Pertinence and Causality

Ramón P. Otero and Pedro Cabalar
AI Lab. - Dept. of Computer Science
University of Corunna
Corunna, Galicia, SPAIN
{otero,cabalar}@dc.fi.udc.es

Abstract

This paper studies the introduction of causality into a logical representation as an embodied ontological feature *per se*. Technically, this purpose is achieved by introducing the concept of *pertinence* as an extra valuation of atoms, parallel to the classical truth valuation. Based on this idea, we propose a narrative-based actions formalism which captures at once several aspects of reasoning about change, providing a simple and elegant solution to the frame and the ramification problems. Finally, we compare this approach to other recent proposals for causal action theories.

1 Introduction

One of the most recurrent topics in the AI literature for commonsense reasoning about actions has been the treatment of causal knowledge. In a recent line of research [McCain and Turner, 1995; 1997; 1998; Turner, 1999; Lifschitz, 1997; Giunchiglia and Lifschitz, 1998; Lin, 1995; 1996; Gustafsson and Doherty, 1996; Thielscher, 1997], motivation for introducing causality into action based formalisms is clear and sufficiently justified. However, causality is usually understood as a particular mechanism for rejecting some undesired models for a given domain description, and in many cases, causal information is simply disregarded after this models selection is done.

The aim of this paper is to emphasize the importance of causal knowledge and to treat causality not as a mere epistemological feature, but as an embodied ontological one. Consequently, we propose a new logical approach, L^2, in which causality is promoted to the category of an *extra valuation* of formulas, parallel to the classical truth valuation. This extra valuation receives the name of *pertinence* and its basic foundations have been extracted from the practical knowledge representation formalism of the Expert Systems Shell *Medtool* [Medtool, 1995], mainly used for temporal projection problems in the medical domain.

The interest of the paper is twofold. On the one hand, we will provide a comparative between pertinence and the different meanings given to the concept of "caused effect" by several well known approaches to causality. In a similar spirit to [Costello, 1997], we will study how these approaches understand the inertia law, but paying special attention here to its relation to causality. On the other hand, though we will not deepen on Medtool features, this paper also constitutes a first contribution of a research line focussed on establishing a strong, mathematically grounded, connection between our work in automated reasoning and the theoretical results for commonsense reasoning about actions.

This paper is organized as follows. In the next section, we introduce some intuitions about the concept of pertinence, providing two motivating examples. Next, we formulate the syntax and semantics of L^2 in the basic static case, adding, for comparison purposes, a circumscriptive encoding into propositional logic. Section 4 contains an elaboration of L^2 for narrative-based actions reasoning, called AL^2. In the next three sections, we establish formal comparisons with other three causal action approaches: Causal Theories [McCain and Turner, 1997], a slight variation of [Lin, 1995; 1996], and Temporal Action Logics [Doherty *et al.*, 1998]. Finally, we discuss the main contributions and draw some conclusions.

2 The concept of *pertinence*

Let us provide some intuitions about the concept of pertinence. Informally speaking, a formula will be pertinent if either we have explicitly asserted it to be so, or it is affected by other pertinent expressions. This general idea will be applied for action theories so that, a fluent will become pertinent when it is *affected by change*, that is, either we just have an observation on the domain stating it, or it is the result of applying suitable inference rules fired by other pertinent fluents.

Sometimes, when handling dynamic domains, it is interesting to understand pertinence as the opposite to persistence. In fact, when a formula is not pertinent, its truth value persists. This property is exploited, for instance, in the (temporal projection) inference mechanism of Medtool, allowing to disregard all the rules which conditions are not pertinent, without need of reevaluating them. However, this identification as non-persistence has to be taken with care, since *pertinence does not always imply a change in truth value.*

Let us study two motivating examples. Think, for instance, in the suitcase scenario introduced in [Lin, 1995] for studying the ramification problem:

Example 1 (The suitcase domain) A suitcase has two locks and a spring-loaded mechanism that will open the suitcase when both locks are in the up position. □

Suppose that, having opened the suitcase, we execute an action *wait* that does not cause any effect. The intuitive expected result is that everything should remain unchanged, that is, there should be no caused effects. As we will see later, formulations in some existing approaches conclude that the suitcase is *caused* to be open, while other assert that *open* is both obtained by inertia and simultaneously by a causal influence. Although in this example, distinguishing if *open* is caused or not has no further use, we can easily suppose an example in which this becomes interesting.

Example 2 (The gong example) A ballerina must start her choreography when the *gong* is struck. If the choreography is particularly short, she may even finish while the gong is still sounding. □

Intuitively, there is no contradiction when the ballerina finishes while the gong is still sounding. There exists an ontological difference between *gong* being true (the gong is sounding) and being true and pertinent (the gong has been struck). In this way, if *gong* remains true for three situations, and has become pertinent in each one, we should conclude that it has been struck *three times*, whereas if it has remained not pertinent, we should conclude that it is due to the persistence of *the same stroke.*

Of course, this example could be formulated by using an inertial fluent *gong_sounding* and a non-inertial fluent *gong_struck*, but as we will see later, the solution with pertinence avoids such distinction, allowing this duality for any fluent, and uses, in an elegant way, the same constructs than those added for solving the frame and ramification problems.

3 L^2 Syntax and semantics

The syntax of L^2 is the classical propositional one plus the new conditional expressions:

$$\psi \Leftarrow \phi \tag{1}$$

being ϕ and ψ classical propositional formulas. Notice that this conditional will be different from the usual material implication we will represent here as $\phi \supset \psi$.

For this simple syntax, we introduce a semantics with *two simultaneous* valuations. Apart from the usual truth value (true or false, $\{\mathtt{t},\mathtt{f}\}$), a formula will also be assigned a *pertinence* value: "pertinent" or "non-pertinent", denoted respectively as $\mathtt{p}$ and $\mathtt{n}$.

