Conditional Logic of Actions and Causation

Laura Giordano

Dipartimento di Informatica

Università del Piemonte Orientale "A. Avogadro"

Corso Borsalino 54 - 15100 Alessandria, Italy, E-mail: laura@mfn.unipmn.it

Camilla Schwind

MAP- CNRS, Luminy

École d'architecture de Marseille

187 Avenue de Luminy, Case 901 13288 Marseille Cedex 9, France,

E-mail: schwind@lim.univ-mrs.fr

Abstract

In this paper we present a new approach to reasoning about actions and causation which is based on a conditional logic. The conditional implication is interpreted as causal implication. This makes it possible to formalize in a uniform way causal dependencies between actions and their immediate and indirect effects. The proposed approach also provides a natural formalization of concurrent actions and of the dependency (and independency) relations between actions. The properties of causality are formalized as axioms of the conditional connectives and a non-monotonic (abductive) semantics is adopted for dealing with the frame problem.

1 Introduction

Causality plays a prominent role in the context of reasoning about actions, as the ramification effects of actions can be regarded as causal dependencies. In this context, causal rules are intended to express causal dependencies among fluents, and, intuitively, their being directional makes them similar to inference rules: if we are able to derive α then we can conclude β. The necessity for and the usefulness of causal rules has been widely recognized in the literature [2, 29, 26, 42]. Many approaches for reasoning about actions have been proposed which allow causal dependencies to be captured [10, 27, 30, 42, 2]. Schwind [38] has studied how causal inferences have been integrated and used in action theories by analyzing four formalisms, which are approaches to action and causality, and comparing them with respect to criteria she established for causality. Namely, the article analyses Lin's approach [26, 27], McCain and Turner's causal theory for action and change [30, 44], Thielscher's theory of ramification and causation [42], and Giordano, Martelli and Schwind's dynamic causal action logic [10].

More recently, Zhang and Foo [45, 4, 5] propose to extend propositional dynamic logic, where actions are modalities, by introducing modalities which are propositions. Sentence "ϕ causes ψ" is represented by the formula $[\phi]\psi$, where $[\phi]$ is a new modality. Note that this representation corresponds to a conditional logic approach, since the EPDL formula $[\phi]\psi$ can be interpreted as the conditional formula $\phi > \psi$. Zhang and Foo's approach has the merit of providing a clean representation of causation as well as a uniform representation of direct and indirect effects of actions.

In this paper we propose an approach to causality based on conditional logics. Causality is represented by a binary logical operator, the conditional operator $>$. A conditional formula $A > B$ is intended to model the causal law: "A causes B".

The properties of causality (as, for instance, those discussed in [38]) are reflected in the axiomatization of the conditional operator.

Traditionally, considering a conditional as a causal implication has frequently attracted the attention of researchers in conditional logic and in AI ([21, 22, 33]). Bennet in [1] proposes a counterfactual analysis of causation which relies on a distinction between *event* and *fact* or *state of affairs* theories of causation. Causality is also a very important concept in the framework of action systems

[29, 26, 42, 38, 11, 45, 15]. In this context, causal implication can occur between different types of assertions:

- An action can cause a fact to become true or
- A fact can cause another fact

In the first case, the causal implication defines also a state transition: if the action is executed then the results caused by it will become true in the "next state". On the other hand, for the second, we assume that the caused fact does not lead to a next state, but it produces modifications on the current state: caused facts are regarded as indirect (or ramification) effects of actions. So we can also find Bennet's distinction in theories of actions. Of course, both, the prerequisite and the consequence of a causal law can be more complex. Two actions can cause together a fact. An action can cause a fact provided that another fact holds. Actions and facts (or more general, formulas) can also cause other causal implications. The only restriction we adopt is that a causal implication cannot itself cause other facts or formulas.

Since we want to model causality in the context of reasoning about actions, our objective is to develop an integrated model of actions and causality: we aim at capturing causal consequences of actions and causal consequences of facts (or formulas) by one single conditional operator: the causal law "A causes B" is represented by the conditional formula $A > B$ and the action law "action a causes proposition C" is represented by the conditional formula $do(a) > C$, where $do(a)$ is a special atomic proposition associated with each action a. This uniform representation of the causal relationship between actions and their results as well as between facts and their effects gives us a great flexibility for handling both concepts in a simple way when representing actions. For example, in this setting, concurrent execution of actions is naturally modelled by conjunctions of the form $do(a_1) \wedge \ldots \wedge do(a_n)$ in the antecedents of conditionals. It is also very natural to express dependency (and independency) relations between actions and actions, actions and propositions, etc.

