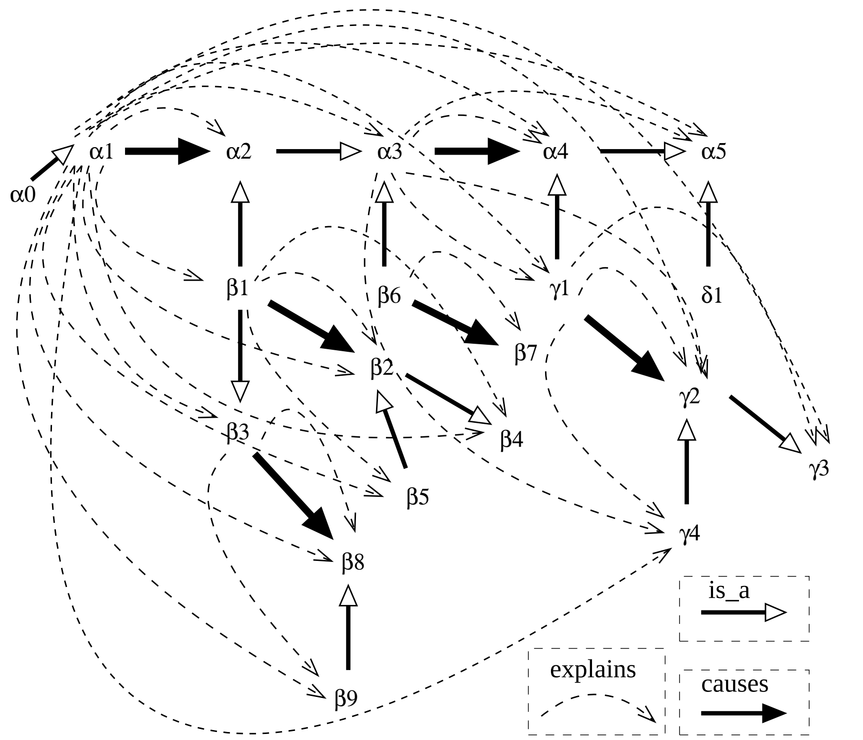
$$\begin{array}{rl} \alpha \textit{ explains } \beta[c] \textit{ because\_possible } \{\alpha\} & \text{by (3a) as upward case} \\ \beta[c] \textit{ explains } \gamma \textit{ because\_possible } \{\beta[c]\} & \text{by (3a) as base case} \\ \alpha \textit{ explains } \gamma \textit{ because\_possible } \{\alpha, \beta[c]\} & \text{by (3b)} \\ \alpha \textit{ explains } \gamma \textit{ because\_possible } \{\alpha\} & \text{by (3c) simplifying the proviso} \end{array}$$

5 A generic diagram

Below an abstract diagram is depicted that summarizes many patterns of inferred explanations from various cases of causal statements and $\rightarrow_{IS-A}$ links.

In this example, for each pair of symbols $(\sigma 1, \sigma 2)$, there is only one "explanation path". E.g., we get α_1 *explains* γ_4 *because_possible* $\{\alpha_1, \gamma_1, \gamma_4\}$, through α_3 and γ_1. Indeed, from α_1 *explains* α_3 *because_possible* $\{\alpha_1\}$ as well as α_3 *explains* γ_1 *because_possible* $\{\alpha_3, \gamma_1\}$ and
γ_1 *explains* γ_4 *because_possible* $\{\gamma_1, \gamma_4\}$, we obtain
α_1 *explains* γ_4 *because_possible* $\{\alpha_1, \alpha_3, \gamma_1, \gamma_4\}$ (using transitivity twice).
Lastly, we simplify the condition set by virtue of $W \vdash \alpha_1 \rightarrow \alpha_3$.

In other examples, various "explanation paths" exist. It suffices that the inference pattern 3a can be applied with more than one "a", or that transitivity (3b) can be applied with more than one β. We have implemented a program in DLV [8] (an implementation of the Answer Set Programming formalism [1]) that takes only a few seconds to give all the results $s1$ *explains* $s2$ *bec_poss* Φ, for all examples of this kind, including the case of different explanation paths (less than one second for the diagram depicted below).



6 Back to ontology

The explanation inferences that we obtain follow the patterns presented before. The inference pattern (3a) is important, and, in it, the direction of the $\rightarrow_{IS-A}$ links $a \rightarrow_{IS-A} b$ and $a \rightarrow_{IS-A} c$ is important: Unexpected conclusions would ensue if other directions, e.g., $b \rightarrow_{IS-A} a$ and $c \rightarrow_{IS-A} a$ were allowed.

We have considered ontological information in the most common form, as a $\rightarrow_{IS-A}$ hierarchy. A closer look at the process of inferring explanations reveals that $\rightarrow_{IS-A}$ links serve as a means to get another kind of ontological information by taking advantage of the property of the unary predicates to "inherit upwards". The basis for an explanation are causal statements, and these apply to facts or events [9]. Indeed, the ontological information which is eventually used in the process of inferring explanations is about facts or events, through the "inherit upwards" property. We eventually resort to an ontology over events.

Example. *Getting cold usually causes Mary to become active. I see Mary jogging. So, Mary getting cold might be taken as an explanation for her jogging. This holds on condition that the weather is possibly cold, otherwise the inference fails to ensue: In the presence of the information that the weather is warm, Mary getting cold is inconsistent.*

$$C = \{Mary\_was\_getting\_cold \ causes \ Mary\_is\_moving\_up\}$$

$$O = \{Mary\_is\_jogging \rightarrow_{IS-A} Mary\_is\_moving\_up\}$$

$$W = \left\{ \begin{array}{c} Mary\_was\_getting\_cold \rightarrow Mary\_is\_moving\_up \\ Mary\_is\_jogging \rightarrow Mary\_is\_moving\_up \end{array} \right\}$$

If allowing such an ontology over events as in O, we could use an extended version of our proof system to infer the atom

$Mary\_was\_getting\_cold$
explains $Mary\_is\_jogging$ *bec_poss* $\{Mary\_was\_getting\_cold\}$.