Given a signature $\mathcal{S}$, a L^2 *interpretation*, I, is a mapping $I : \mathcal{S} \longrightarrow \{\mathtt{t},\mathtt{f}\} \times \{\mathtt{p},\mathtt{n}\}$ assigning truth and pertinence to each atom. In this way, though the underlying concept remains as a dual valuation, we will technically handle a multivalued logic with the four combinations $\{\mathtt{tp},\mathtt{tn},\mathtt{fp},\mathtt{fn}\}$. An interpretation I may be alternatively described by a pair $\langle S, P\rangle$ of sets of atoms, for representing truth and pertinence respectively:

$$\begin{array}{ll} I(q) \in \{\mathtt{tp},\mathtt{tn}\} & \text{iff } q \in S \\ I(q) \in \{\mathtt{tp},\mathtt{fp}\} & \text{iff } q \in P \end{array}$$

$$(\neg\phi)^I = \begin{cases} \mathtt{t}\alpha & \text{if } \phi^I = \mathtt{f}\alpha \\ \mathtt{f}\alpha & \text{if } \phi^I = \mathtt{t}\alpha \end{cases}$$

$$(\psi \Leftarrow \phi)^I = \begin{cases} \mathtt{tn} & \text{if } \phi^I \neq \mathtt{tp} \\ \mathtt{fn} & \text{if } \phi^I = \mathtt{tp} \text{ and } \psi^I \neq \mathtt{tp} \\ \mathtt{tp} & \text{otherwise} \end{cases}$$

$(\phi \wedge \psi)^I$	fn	fp	tn	tp
fn	fn	fp	fn	fp
fp	fp	fp	fp	fp
tn	fn	fp	tn	tp
tp	fp	fp	tp	tp

$(\phi \vee \psi)^I$	fn	fp	tn	tp
fn	fn	fp	tn	tp
fp	fp	fp	tp	tp
tn	tn	tp	tn	tp
tp	tp	tp	tp	tp

Figure 1: Interpretation of a formula.

Interpretations are extended to non-atomic formulae by a variation of standard structural induction, shown in Figure definition 1. There, ϕ^I stands for the pair of values that interpretation I assigns to formula ϕ. Also, variable α is used to represent any pertinence value.

In short, truth for propositional formulas is interpreted in the classical way, whereas their pertinence is independently measured by checking whether any atom occurring in the formula is pertinent. The relation between truth and pertinence can only be established by using causal rules. A causal rule $(\psi \Leftarrow \phi)$ is understood as: "if ϕ is true and pertinent, then ψ is true and pertinent." Causal rules act as a pertinence propagation, reflecting how changes may affect to other changes.

As in any multivalued logic, in order to define the concept of model, we will need to select a value to understand that a formula is satisfied. If nothing else is specified, we will assume that assigning any value from $\{\mathtt{tp},\mathtt{tn}\}$, *satisfies* a formula, that is, it is enough with making it true. However, we sometimes will be interested in being more specific. To this aim, we adopt the notation $\phi\langle v\rangle$, where $v \in \{\mathtt{tp},\mathtt{tn},\mathtt{fp},\mathtt{fn},\mathtt{t},\mathtt{f},\mathtt{p},\mathtt{n}\}$, to stand for "formula ϕ is satisfied if it is assigned value v."

An interpretation I is said to be a *model* of a theory T when it satisfies all the formulas in T, that is, it assigns to each formula its selected value. A set of formulas T *entails* a formula ϕ, if all the models of T are models of

ϕ.

Consider the theory $T_1 = \{c \Leftarrow \neg a \wedge b,\ \ \neg a,\ \ b\langle\mathtt{tn}\rangle\}$. The condition of the causal rule is true, since both $\neg a$ and b hold. Its pertinence, however, depends exclusively on the pertinence of atom a (b is non-pertinent). We have five models: one where a is pertinent, the rule is applied and we obtain $\{c\langle\mathtt{tp}\rangle\}$; and the rest where a is non-pertinent, leaving c free.

This basic monotonic framework must be extended to capture the default: "any atom should only be pertinent when there exists some evidence for that." In our example, since nothing is said about a, we should conclude it is not pertinent. In the same way, being a non-pertinent, the conclusion c is free, and we should assume again non-pertinence. The following definitions capture this idea:

Definition 1 ($\leq_\mathtt{p}$) Given two L^2 interpretations $I = \langle S, P\rangle$ and $I' = \langle S', P'\rangle$, we say that I is lower than I', $I \leq_\mathtt{p} I'$, iff $S = S'$ and $P \subseteq P'$. □

Definition 2 (L^2 **selected model**) A model $I = \langle S, P\rangle$ of a theory T is said to be a *selected model* iff it is a minimal model, with respect to the $\leq_\mathtt{p}$ ordering relation. □

Selected models for T_1 correspond to the two cases for truth value of c: $\{\neg a, b, c\}$ and $\{\neg a, b, \neg c\}$, being $\{a\langle\mathtt{n}\rangle, b\langle\mathtt{n}\rangle, c\langle\mathtt{n}\rangle\}$ satisfied by both models. Non-monotonicity can be easily shown. Theory T_1 entails $c\langle\mathtt{n}\rangle$, whereas $T_1 \cup \{a\langle\mathtt{p}\rangle\}$ has as single model $\{a\langle\mathtt{fp}\rangle, b\langle\mathtt{tn}\rangle, c\langle\mathtt{tp}\rangle\}$ and so, does not entail $c\langle\mathtt{n}\rangle$ any more.

3.1 Propositional circumscriptive encoding

In this section we show how the semantic features of L^2 can be translated into propositional logic with Circumscription, for making easier the comparison to other approaches. It must be taken into account that the base propositional formalism provides more representational freedom than the one obtained by directly translating L^2, which exploits as one of its features, that causal rules have a restricted form. Apart from this remark, we could have presented L^2 as a surface language (in the sense, for instance, of $\mathcal{L}(\mathcal{SD})$ in [Doherty *et al.*, 1998]), providing its semantics in terms of translation to classical logic. However, we consider an advantage the availability of a simple "direct" semantics, which is closer to the way in which Medtool practical formalism works.