In this paper we define the properties of the causality operator by introducing suitable axioms which rule the conditional implication. Our causality operator turns out to be non-monotonic and weaker than the one proposed in [45], as it does not entail material implication, which is accepted in [45].

In the following subsection we provide some motivations for the properties we have chosen for causality.

Motivations for the axiom system

We have not adopted a standard conditional logic like Lewis system VCU [31, 23], but we have rather chosen the axioms of this logic which, in our opinion, represent wanted properties of causality and omitted other axioms which instead represent unwanted (or, at least, doubtful) properties. Moreover, we have introduced one axiom, CE, which is new for conditional logic. And we have a modal conditional logic, since in addition to the conditional operator, we have a modal operator for representing general world laws (supposed to be true in every state).

The axiom system we define is motivated by its representational properties as well as by its logical properties. From the action and causality theory viewpoint, our choice is motivated by the following considerations:

1. Causality is certainly NOT reflexive. We assume that no fact should cause itself, unless it is a tautology. Observe, that $A > A$ is not acceptable for the causal laws in which the effects are not simultaneous with their causes. And we do not want to postulate that effects must be simultaneous with their causes: an action causes its results, which become true after its execution. Therefore, the identity axiom $(ID) A > A$ does NOT belong to our axiom system. Moreover, in presence of the identity axiom supra-classicality would be derivable, i. e. from $\vdash A \rightarrow B$, $\vdash A > B$ can be deduced. We think it is not very intuitive to assume that any tautological implication $A \rightarrow B$ defines a causal link $A > B$.

2. We are ready to accept that, for all tautologies A and for all formulas B, $B > A$ holds, that is, a tautology is caused by everything. We do not regard this inclusion to be harmful as all tautologies hold anyhow in each state of the world. Moreover, this property follows from rule (RCK) (see below), which is one of the basic inference rules of conditional logic and which we certainly want to keep.

3. (CS) $A \wedge B \rightarrow (A > B)$ should not be a property of causal implication: A and B could both hold conjunctively without A being a cause of B. From the fact that in the current state ”there is sun” and ”I have a cold” we do not want to conclude that: ”sun causes cold”.

4. Monotonicity
   Classical implication is monotonic. That means from $A \rightarrow B$ we can deduce $A \wedge C \rightarrow B$. But we have many reasons to assume, that causal implication is NOT monotonic. For example, from *raining causes wet*, it does not follow that *raining and being under my umbrella causes wet*. We can obviously imagine many examples of this form: if a fact causes another fact then there might

very frequently exist another fact which hinders the causal result to be produced. Therefore, we think, that the law $A > B \to (A \wedge C) > B$ should not be a property of causal implication.

5. Nevertheless, under some preconditions, we may accept or even want a weakened form of monotonicity. In the case when the added precondition does not contradict the original precondition, we want to continue to conclude the causal consequence of a formula after adding the new precondition. For taking into account this property, we use Axiom $(CV)\neg(A > \neg C) \wedge (A > B) \to (A \wedge C > B)$, which allows to weaken a precondition of a causal law thus introducing a weak form of monotonicity. As we will see, axiom (CV) is very useful for describing interactions (such as independency) among actions and facts.

6. Axiom (MP) $(A > B) \to (A \to B)$ has been widely used implicitly or explicitly in the literature on causality [29, 26, 6, 45, 15]. Makinson discusses its use for nonmonotonic reasoning [28]. We do not want to include MP because it allows to derive material (classical) implication from causal implication and annihilates its temporal aspect. An action or a fact has a causal consequence but this caused consequence may be delayed: the action occurs or the fact holds and this makes its caused results true. Including (MP) would lead to unwanted conclusions when contraposition, modus tollens, monotonicity or any other property of classical logic is applied to the material implication which can be derived by (MP) from the causal implication. For this reason we do not include (MP) in our logic.