That is to say, $Mary\_was\_getting\_cold$ *would be inferred as an explanation for* $Mary\_is\_jogging$. *Also, the inference would break down if for example both* $Warm\_weather$ *and* $Mary\_was\_getting\_cold \rightarrow \neg Warm\_weather$ *were in* W.

7 Conclusion

We have provided a logical framework allowing predictive and abductive reasoning from *causal* information. Indeed, our formalism allows to express causal information in a direct way. Then, we deduce so-called *explanation atoms* which capture what might explain what, in view of some given information. We have resorted to *ontological* information. Not only is it generally useful, it is key in generating sensible explanations from causal statements.

In our approach, the user provides a list of *ontological atoms* $a \rightarrow_{IS-A} b$ intended to mean that object a "is_a" b. The basic terms, denoted α, β are then concrete atoms built with unary predicates, such as $P(a)$. The user also provides causal information in an intuitive form, as causal atoms α *causes* β (which can occur in more complex formulas).

This makes formalization fairly short and natural. The ontology provided is used in various patterns of inference for explanations. In our approach, such information is rather easy to express, or to obtain in practice due to existing ontologies and ontological languages. If we were in a purely propositional setting, the user should write $Own\_small\_car \rightarrow_{IS-A} Own\_car$, $Own\_big\_car \rightarrow_{IS-A} Own\_car$, and, when necessary $Heard\_small\_car \rightarrow_{IS-A} Heard\_car$, $Heard\_big\_car \rightarrow_{IS-A} Heard\_car$, and so on. This would be cumbersome, and error prone. This contrasts with our setting, which, moreover, is "essentially propositional" in that, for what concerns the causal atoms, it is as if $Own(small\_car)$ were a propositional symbol Own_small_car, while, for what concerns the ontology, we really use the fact that $Heard$ and Own are predicates.

As always with knowledge representation, some care must be exercized as to the vocabulary used. E.g., minimization formalisms such as circumscription or predicate completion (see logic programming and its "offspring" answer set programming) require to distinguish between "positive notions" (to be minimized) from negative notions: Writing Fly vs. *not* Fly yields a different behavior than writing *not* $Unfly$ and $Unfly$. Here, we have a similar situation, since the inference patterns would not work properly if we were to use predicates that do not "inherit upwards" with respect to the ontology provided: We should not expect to infer $Dislike\_small\_car \rightarrow_{IS-A} Dislike\_car$, since *Dislike* obviously fails to "inherit upwards" when the ontology contains $small\_car \rightarrow_{IS-A} car$.

It does seem to us, that what we propose here is a good compromise between simplicity, as well as clarity, when it comes to describing a situation, and efficiency and pertinence of the results provided by the formalism.

Our work differs from other approaches in the literature in that it strictly separates causality, ontology and explanations. The main advantages are that information is more properly expressed and that our approach is compatible with various accounts of these notions, most notably causality. In particular, we need no special instances of α *causes* α to hold nor γ (equivalent with β) to be an effect of α whenever α *causes* β holds (contrast with [3–7, 10] although in the context of actions such confusion is less harmful). In our approach, these are strictly limited to being where they belong, i.e., explanations. Space restriction prevent us from giving more details on the differences with other approaches.

Acknowledgements

It is a pleasure for us to thank the referees for their constructive remarks.

References

1. Baral, Chitta *Knowledge representation, reasoning and declarative problem solving.* Cambridge University Press, 2003.
2. Besnard Ph. and Cordier M.-O. Inferring Causal Explanations. In A. Hunter and S. Parsons (eds), *ECSQARU-99*, pp. 55–67, LNAI 1638, London, UK. Springer, 1999.
3. Bell J. Causation as Production. In G. Brewka, S. Coradeschi, A. Perini and P. Traverso (eds), *ECAI-06*, pp. 327-331, Riva del Garda, Italy. IOS Press, 2006.
4. Bochman A. A Logic for Causal Reasoning. In G. Gottlob and T. Walsh (eds), *IJCAI-03*, pp. 141–146, Acapulco, Mexico. Morgan Kaufmann, 2003.
5. Giunchiglia E., Lee J., Lifschitz V., McCain N., Turner H. Nonmonotonic Causal Theories. *Artificial Intelligence* 153(1–2):49–104, 2004.
6. Halpern J. and Pearl J. Causes and Explanations: A Structural-Model Approach. Part I: Causes. In J. S. Breese and D. Koller (eds), *UAI-01*, pp. 194–202, Seattle, Wa., USA. Morgan Kaufmann, 2001.
7. Halpern J. and Pearl J. Causes and Explanations: A Structural-Model Approach - Part II: Explanations. In B. Nebel (ed), *IJCAI-01*, pp. 27-34, Seattle, Wa., USA. Morgan Kaufmann, 2001.
8. Leone N., Pfeifer G., Faber W., Eiter T., Gottlob G., Perri S., and F. Scarcello. The DLV System for Knowledge Representation and Reasoning. *ACM Trans. on Computational Logic (TOCL)*, 7(3):499–562, 2006.
9. Mellor D. H. *The Facts of Causation.* Routledge, 1995.
10. Shafer G. Causal Logic. In H. Prade (ed), *ECAI-98*, pp. 711-720, Brighton, UK. Wiley, 1998.