Encoding of L^2 into propositional logic is simply done by an a priori syntactic transformation that reduces pertinence to a new kind of renamed atoms of the shape[1] $q\langle\mathtt{p}\rangle$, for any $q \in \mathcal{S}$. Then, for any propositional formulas ϕ, ψ, we define $\langle\mathtt{p}\rangle$ as the recursive syntactic transformation:

$$\begin{aligned}(\neg\phi)\langle\mathtt{p}\rangle &\stackrel{\text{def}}{=} \phi\langle\mathtt{p}\rangle\\ (\phi\vee\psi)\langle\mathtt{p}\rangle &\stackrel{\text{def}}{=} \phi\langle\mathtt{p}\rangle \vee \psi\langle\mathtt{p}\rangle\\ (\phi\wedge\psi)\langle\mathtt{p}\rangle &\stackrel{\text{def}}{=} \phi\langle\mathtt{p}\rangle \vee \psi\langle\mathtt{p}\rangle\end{aligned}$$

[1] Note that we handle $\langle\mathtt{p}\rangle$ as a suffix operator with the greatest precedence.

and introduce the derived operators:

$$\begin{aligned}\phi\langle\alpha\beta\rangle &\stackrel{\text{def}}{=} \phi\langle\alpha\rangle \wedge \psi\langle\beta\rangle\\ \phi\langle\mathtt{t}\rangle &\stackrel{\text{def}}{=} \phi\\ \phi\langle\mathtt{f}\rangle &\stackrel{\text{def}}{=} \neg\phi\\ \phi\langle\mathtt{n}\rangle &\stackrel{\text{def}}{=} \neg\phi\langle\mathtt{p}\rangle\end{aligned}$$

Finally, causal rules are translated as:

$$\begin{aligned}\psi \Leftarrow \phi &\stackrel{\text{def}}{=} \phi\langle\mathtt{tp}\rangle \supset \psi\langle\mathtt{tp}\rangle\\ (\psi \Leftarrow \phi)\langle\mathtt{p}\rangle &\stackrel{\text{def}}{=} \phi\langle\mathtt{tp}\rangle\end{aligned}$$

Theorem 1 *Each model of a L^2 theory corresponds to a model of its propositional encoding and vice versa.*

Proof Any L^2 interpretation $\langle S, P\rangle$ corresponds to a propositional interpretation $I(S, P) = S \cup \{q\langle\mathtt{p}\rangle : q \in P\}$ and vice versa. The syntactic transformations are a straightforward rephrasing of the tables in Figure definition 1. □

The translation of the example theory T_1 would be:

$$T_1 = \{\neg a \wedge b \wedge (a\langle\mathtt{p}\rangle \vee b\langle\mathtt{p}\rangle) \supset (c \wedge c\langle\mathtt{p}\rangle),\ \neg a,\ b \wedge \neg b\langle\mathtt{p}\rangle\}$$

Thanks to theorem 1, from now on we will indistinctly use any L^2 theory T as a propositional theory, and any L^2 model $\langle S, P\rangle$ as its corresponding propositional structure $I(S, P)$. This translation, however, does not cover yet the definition 2 of L^2 selected models.

Theorem 2 *There exists a one to one correspondence between selected models of a L^2 theory T with signature $\mathcal{S} = \{q^1, q^2, \ldots, q^n\}$ and models of the parallel circumscription* $\mathrm{CIRC}[T;\ q^1\langle\mathtt{p}\rangle, q^2\langle\mathtt{p}\rangle, \ldots, q^n\langle\mathtt{p}\rangle;]$

Proof It follows from the models selection of parallel propositional circumscription as described in section 7.1 in [Lifschitz, 1993]. The order relation $\leq_\mathtt{p}$ in L^2 corresponds to $\leq_{P;Z}$ defined there with P as the vector of pertinence atoms $q^i\langle\mathtt{p}\rangle$, and Z the empty set (the rest of atoms are fixed). □

4 Pertinence for Reasoning About Actions: AL^2

Once we have seen how pertinence works in a static domain, we will extend this concept to a narrative-based actions formalism inspired to the work of [McCain and Turner, 1997]. We will introduce a new logic AL^2 (from *Action* L^2) where we will distinguish between action and fluents and we will locate their occurrences along a narrative denoted with nonnegative integer subindices.

Intuitively, the use of pertinence in a dynamic domain will bring out which fluents (and formulas) have been affected by the changes due to an action performance. As a consequence, two new postulates will be introduced: (1) an action will always be assumed to be pertinent; and (2) a non-pertinent fluent must maintain its previous truth value.

From a syntactical point of view, AL^2 is just a specialization of L^2, separating the signature into two disjoint sets: the set $\mathcal{F}$ of *fluent names* and the set $\mathcal{A}$ of *action* *names.* Atoms will have the shape q_t where q is either a fluent or an action name and t is a non-negative integer ranging from 0 to the number of stituations in the narrative.

A causal rule $F \Leftarrow \phi$ in AL^2 must satisfy two new syntactic restrictions: (1) F must have the shape f_i or $\neg f_i$ (for some fluent f), and (2) there cannot occur a subindex $j > i$ in ϕ. The first restriction intuitively says that actions are never "concluded" as a result of a rule. We will be only interested in concluding fluent literals. The second condition means that a fact for a fluent can only be concluded from facts in the same or in previous situations. It would be difficult to think about causal influences from future to past.

Typically, an AL^2 domain description will consist of two sets of formulas $\langle B, O\rangle$: the set B called *background knowledge* which contains causal rules, and the set O of *observations* which are non-causal formulas. We will be interested in causal rules that are satisfied for any situation in the narrative. To this aim, we use an abbreviation allowing causal rules with variable subindices and expressions relating them. Thus, a rule like:

$$f_t \Leftarrow a_t \wedge g_{t-1}$$

actually represents:

$$\begin{array}{lcl} f_1 & \Leftarrow & a_1 \wedge g_0 \\ f_2 & \Leftarrow & a_2 \wedge g_1 \\ & \cdots & \\ f_n & \Leftarrow & a_n \wedge g_{n-1} \end{array}$$

Pertinence of actions can be captured now by the axiom schemata:

$$a_t\langle \mathtt{p}\rangle \tag{2}$$

for any action $a \in \mathcal{A}$. Persistence of non-pertinent fluents is represented by the inertia law:

$$\neg f_t\langle \mathtt{p}\rangle \supset (f_t \equiv f_{t-1}) \tag{3}$$

for any fluent f and for $t > 0$. We will apply a *filter preferential entailment* technique (see for instance [Sandewall, 1989]), imposing inertia *outside* the minimization for selected models.