   To clarify the problem, let us refer to the *suitcase problem* presented in [26]: There is a suitcase with two locks and a spring loaded mechanism which will open the suitcase when both of the locks are in the up position. Consider the following causal law: "lock 1 open and lock 2 open causes the suitcase to open" ($up_1 \wedge up_2 > open$). Assume that in the initial state lock 1 is up and lock 2 is down and the suitcase is closed ($\{up_1, \neg up_2, \neg open\}$). We would expect that flipping lock 2 in the up position would cause the suitcase to open (assuming that lock 1 persists in the up position). This solution can be obtained by applying the causal law above to conclude $open$ from $up_1 \wedge up_2$. If we accept (MP), we can derive $up_1 \wedge up_2 \to open$, which is equivalent to $\neg open \wedge up_2 \to \neg up_1$. Then we could get the solution that flipping lock 2 in the up position causes lock 1 to flip in the down position (assuming that $\neg open$ persists). This is certainly an incorrect conclusion.

   Observe that (MP) can either lead to unwanted conclusions or not, depending on the way causal laws are used to compute the immediate and indirect effects of actions. As we will see, in our approach all the formulas in the causal theory and their consequences are used to compute the immediate and indirect effects of actions. Using the material implications, that are derivable from causal laws in the presence of (MP), to compute indirect effects of actions gives the unintended outcomes.

   Moreover, axiom (MP), which is $(A > B) \to (A \to B)$, makes it possible to derive $(A > B) \wedge \neg B \to (A \to C)$ for any formula C. Hence by using MP, we get that for any A and B, such that "A causes B" holds, whenever B is not true, A implies everything. To illustrate this property, let us consider the following set of sentences:

(a) "raining causes Tim to become wet"

(b) "Tim does not become wet"

From this we conclude then "If it rains the moon is made of green cheese"

The second reason for not including (MP) in our axioms, comes from the fact that we do not want to exclude causal relations in which the effects are not simultaneous with the causes (though our formalism does not represent explicitly all the intermediate states). In such cases the material implication between the causes and effects appears to be unintended. Given the sentence: "Yellow fever causes death", which says that death is caused by yellow fever (but possibly with a delay with respect to the contraction of illness), we are not ready to accept the sentence "Yellow fever implies death" or "Not death implies not yellow fever".

7. Given a causal law A causes B, we want to be able to derive B whenever A holds. This is obviously a useful property of causal implication and we want to retain it even without keeping (MP). For that purpose, we introduce an new axiom, (CE), which is weaker than (MP) and allows for a sort of modus ponens for causal implication: $(ca > (A > B)) \wedge (ca > A) \rightarrow (ca > B)$. (CE) allows to deduce that B holds after an action ca whenever ca causes $(A > B)$ and ca causes A. Note, that (CE) is a logical consequence of (MP) in the system CK.

8. Reasoning by cases is an important property of causal inference. Consider a circuit with two switches and a lamp. If we know that toggling one of the two switches causes the lamp not to be alight, and we do not know which of the switches has been toggled, we only know $sw_1 \vee sw_2$, then we want to be able to derive $\neg light$. For that we need axiom (CA), $(A > C) \wedge (B > C) \rightarrow ((A \vee B) > C)$, which allows to deduce $sw_1 \vee sw_2 > \neg light$ from $sw_1 > \neg light$ and $sw_2 > \neg light$.

9. The $\Box$ operator is introduced for expressing that a law always holds (in every state of the world). This also makes it possible to represent domain constraints. $\Box$ is a modal operator and has the properties of the modal system $(S4)$ (reflexivity and transitivity).

10. Domain constraints must hold in all states of the world and more specifically after each action execution. "True in all states" is expressed by means of $\Box$. The (MOD) axiom $\Box A \rightarrow (do(a) > A)$ relates the modal operator $\Box$ to the causal operator $>$ and makes it possible to derive that a constraint, which is true in all states, is also true after the execution of an action a.

The paper is organized as follows. In the next section, we introduce the conditional action logic and we show its completeness and decidability. The decidability result gives also a complexity bound. In section 3 we show how this logic is used to formalize action theories and we illustrate some properties of our formalism by examples. In section 4 we compare our approach to related work and section 5 concludes.

2 The Causal Action Logic AC

The language $\mathcal{L}_>$ of our action logic is that of propositional logic $\mathcal{L}$ augmented with a conditional operator $>$ and the modal operator $\Box$. The set of propositional variables in $\mathcal{L}_>$, Var, includes the set $\{do(a) : a \in \Delta_0\}$, where Δ_0 is a set of *elementary actions* including the "empty" action ϵ. In the following, we assume that the conditional $>$ has higher precedence than the material implication $\rightarrow$ and that all other boolean connectives have higher precedence than $>$. Also we assume that $>$ is right associative.