Definition 3 (Causal model) A model I of a theory T is said to be *causal* iff it is a selected model of T and satisfies the inertia law schema (3). $\square$

In other words, causal models correspond to the theory:

$$\text{CIRC}[T;\ q^1\langle \mathtt{p}\rangle, q^2\langle \mathtt{p}\rangle, \ldots, q^n\langle \mathtt{p}\rangle;] \cup (3)$$

being q^i all the subindexed atoms for actions and fluents. The AL^2 formulation of example 1 consists of a set of actions $\mathcal{A} = \{toggle(l1), toggle(l2)\}$, a set of fluents $\mathcal{F} = \{up(l1), up(l2), open\}$ and a set B of causal rules:

$$\begin{array}{rcl} open_t & \Leftarrow & up(l1)_t \wedge up(l2)_t \\ up(L)_t & \Leftarrow & toggle(L)_t \wedge \neg up(L)_{t-1} \\ \neg up(L)_t & \Leftarrow & toggle(L)_t \wedge up(L)_{t-1} \end{array}$$

where L is a metavariable ranging on $\{l1, l2\}$, to be used from now on. The propositional translation would amount to:

$$up(l1)_t \wedge up(l2)_t \wedge (up(l1)_t\langle \mathtt{p}\rangle \vee up(l1)_t\langle \mathtt{p}\rangle) \supset open_t \wedge open_t\langle \mathtt{p}\rangle \tag{4}$$

$$toggle(L)_t \wedge \neg up(L)_{t-1} \supset up(L)_t \wedge up(L)_t\langle \mathtt{p}\rangle \tag{5}$$

$$toggle(L)_t \wedge up(L)_{t-1} \supset \neg up(L)_t \wedge up(L)_t\langle \mathtt{p}\rangle \tag{6}$$

Let us study the case where only lock 2 was up, the suitcase was closed and we toggle lock 1. This is captured by the observation:

$$\neg up(l1)_0 \wedge up(l2)_0 \wedge \neg open_0 \wedge toggle(l1)_1 \wedge \neg toggle(l2)_1 \tag{7}$$

In this example we obtain the following 3 selected models corresponding to:

$$\{up(l2)_1, open_1, \neg up(l2)_1\langle \mathtt{p}\rangle, open_1\langle \mathtt{p}\rangle\}$$
$$\{\neg up(l2)_1, open_1, \neg up(l2)_1\langle \mathtt{p}\rangle, \neg open_1\langle \mathtt{p}\rangle\}$$
$$\{\neg up(l2)_1, \neg open_1, \neg up(l2)_1\langle \mathtt{p}\rangle, \neg open_1\langle \mathtt{p}\rangle\}$$

As we can see, the last two models do not satisfy inertia, since in spite of making both fluents non-pertinent, they change their truth value with respect to initial situation. Therefore, the first model is the unique causal one for this domain description. Let us suppose that, as explained in section 3.1 we extend the example by executing an action *wait* in situation 2. The unique causal model entails now $open_2 \wedge \neg open_2\langle \mathtt{p}\rangle$, that is, the suitcase persists open. As for example 2 a possible formulation could consist of set of actions $\mathcal{A} = \{strike, finish, wait\}$, a set of fluents $\mathcal{F} = \{dancing, gong\}$ and a set B of causal rules:

$$gong_t \Leftarrow strike_t \tag{8}$$

$$dancing_t \Leftarrow gong_t \tag{9}$$

$$\neg dancing_t \Leftarrow finish_t \tag{10}$$

Given the observations:

$$\neg dancing_0, strike_1, wait_2, finish_3$$

if we assume no other action is executed, we obtain as unique causal model:

$$\{gong_1, gong_1\langle \mathtt{p}\rangle, dancing_1, dancing_1\langle \mathtt{p}\rangle,$$
$$gong_2, \neg gong_2\langle \mathtt{p}\rangle, dancing_2, \neg dancing_2\langle \mathtt{p}\rangle,$$
$$gong_3, \neg gong_3\langle \mathtt{p}\rangle, \neg dancing_3, dancing_3\langle \mathtt{p}\rangle\}$$

As we can see, the facts $gong_3$ and $\neg dancing_3$ do not contradict rule (9) since its condition is not pertinent.

4.1 Encoding into Causal Action Theories

In this section we study the relationship between AL^2 and Causal Action Theories, as proposed in [McCain and Turner, 1997]. In a first step, we study the static case, showing that L^2 can be encoded into Causal Theories instead of the already presented translation using

Circumscription. Then, we study how the additional assumptions in AL^2 can be incorporated and compare the result to Causal Action Theories.

A Causal Theory is a propositional theory together with a new kind of (non-nested) conditional formulas of the shape $\phi \rightarrow \psi$. Given a set of causal rules D and a propositional interpretation I, the theory D^I is defined as:

$$D^I = \{\psi : \phi \rightarrow \psi \in D \text{ and } I \models \phi\}$$

That is, D^I is a classical theory containing the consequents of all those rules which condition are true in I. An interpretation I is said to be a *causally explained model* of a causal theory D iff it is the unique model of D^I.

The following definitions constitute a rephrasing of definition 2 of L^2 selected models which is closer to Causal Theories:

Definition 4 (I-model) Given a L^2 theory T and an interpretation $I = \langle S, P \rangle$, a model $I' = \langle S, P' \rangle$ of T is said to be an *I-model* iff $P' \subseteq P$. □

That is, an I-model is any L^2 model that has the same truth than I and *less or equal* pertinent atoms.