Formulas are defined as usual except that we assume that only propositional formulas in $\mathcal{L}$ can occur as antecedents of conditional formulas. Hence, in a conditional $A > B$ nested conditionals (or modalities) can only occur in B, and we do not allow statements like $(C > D) > B$. As we mentioned in the introduction, we do not admit that a causal law causes anything. In fact, we believe that the interpretation of the statement $(C > D) > B$ as a causal law is not straightforward.[1]

Intuitively, $\Box A$ means that A necessarily holds, i.e. holds in every state of the world. $A > B$ means that A causes B. In particular, when A is an action predicate $do(a)$, $do(a) > B$ means that executing action a causes B to hold. Let us point out, that $do(\epsilon) > B$ is NOT equivalent to B (and therefore does not entail B).

Let ca represent a finite conjunction of action formulas $do(a_1) \wedge \ldots \wedge do(a_n)$ for $a_i \in \Delta_0, 1 \leq i \leq n$. We introduce the following axiom system for logic AC.

Definition 1 [AC] The conditional logic AC is the smallest logic containing the following axioms and deduction rules:

(CLASS) All classical propositional axioms and inference rules

(CV) $\neg(A > \neg C) \wedge (A > B) \rightarrow (A \wedge C > B)$

(CA) $(A > C) \wedge (B > C) \rightarrow ((A \vee B) > C)$

(CE) $(ca > B) \wedge (ca > (B > C)) \rightarrow (ca > C)$, where $C \in \mathcal{L}$ a propositional formula

(MOD) $\Box A \rightarrow (ca > A)$

(K) $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$

(4) $\Box A \rightarrow \Box\Box A$

(T) $\Box A \rightarrow A$

(RCEA) if $\vdash A \leftrightarrow B$, then $\vdash (A > C) \equiv (B > C)$ where A and B are propositional formulas.

(RCK) if $\vdash A_1 \wedge \ldots \wedge A_n \rightarrow B$, then $\vdash (C > A_1) \wedge \ldots \wedge (C > A_n) \rightarrow (C > B)$ for any propositional formula $C \in \mathcal{L}$

(NEC) if $\vdash A$ then $\vdash \Box A$

[1] A possible interpretation is: "under the hypothesis that $C > D$ holds, B is caused", which, however, is a "hypothetical" interpretation of the (external) conditional, rather than a "causal" interpretation.

Note that all axioms and inference rules are standard in conditional logics and, in particular, they belong to the axiomatization of Lewis's logic VCU (see [23]). As a difference, we have excluded several of the standard axioms of conditional logics such as (ID), (MP) and (CS) based on the motivations we have presented in the previous section.

Referring to the discussion in section 1, let us see that (ID) $A > A$ together with (RCK) entails the rule of supra-classicality:

$$if \vdash A \rightarrow B \;\; then \vdash A > B$$

by

(1) $A \rightarrow B$ hypothesis
(2) $(A > A) \rightarrow (A > B)$ from (1) by RCK
(3) $(A > A)$ (ID)
(4) $A > B$ from (2) and (3) by (CLASS)

Observe also that (MP) $(A > B) \rightarrow (A \rightarrow B)$, which is not an axiom of our logic, allows a contrapositive use of causal laws. In fact, (MP) is classically equivalent to $(A > B) \wedge \neg B \rightarrow \neg A$ and it entails, by classical inference,

(5) $(A > B) \wedge \neg B \rightarrow (A \rightarrow C)$

for any formula C. (5) states that whenever causal law $A > B$ belongs to a theory, A implies everything, provided that B does not hold. We think that this is not acceptable in our logic and it does not hold in it.

Let us now explain the axioms of AC. Our logic contains axioms and rules of the standard minimal conditional logic (CK) as well as additional axioms (CA) and (CV), which are also standard in conditional logic systems. Moreover, we have added one new axiom (CE), which is not standard in conditional logics. (CE) allows action laws and causal laws to interact, it provides the chain effects between causal laws and action laws. (CE) says that the causal consequences of action effects are in turn action effects: If the execution of the concurrent actions ca causes B to become true and if ca also causes the causal implication $(B > C)$ then ca also causes C. (CE) weakens (MP) as it is clear from the following formulation of (CE)