Definition 5 (L^2 selected model) A model $I = \langle S, P \rangle$ of a L^2 theory T is said to be a *selected model* iff it is the unique I-model of T. □

Theorem 3 *An interpretation I is a model under definition 2 iff it is a model under definition 5.*

Proof The proof is trivial. We have that I is the unique I-model iff there cannot be a different model $I' = \langle S, P' \rangle$ with $P' \subseteq P$, and this is equivalent to asserting $I \leq_{\mathrm{p}} I'$, for any L^2-model I'. □

As we can see, this new definition bears a strong resemblance to definition of causally explained model and directly leads to the following result:

Theorem 4 *There exists a one to one correspondence between selected models of a L^2 theory T (in its propositional encoding) and models of the Causal Theory D consisting of:*

1. *a rule $(true \rightarrow \phi)$, for any $\phi \in T$*
2. *the rules:*

$$\begin{aligned} q &\rightarrow q & (11)\\ \neg q &\rightarrow \neg q & (12)\\ \neg q\langle \mathrm{p} \rangle &\rightarrow \neg q\langle \mathrm{p} \rangle & (13) \end{aligned}$$

for each symbol $q \in \mathcal{S}$.

Proof Given any model I of T, the theory D^I is T plus all the literals for truth and the negative pertinence literals satisfied by I. Therefore, models of D^I correspond to models of T with same truth than I but less or equal pertinence atoms, which is exactly the definition 4 of I-models. Then, requiring that the unique model for D^I is I corresponds exactly to the uniqueness condition for selected model in definition 5. □

Though this result allows us to recast any L^2 theory into a Causal Theory, some details must be taken into account. From the Causal Theories point of view, the translation of a L^2 theory may introduce two problems. First, we have had to introduce a new kind of atoms, $q\langle \mathrm{p} \rangle$, for expressing pertinence. This brings out that the L^2 understanding of a "causal rule" handles more information than a typical rule from Causal Theories.

Second, the resulting translated theories are not suitable, at least directly, for applying completion. In [McCain and Turner, 1997], a simple and powerful method called *literal completion* was introduced for converting any causal theory into a classical propositional theory. Literal completion can be applied when the causal theory is *definite*, that is, the consequent of each rule is a literal. The encoding of a L^2 rule, however, leads to a rule $true \rightarrow \phi$, being ϕ a complex propositional formula.

Once we have seen how to encode static L^2 theories as Causal Theories, the only additional issue for encoding AL^2 is how to obtain the AL^2 causal models (definition 3), that is, how to require satisfaction of inertia afterwards the models selection. We will not be able to introduce inertia as any L^2 causal rule, using the formulation:

$$true \rightarrow \phi \tag{14}$$

Instead, we have to use:

$$\neg\phi \rightarrow false \tag{15}$$

to point out that the inertia law is checked *after* the causally explained models are fixed. The difference between causal rules of the shape (14) and (15) is commented in [McCain and Turner, 1997] to cope with classical ramification and qualification constraints respectively.

Theorem 5 *Let T be an AL^2 theory, and D its encoding into Causal Theories. Then, there exists a one to one correspondence between causal models of T and causally explained models for the theory D' which results from adding to D the rules:*

$$\neg(\neg f_t\langle \mathrm{p} \rangle \supset (f_t \equiv f_{t-1})) \rightarrow false \tag{16}$$

for any fluent f and $t > 0$.

Proof If a model I does not satisfy inertia, then some of the conditions (16) will introduce a $false$ in the theory D'^I, which will cause it to have no models. So I cannot be a causally explained model of D' even if it was so for D. Now, we have to show that all the rest of causally explained models of D will also be so for D'. This is clear because, if I was a selected model of T which satisfies inertia, then all the introduced rules (16) would have a false condition with respect to I, and so $D'^I = D^I$. □

The AL^2 formulation of the suitcase domain, encoded in Causal Theories, would amount to the theory D:

$$\begin{aligned} true \;\rightarrow\; & (up(l1)_t \wedge up(l2)_t \wedge [up(l1)_t\langle \mathrm{p} \rangle \vee \\ & up(l1)_t\langle \mathrm{p} \rangle] \supset [open_t \wedge open_t\langle \mathrm{p} \rangle]) \quad (17)\\ true \;\rightarrow\; & (toggle(L)_t \wedge \neg up(L)_{t-1} \supset \end{aligned}$$

$$[up(L)_t \wedge up(L)_t\langle \mathtt{p}\rangle]) \quad (18)$$

$$true \rightarrow (toggle(L)_t \wedge up(L)_{t-1} \supset [\neg up(L)_t \wedge up(L)_t\langle \mathtt{p}\rangle]) \quad (19)$$

$$true \rightarrow (\neg up(l1)_0 \wedge up(l2)_0 \wedge \neg open_1 \wedge toggle(l1)_1) \quad (20)$$

together with the general rules for fixing truth ((11) and (12)), minimizing pertinence (13) and introducing inertia outside (16). Now, let us have a look at the original formulation, D', by McCain and Turner:

$$up1_t \wedge up2_t \rightarrow open_t \quad (21)$$

$$toggle(L)_t \wedge up(L)_t \rightarrow \neg up(L)_{t+1} \quad (22)$$

$$toggle(L)_t \wedge \neg up(L)_t \rightarrow up(L)_{t+1} \quad (23)$$

together with the general axiom schemata:

$$a_t \rightarrow a_t \qquad \neg a_t \rightarrow \neg a_t \quad (24)$$

$$f_0 \rightarrow f_0 \qquad \neg f_0 \rightarrow \neg f_0 \quad (25)$$

$$f_t \wedge f_{t+1} \rightarrow f_{t+1} \quad (26)$$

$$\neg f_t \wedge \neg f_{t+1} \rightarrow \neg f_{t+1} \quad (27)$$

for any action a and fluent f.

A first minor difference between D and D' is the way in which situations and actions are located. Notice that rules in D are *oriented* [2] from situation $t-1$ to t, whereas in D' the formulation is from t to $t+1$. In the same way, actions in D are located in the same situation than their effects, whereas in D' they are located in a previous situation.

Apart from this minor distinction in the representation conventions, we can outline three main differences between the use of causal rules in both theories:

1. Causal relations in D are treated as ramification constraints (the antecedent is always *true*), whereas in D' they are represented as "genuine" causal rules.
2. Inertia in D' is represented as a default (atom f_{i+1} occurs in both sides of the rule) while in D, it is represented as a qualification constraint (a rule with consequent $false$).
3. Free assumptions in D' (formulas (24) to (25)) allow explaining any action occurrence or any fact at the initial situation, whereas in D, we may also freely assume a truth literal (formulas (11) and (12)) and a negated pertinence (formula (13)) *at any situation*.