$$(ca > (B > C)) \rightarrow ((ca > B) \rightarrow (ca > C)),$$

which can be obtained from (MP) by (RCK). (CE) has similarities with the property of transitivity ($TRANS_{>}$) of $>$: $(ca > B) \wedge (B > C) \rightarrow (ca > C)$. For standard conditional logic with reflexivity (ID), adding ($TRANS_{>}$) would collapse the conditional implication to material implication. But this is not the case for our causal action logic AC, since identity $A > A$ is not an axiom. (CE) requires that the causal law $B > C$ holds as a consequence of the execution of the actions ca. As we will see in our action theory causal laws do not necessarily hold in all possible states, as they may have preconditions which make them hold in some states only. As an example of a causal law with precondition consider the following one: $\Box(at(y, r) \rightarrow (at(z, r) > at(y, next(r))))$ which says that if block y is at r then moving block z to position r causes y to move to a next position (see example 4 below). It should also be noted that the instance of (CE) with $ca = do(\epsilon)$, $(do(\epsilon) > B) \wedge (do(\epsilon) > (B > C)) \to (do(\epsilon) > C)$ does not entail (MP), since $do(\epsilon) > X$ is not equivalent to X, as pointed out at the beginning of this section. Observe also that in axiom (CE) the formula C is restricted to be a propositional formula.

(MOD), (4) and (T) define the properties of the necessity operator $\Box$, where (MOD) defines the relationship between the conditional and the modal operator. In particular, (4) and (T) say that $\Box$ has S4-properties. The three axioms allow to deduce $\Box A \to (ca_n > (ca_{n-1} \ldots > (ca_1 > A) \ldots)$ for any finite sequence of concurrent actions $ca_1, \ldots ca_n$ ($n \geq 0$) including the empty sequence, meaning that a formula A which is *always* true is also true after the occurrence of any finite sequence of concurrent actions. The deduction is by induction on n:

For $n = 1$, we have $\Box A \to (ca_1 > A)$ (MOD)
For $n > 1$:
(1) $\Box A \to \Box\Box A$ (4)
(2) $\Box\Box A \to (ca_n > \Box A)$ (MOD)
(3) $\Box A \to (ca_{n-1} > \ldots (ca_1 > A) \ldots)$ induction hypothesis
(4) $(ca_n > \Box A) \to (ca_n > (ca_{n-1} > \ldots (ca_1 > A) \ldots)$ from (3) by (RCK)
(5) $\Box A \to (ca_n > (ca_{n-1} > \ldots (ca_1 > A) \ldots)$ from (1), (2) and (4) by (CLASS)

So, the subsequent occurrence of actions determines subsequent states of the world according to time, although time is not represented explicitly in our formalism. MOD requires action execution formulas ca as causal implicant. $\Box A \to (B > A)$ is not valid in our logic when B is not a conjunction ca of action execution formulas. We restricted MOD to action formulas since we do not think that general laws, which hold in every state of the world, should be causal consequences of any formula. Instead they should hold after any execution of actions.

Observe that, as a difference with VCU in which a $\Box$ modality is defined through the conditional operator, as $\Box A \equiv (\neg A > \bot)$, we have introduced an independent modality $\Box$, characterized by (4) and (T) and we have related it to the $>$ operator through the interaction axiom (MOD). This makes this causal logic weaker than the one we introduced in [13], which instead adopts the definition of $\Box$ through the conditional connective. While in [13] we wanted to stay as close as possible to a standard conditional logic like VCU, in this paper we have preferred to include in the logic the less axioms as possible, namely, those axioms which are motivated by properties of causality. In particular, we have neither included the definition of $\Box$ in terms of the conditional operator $>$, nor the axiom $\Diamond A \to \Diamond\Box A$ (which would give S5 structures, rather then S4 structures). Their introduction, in fact, is not needed as it cannot be motivated by properties of causality.

Entailment $\vdash$ is defined as usual and given a set of formulas E, the deductive closure of E is denoted by $Th(E)$. AC is characterized semantically in terms of selection function models.