Finally, notice how introducing pertinent atoms avoids representing inertia as a default. In fact, the default is applied now to the minimization of pertinence (rule (13)). Besides, causal rules (formulas from (17) to (19)) are only applied when the condition has some new evidence (that is, the condition is always required to be pertinent. Think for instance in the situation introduced in section 3.1. When we execute action *wait* with an open suitcase, D' obtains that it remains open. However, in D only inertia makes us conclude that *open* must be true, whereas in D' this is obtained simultaneously by inertia and by causal rule (21). As for example 2, it cannot be properly represented without the introduction of pertinence, since the axioms:

$$strike_t \rightarrow gong_t$$

$$gong_t \rightarrow dancing_t$$

$$finish_t \rightarrow \neg dancing_t$$

would lead to an inconsistence when the choreography finishes while the gong remains sounding.

4.2 Lin's Action Theories

In [Lin, 1995], a causal approach was introduced for reasoning about actions based on Situation Calculus and Circumscription. In short, Lin's method consists in the introduction of a new predicate $caused(f, v, s)$, where f is a fluent, v a truth value ($true$ or $false$) and s a situation. This predicate is then minimized (using a fixed circumscription) and in a final step, inertia is applied to rule out undesired models. As we can see, Lin's method bears a strong resemblance to the application of inertia outside the models selection process made in AL^2.

Lin's work has also been compared to Causal Theories. More accurately, the comparison was done with respect to a modal generalization of Causal Theories called *Universal Causation Logic* (UCL), introduced by Turner in [Turner, 1999]. For the sake of comparison, we will follow similar steps to those done there, simplifying Lin's proposal for the propositional narrative case, instead of using Situation Calculus.

Let us define a PNLT (Propositional Narrative Lin Theory) as a set of propositional formulas for the sorted signature $\mathcal{A} \cup \mathcal{F}$. We extend the set of atomic formulas to symbols with subindices q_t (like in section 4) and two new kinds of atoms: for each fluent symbol f_t, we will have $C^+ f_t$ and $C^- f_t$, standing for "caused true" and "caused false" respectively.

Since $C^+ f_t$ and $C^- f_t$ atoms imply some truth value for f_t, we will include the axiom schemata:

$$C^+ f_t \supset f_t \quad (28)$$

$$C^- f_t \supset \neg f_t \quad (29)$$

A direct effect axiom in PNLT is a formula of the shape:

$$a_t \wedge H_{t-1} \supset C f_t \quad (30)$$

where a is an action, H is a propositional formula with fluent atoms f^i_{t-1} and $C f_t$ is either $C^+ f_t$ or $C^- f_t$. Following this same notation, a "causal rule" axiom corresponds to a formula like:

$$H_t \supset C f_t \quad (31)$$

An explicit precondition would have the shape:

$$a_t \supset H_t \quad (32)$$

and a qualification constraint would look like:

$$H_t \quad (33)$$

[2] This matches the syntax of the practical system.

Using this modification of Lin's Action Theories, the formalization of the suitcase example presented in [Lin, 1995] would lead to the formulas:

$$up(l1)_t \wedge up(l2)_t \supset C^+ open_t \quad (34)$$
$$toggle(l1)_t \wedge \neg up(l1)_{t-1} \supset C^+ up(l1)_t \quad (35)$$
$$toggle(l1)_t \wedge up(l1)_{t-1} \supset C^- up(l1)_t \quad (36)$$

Let T include the formulas (28), (29) and from (34) to (36). In Lin's original formulation, all the domain formulas where minimized by circumscribing predicate *caused*, fixing the truth for all fluents (predicate *holds*). In our propositional version, we make instead a parallel circumscription minimizing all the atoms C^+ and C^-, while fixing the rest of atoms. Thus, we would have:

$$T^* = CIRC[T; C^+ f^i, C^- f^i, \ldots, C^+ f^n, C^- f^n;]$$

being f^i each possible subindexed fluent atom. The final models are obtained in a second step where, outside the minimization, we add the inertia axiom requiring persistence for everything not caused. In this case, inertia will constitute an axiom schemata:

$$(\neg C^+ f_t \wedge \neg C^- f_t) \supset (f_t \equiv f_{t-1}) \quad (37)$$

for any fluent f and for $t > 0$. Therefore, Lin's selected models would correspond to those from the theory: $T^* \cup$ (37).

At a first glance, there exist two main differences between these narrative Lin Theories and AL^2:

1. Instead of using pertinence of an atom, we have two kind of atoms C^+ and C^- that seem to capture simultaneously both the concept of "caused" and the truth value of the atom.
2. Conditions of causal rules are just required to be true. Caused atoms do not occur in the condition of a rule.

The first difference is that AL^2 *avoids reifying* the truth values. This is more economic, in the sense that axioms (28) and (29) are not necessary any more. Furthermore, it can be easily seen that we can establish the relation:

$$C^+ f \equiv (f \wedge f\langle \mathrm{p} \rangle) \quad (38)$$
$$C^- f \equiv (\neg f \wedge f\langle \mathrm{p} \rangle) \quad (39)$$

Besides, we define causality in a broader sense, extended for complex formulas and not only for literals.

The second difference is also related to the idea of pertinence and depends on the shape of the Lin Theory we are studying. Although in the original formalization of the suitcase example presented by Lin, caused literals did not occur in the condition of rules, this was possible in their general shape. The problem is that it was not enough clear when these atoms should appear or not. In L^2 there is not any choice for that, since a condition of a rule *must necessarily be pertinent* in order to be applied.

As an example of this, let us study again the situation for the suitcase domain presented in section 3.1. Both Lin's and AL^2 formulations will obtain that $open_t$ is true. However, Lin's formulation would also obtain that $open_t$ has become caused, when there was not any reason for that (it was not affected by change). This objection is basically the same presented in comparison to Causal Theories. Nevertheless, it must be noticed that in Lin's case, there actually exists the possibility of distinction between facts affected and not affected by change. In fact, it would have been enough to add to the condition of the causal rule the requirement that some of its effects were caused:

$$C^+ up(l1)_t \vee C^+ up(l2)_t$$

Something similar would happen in the gong example, that should be formulated as:

$$strike_t \supset C^+ gong_t$$
$$C^+ gong_t \supset C^+ dancing_t$$
$$finish_t \supset C^- dancing_t$$

in order to obtain the expected result.