Definition 2 An AC-structure M is a quadruplet $\langle W, f, R, [[\,]] \rangle$, where W is a non-empty set, whose elements are called possible worlds, f, called *the selection function*, is a function of type $\mathcal{L} \times W \rightarrow 2^W$, $R \subseteq W \times W$ is the *accessibility relation* for $\Box$, $[[\,]]$, called the evaluation function, is a function of type $\mathcal{L}_> \rightarrow 2^W$ that assigns a subset of W, $[[A]]$ to each formula A. Let us note $R(w) = \{w' : R(w, w')\}$. The following conditions have to be fulfilled by $[[\,]]$:

(1) $[[A \wedge B]] = [[A]] \cap [[B]]$;
(2) $[[\neg A]] = W - [[A]]$;
(3) $[[A > B]] = \{w : \; f(A, w) \subseteq [[B]]\}$
(4) $[[\Box A]] = \{w : \; R(w) \subseteq [[A]]\}$.

Using the standard boolean equivalences, we obtain $[[A \vee B]] = [[A]] \cup [[B]]$, $[[A \rightarrow B]] = (W - [[A]]) \cup [[B]]$, $[[\top]] = W$, $[[\bot]] = \emptyset$.

We assume that the selection function f satisfies the following properties which correspond to the axioms of our logic AC:

(S-RCEA) if $[[A]] = [[B]]$ then $f(A, w) = f(B, w)$

(S-CV) if $f(A, w) \cap [[C]] \neq \emptyset$ then $f(A \wedge C, w) \subseteq f(A, w)$

(S-CA) $f(A \vee B, w) \subseteq f(A, w) \cup f(B, w)$

(S-CE) if $f(ca, w) \subseteq [[B]]$ then $ValProp(f(ca, w)) \subseteq ValProp(f(B, f(ca, w)))$

(S-MOD) $f(ca, w) \subseteq R(w)$

(S-4) if $R(w, w')$ and $R(w', w'')$ then $R(w, w'')$, for all $w, w', w'' \in W$

(S-T) $R(w, w)$, for all $w \in W$

where ca is a finite conjunction of actions $do(a_1) \ldots do(a_n)$ for $a_i \in \Delta_0$ and $f(B, f(ca, w))$ represents the set of worlds $\{z \in f(B, x) \;:\; x \in f(ca, w)\}$. Moreover, given a set of worlds S, $ValProp(S)$ is the set of all the propositional valuations at the worlds in S.

We say that a formula A is true in an AC-structure $M = \langle W, f, R, [[]]\rangle$ if $[[A]] = W$. We say that a formula A is AC-valid ($\models A$) if it is true in every AC-structure. Given a AC-structure M, a set of formulas S and a formula A, $S \models_M A$ means that for all $w \in M$ if $w \in [[B]]$ for all $B \in S$, then $w \in [[A]]$.

The above axiom system is sound and complete with respect to the semantics

Theorem 1 $\models A$ *iff* $\vdash A$

The completeness proof is shown by the canonical model construction [39] and can be found in Appendix A. Moreover, the axiomatization of the logic AC is consistent and the logic is decidable. The *consistency* of the axiomatization comes from the fact that, if we replace the modality $\Box$ with the formula $\neg A > \bot$ in all axioms, we get a subset of the axioms of VCU, which is known to be consistent. This also shows that the logic AC is ”non-trivial” in some sense.

For the proof of *decidability* we refer to the Appendix B. We only mention that it proves the finite model property for the logic AC, by showing that, if there is a model satisfying a formula F, then there is a *finite* model satisfying it. The decidability proof constructs a model of double exponential size. Hence, this provides an upper bound for the complexity of satisfiability in AC. The problem is non-deterministic double exponential in time (with respect to the number of propositional variables in F).

In logic AC, formula $A > \perp$ is not inconsistent. It is easy to see that according to our semantics, $A > \perp$ is true in a state w of a model M iff $f(A, w) = \emptyset$. If A is an action formula $do(a)$, the intended meaning of $do(a) > \perp$ being true in w is that executing a in state w does not yield any resulting state: executing a is not possible in w! This is a powerful property of our logic. For example, it makes it possible to express that two actions a and b cannot occur together by formulating $do(a) \wedge do(b) > \perp$.

3 Action Theories

So far, we have introduced the logical language for action theories together with its logical axiom system. In this section, we show how this logic is used for describing systems and worlds where actions occur and causality laws hold.