Another interesting advantage with respect to Lin's domain is a greater commodity and a clearest definition for requiring a complex condition to be caused. Thus, for instance, if we want the condition $up(l1)_t \wedge up(l2)_t$ to be caused instead of just to be hold, we cannot simply require:

$$C^+ up(l1)_t \wedge C^+ up(l2)_t \supset C^+ open_t \quad (40)$$

since it would have a completely different meaning (as also pointed out in [Turner, 1999]). Instead, we should understand the condition to be caused when at least one of the conjuncts is caused, having the two rules:

$$C^+ up(l1)_t \wedge up(l2)_t \supset C^+ open_t$$
$$up(l1)_t \wedge C^+ up(l2)_t \supset C^+ open_t$$

In AL^2 these two rules simply correspond to the single causal rule:

$$open_t \Leftarrow up(l1)_t \wedge up(l2)_t$$

In order to enforce that this interpretation of pertinence for complex formulas is a natural and intuitive concept, it is interesting to note the surprising similarity to a recent proposal introduced by Denecker [Denecker *et al.*, 1998] where the concept of *initiation* for complex formulas is defined in an analogous way.

Finally, since we have seen that AL^2 causal models correspond to (a slight variation of) Lin theories, it directly follows that Lin theories may be encoded into Causal Theories. A similar comparison study was made by Turner in [Turner, 1999]. In that work, Turner developed a generalization of Causal Theories consisting in a modal approach called *Universal Causation Logic* (UCL) which not only subsumed Causal Theories, but also a subset of Default Logic. The shape of UCL is that of S5 modal logic, replacing the necessity operator by the letter C, meaning that a formula is caused. Turner developed an elaboration of Lin's theories similar to PNLT, and proved that any model of a Lin theory has a partial correspondence to a model of an UCL theory (theorem 4 in [Turner, 1999]). However this correspondence was not complete, since it was only established for the "caused" part of the Lin's model.

The partiality of this result is not surprising since one of the postulates of UCL is that, in the final selected models, everything must result to be caused, which of course, is not true in any Lin theory – otherwise, inertia would simply be never applicable.

The encoding obtained here extends Turner's result and covers complete and accurately the models of any Lin theory. This is possible because we have incorporated special atoms for representing Lin's causation, apart from the causal conditionals used in Causal Theories. It follows from this that the meanings of "caused" in Lin and in UCL are actually different.

4.3 Temporal Action Logics

Another well known research trend for Reasoning About Actions is the one initiated in [Sandewall, 1994] as a general framework for assessment from which several approaches have emerged for particular classes of action scenario descriptions. Inside this trend, we will focus on two recent papers [Gustafsson and Doherty, 1996; Doherty *et al.*, 1998] describing *Temporal Action Logics* (TAL).

As an example, following TAL syntax, a possible representation of the suitcase scenario using the surface language $\mathcal{L}(\mathcal{SD})$ could be:

$$[t, t+1]toggle(L) \leadsto ([t]up(L) \supset [t+1]\neg up(L)) \quad (41)$$

$$[t, t+1]toggle(L) \leadsto ([t]\neg up(L) \supset [t+1]up(L)) \quad (42)$$

$$\forall t[t]((up(l1) \wedge up(l2)) \gg [t]open) \quad (43)$$

In order to translate it into the base language, we will adopt the same notation policy as for the other approaches. We assume a sorted signature $\mathcal{A} \cup \mathcal{F}$ with actions and fluents. Time points are directly represented as subindices. Finally, the *occlude* predicate is replaced by atoms of the shape Xf_t, meaning "f is occluded at point t." The result would be a theory T with the formulas:

$$toggle(L)_t \wedge \neg up(L)_t \supset up(L)_{t+1} \wedge Xup(L)_{t+1} \quad (44)$$

$$toggle(L)_t \wedge up(L)_t \supset \neg up(L)_{t+1} \wedge Xup(L)_{t+1} \quad (45)$$

$$up(l1)_t \wedge up(l2)_t \supset open_t \quad (46)$$

$$\text{CHANGED}(up(l1)_{t+1} \wedge up(l2)_{t+1}) \supset Xopen_{t+1} \quad (47)$$

where:

$$\text{CHANGED}(\phi_{t+1}) \stackrel{\text{def}}{=} \neg\phi_t \wedge \phi_{t+1}$$

The semantics consists in a filter preferential entailment, circumscribing the Xf_t atoms for these resulting formulas and conjoining afterwards the inertia axiom schemata:

$$\neg(f_t \equiv f_{t+1}) \supset Xf_{t+1} \quad (48)$$

That is, the final selected models correspond to the theory: $CIRC[T; Xf_1, \ldots, Xf_n;] \wedge (48)$

Again, the minimization technique is analogous to that in AL^2 and Lin's theories. In this case, the similarity to AL^2 is greater for the way of representing a "caused" fact, since the occlusion atoms do not contain information about truth, like happens with Lin's notation C^+ and C^-. Thus, at a first sight, we can simply make the identification:

$$Xf_t \equiv f_t\langle \text{p} \rangle$$

Under this correspondence, effect axioms in AL^2, (in the example, (5) and (6), have exactly the same shape than their counterparts in TAL ((44) and (45) respectively), and both formulations of inertia (3) and (48) are directly equivalent. Despite of this great similarity, there exist several main differences between TAL and AL^2 related to the way in which the concepts of occlusion and pertinence are understood, specially with respect to application of causal rules.

The first difference is the way in which conditions of causal rules are translated. This also happened for Lin's theories. However, formula (47) has a stronger condition than that in (34), since it not only requires the condition to be true, but also to have *changed from false to true.* For instance, for the suitcase example, this avoids *open* to become occluded when *wait* is executed and nothing changes.

With respect to AL^2 causal rules, we can see that from (46) and (47) we can conclude:

$$\text{CHANGED}(up(l1)_{t+1} \wedge up(l2)_{t+1}) \supset open_{t+1} \wedge Xopen_{t+1}$$

which can be seen as an approximation of (4), bearing in mind that CHANGED(ϕ) implies that ϕ is pertinent. Thus, TAL condition is more restrictive than that in AL^2 — the rule is applied in less cases. This can be easily seen in the gong example:

$$finishes_t \supset \neg dancing_t \quad (49)$$

$$CHANGED(finishes_{t+1}) \supset Xdancing_{t+1} \quad (50)$$

$$gong_t \supset dancing_t \quad (51)$$

$$CHANGED(gong_{t+1}) \supset Xdancing_{t+1} \quad (52)$$

Suppose this time that the ballerina has entered into the stage, while the gong was sounding from an earlier struck. Without silencing it, the gong is struck again. Under this formulation, the ballerina would start dancing before she could hear a gong stroke. Thus, the behaviour of TAL is different from all the other causal approaches in this point, since indirect effects repeating the same truth (*gong* in the example) are not occluded.

A second important difference has to do with the way in which pertinence and occlusion are understood. On the one hand, occlusion of a complex formula is defined for occluding the *consequent* of a causal rule, disabling inertia for *all* the occurring fluents. In some cases, it could perhaps be enough with requiring only some of them to be occluded in order to allow a change of value in the whole formula. On the other hand, pertinence of a complex formula is mainly thought for the *antecedent* of a causal rule, where, in TAL case, occlusion is never used. The condition of the rule is pertinent when *at least one* of the occurring fluents is so. As we saw with formula (40), it would not make sense requiring all of the occurring fluents to be pertinent.

5 Discussion

We have presented a complete approach for reasoning about actions where the introduction of causality is finally understood as an interesting feature to be used, in the same level than truth information for each fluent. As we have seen, from the comparison to other causal action approaches, this concept is either not present in the formalization (as happens with McCain and Turner Causal Theories) or it is not used with this intention, as happens in Lin's formalization of the suitcase domain or in TAL formalization for some cases of the gong example.

As an important outcome of this work, we have obtained formal comparisons among our logic and three main used causal approaches, from which lateral conclusions can be directly obtained. For instance, from the fact that AL^2 can be encoded into Causal Theories, and from the exhaustive comparison studies from AL^2 to Lin's approach and to TAL, an encoding of these two approaches into Causal Theories can be easily outlined. This fact seems to point out that, for the existing causal approaches, the formalism of Causal Theories can be used as an alternative to Circumscription, being a general nonmonotonic framework. However, it also seems that Causal Action Theories lack from additional information about nature of change, which is present, for instance, in all the other approaches in the shape of *pertinence*, *caused* atoms or *occlusion*.

Acknowledgements Very special thanks to Vladimir Lifschitz for his helpful technical orientations in an earlier version of this work. We would like to thank Ale Provetti for his valuable help, and Chitta Baral for useful comments. Thanks to Jose M. Barja, Alvaro Barreiro and David Lorenzo for their advice. This research is partially supported by the Government of Spain, grant PB97-0228 and in part by the Government of Galicia (Spain) grant 10501B98.

References

[Costello, 1997] T. Costello. Change, change, change: three approaches. In C. S. Mellish, editor, *Proc. of the Intl. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 1426–1431. Morgan Kaufmann, 1997.

[Denecker *et al.*, 1998] M. Denecker, D. Theseider, and K. van Belleghem. An inductive definition approach to ramifications. *Linkoping Electronic Articles in Computer and Information Science*, 3(7), 1998. URL: http://www.ep.liu.se/ea/cis/1998/007/.

[Doherty *et al.*, 1998] P. Doherty, J. Gustafsson, L. Karlsson, and J. Kvarnström. Temporal action logics language specification and tutorial. *Linköping University Electronic Press, Series in Computer and Information Science*, 15, 1998.

[Giunchiglia and Lifschitz, 1998] E. Giunchiglia and V. Lifschitz. An action language based on causal explanation: preliminary report. In *Proc. of the AAAI-99*, 1998.

[Gustafsson and Doherty, 1996] J. Gustafsson and P. Doherty. Embracing occlusion in specifying the indirect effects of actions. In *Proc. of the 5th Intl. Conf. on Principles of Knowledge Representation and Reasoning*, 1996.

[Lifschitz, 1993] V. Lifschitz. *Circumscription*, volume 3, pages 298–352. Oxford University Press, 1993.

[Lifschitz, 1997] V. Lifschitz. On the logic of causal explanation. *Artificial Intelligence Journal*, 96:451–465, 1997.

[Lin, 1995] F. Lin. Embracing causality in specifying the indirect effects of actions. In C. S. Mellish, editor, *Proc. of the Intl. Joint Conf. on Artificial Intelligence (IJCAI)*, Montreal, Canada, August 1995. Morgan Kaufmann.

[Lin, 1996] F. Lin. Nondeterminism in causal theories of actions. In *Proc. of the AAAI-96*, pages 670–676, 1996.

[McCain and Turner, 1995] N. McCain and H. Turner. A causal theory of ramifications and qualifications. In C. S. Mellish, editor, *Proc. of the Intl. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 1978–1984, Montreal, Canada, August 1995. Morgan Kaufmann.

[McCain and Turner, 1997] N. McCain and H. Turner. Causal theories of action and change. In *Proc. of the AAAI-97*, pages 460–465, 1997.

[McCain and Turner, 1998] N. McCain and H. Turner. Satisfiability planning with causal theories. In *Proc. of the 6th Intl. Conf. of Knowledge Representation and Reasoning*, 1998.

[Medtool, 1995] The medtool project and related activities are described in web documents http://www.dc.fi.udc.es/ai/medtool.html, 1995.

[Sandewall, 1989] E. Sandewall. Filter preferential entailment for the logic of action in almost continuous worlds. In C. S. Mellish, editor, *Proc. of the Intl. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 894–899. Morgan Kaufmann, 1989.

[Sandewall, 1994] E. Sandewall. *Features and Fluents*. Oxford Science, 1994.

[Thielscher, 1997] M. Thielscher. Ramification and causality. *Artificial Intelligence Journal*, 1-2(89):317–364, 1997.

[Turner, 1999] H. Turner. A logic of universal causation. *Artificial Intelligence Journal*, 1999. (to appear).