3.1 Domain descriptions

We use atomic propositions $f, f_1, f_2, \ldots \in Var$ for *fluent names*. A *fluent literal*, denoted by l , is a fluent name f or its negation $\neg f$. Given a fluent literal l, such that $l = f$ or $l = \neg f$, we define $|l| = f$. We denote by $ca, ca_1, ca_2, \ldots$ concurrent actions $do(a_1) \wedge \ldots do(a_n)$ for $a_i \in \Delta_0$ (including the single action do(a) for $n = 1$). We will denote by $\mathcal{F}_{\Delta_0}$ the set of all fluent names of the form $do(a)$, for $a \in \Delta_0$ and by $\mathcal{F}$ the set of all fluent names different from $do(a)$, for $a \in \Delta_0$. Moreover, we will denote by Lit_{Δ_0} the set of all fluent literals built from $\mathcal{F}_{\Delta_0}$, and by Lit the set of all fluent literals built from $\mathcal{F}$. Finally, we will denote by small greek letters α, β, ... any formula not containing conditional formulas and by upper case latin letters A, B, ... arbitrary formulas.

Our action theory refers to the same ontology as the *Situation Calculus* [37]. The Situation Calculus represents states of the world (*situations*) as sequences of *actions*, and *fluents* as relations whose truth values vary from state to state. The situation calculus is formulated in first-order logic: Situations are represented with by such as $do(a_1, do(a_2, s))$, while fluents are extended with an extra argument denoting a situation (for instance, $f(do(a_1, do(a_2, s)))$). In our action theory, conditional formulas are used to describe the values of fluents at the states: $do(a_1) > do(a_2) > f$ says that fluent f holds in the state obtained by executing action a_1 and then action a_2 (f is caused by executing a_1 and then a_2).

We define a *domain description* as a tuple $(\Pi, Frame_0, Obs)$.

Π is a set of laws and constraints containing *action laws*, *causal laws*, *precondition laws*, *domain constraints* and *causal independency constraints*.

Action laws have the form:

$$\Box(\pi \rightarrow (do(a) > R)),$$

for an action a with precondition π and effect R: executing action a in a state where π holds *causes* R to hold in the resulting state. It should be noted that our theory allows for complex action effects, namely an action can have a causal formula $A > B$ as result as it will be illustrated by the example 3. An action law with no precondition, i.e. $\pi = true$, simply becomes $\Box(do(a) > R)$.

Causal laws have the form:

$$\Box(\pi \rightarrow (\alpha > B)),$$

meaning that "if π holds, then α *causes* B".

Precondition laws have the form:

$$\Box(\pi \equiv \neg(do(a) > \bot)),$$

meaning that "action a is *executable* iff π holds".

Domain constraints include formulas of the form:

$$\Box\alpha,$$

(meaning that "α always holds").

Causal independency constraints have the form:

$$\Box(\neg(A > \neg B)),$$

meaning that A does not cause $\neg B$ (that is, B might be true in a possible situation caused by A).

In particular, when the above constraints concern action execution, we have

$$\Box\neg(do(a) > \neg do(b)),$$

meaning that the execution of action a does not prevent action b from being executed (does not interfere with its execution). Note that as a consequence of this constraint we have, by (CV), that

$$(do(a) > R) \rightarrow (do(a) \wedge do(b) > R),$$

namely, the effects of action a are also effects of the concurrent execution of a and b, as a does not interfere with b. Moreover, by taking $R = \bot$, we get:

$$(do(a) > \bot) \rightarrow (do(a) \wedge do(b) > \bot),$$

meaning that if a is not executable it cannot be executed concurrently with b.

All domain description laws in Π are of the form $\Box A$, since they hold in all states. By using axiom (MOD), from $\Box A$, we can deduce $do(a_1) > \ldots > do(a_n) > A$, for any finite sequence of actions $a_1, \ldots a_n$ (or $ca_1 > \ldots > ca_n > A$ for any finite sequence of concurrent actions $ca_1, \ldots ca_n$), as we have shown in the preceding section. In the following we will frequently use this form of the laws in Π.

$Frame_0$ is a set of pairs $(f, do(a))$, where $f \in \mathcal{F}$ is a fluent and $a \in \Delta_0$ is an elementary action, meaning that f is a *frame fluent* for action a, that is, f is a fluent to which persistency applies when a is executed. Fluents which are non-frame with respect to a do not persist and may change value in a nondeterministic way when a occurs[2].

[2]Observe that the fluents in $\mathcal{F}_{\Delta_0}$ will not be subject to persistency and they can take any value in a state independently from its value in the preceeding state

The set $Frame_0$ defines a sort of *independence* relationship between elementary actions and fluents. It is closely related to dependency (and influence) relations that have been used and studied by several authors including Thielscher [42], Giunchiglia and Lifschitz [14], and Castilho, Gasquet and Herzig [18]. In the next section, when addressing the frame problem in a non-monotonic formalism, we will make use of $Frame_0$ for defining persistency rules of the form $l \rightarrow (do(a) > l)$ (or $l \rightarrow (ca > l)$, for a concurrent action ca) for every literal l, such that $(|l|, a) \in Frame_0$. The meaning of such a rule is that "if fluent l holds at a state, it will persist after the execution of action a in that state".

As we will see, these persistency rules are introduced state by state and they behave like *defaults*: they belong to an "action extension" whenever no inconsistency arises. The $Frame_0$-relationship is extended to concurrent actions. Let us denote by $Frame$ the extension of $Frame_0$ to concurrent actions, which is the smallest set satisfying the following two conditions:

1. $Frame_0 \subset Frame$,
2. If $(f, do(a_1)), \ldots, (f, do(a_n)) \in Frame$ then $(f, do(a_1) \wedge \ldots \wedge do(a_n)) \in Frame$.

As mentioned above, *states* in our action theory are represented by action sequences. Each action execution leads from one state to a new state. Therefore a state is identified by the sequence $ca_1, \ldots, ca_n$ of (possibly concurrent) actions which lead to it from the initial state. We will denote by $S_{ca_1,\ldots,ca_n}$ the set of fluent literals which hold at the state obtained by executing the actions $ca_1, \ldots, ca_n$ in the sequence. As we will see in the next section, the set of facts which hold at a state depend on the extension of the domain description that we are considering. Once the extension is fixed, we will refer to the set $S_{ca_1,\ldots,ca_n}$ as a state. It is the set of the literals l such that the conditional formula: $ca_1 > \ldots > ca_n > l$ holds in the extension. We say that literal l holds at the state obtained by executing actions $ca_1, \ldots, ca_n$. More generally, we will say that a formula α holds at a state obtained by executing actions $ca_1, \ldots, ca_n$ in a given extension when the conditional formula $ca_1 > \ldots > ca_n > \alpha$ holds in the extension. We have a special empty action ϵ to represent the *initial state* $do(\epsilon) > \alpha$ means that α holds in the initial state.

Obs is a set of observations about the value of fluents in different states. They are formulas of the form: $ca_1 > \ldots > ca_n > \alpha$ (where each ca_i is a possibly concurrent action formula of the form $do(a_1) \wedge \ldots \wedge do(a_n)$), meaning that α holds after the concurrent execution of the actions in ca_1, then those in ca_2, ..., then those of in ca_n. Observations about fluents in the initial state have the form $do(\epsilon) > \alpha$. In the following, when identifying a state with an action sequence $ca_1, \ldots, ca_n$, we will implicitly assume that $ca_1 = do(\epsilon)$.

Observe that, given a state $S_{ca_1,\ldots,ca_n}$, $ca_1, \ldots, ca_n$ are the actions which have been explicitly executed to reach the state. We can call them external actions, and they do not include the hidden actions which might be automatically triggered as effect of the execution of other actions (consider, for instance, the causal law $do(a_1) > do(a_2)$: the execution of action a_1 triggers the execution of action a_2). Though our language does not provide an explicit representation of time, as we abandon (MP), time can be embedded in the operator $>$. Given the properties of $>$ we explicitly model a delay between happening of an action and occurrence of its effects, for external actions, while we do not model explicitly the delay between causes and their effects in causal laws as well as in those actions which are automatically triggered by other actions.

Let us explain the reason why we have introduced the empty action to denote the initial state. Though we have not included (MP) in our axioms, we want that, for each causal law $A > B$ which holds in a state, if A holds then B also holds. For all the states except the initial one this is enforced by axiom (CE). In fact, if A holds at state $S_{do(a_1)\ldots,do(a_n)}$, i.e. $do(a_1) > \ldots > do(a_n) > A$ holds, then by (CE) we also have that $do(a_1) > \ldots > do(a_n) > B$ holds. Similarly, for the initial state, from $do(\epsilon) > A$, by (CE), from the causal law $do(\epsilon) > (A > B)$, we get $do(\epsilon) > B$. Otherwise, it would be possible to have an initial state in which $A \wedge \neg B$ holds though the causal law $A > B$ also holds in that state.

Sometimes, when we do not want to consider observations, we will use the notion of *domain frame*, which is a pair $(\Pi, Frame_0)$.

Let us consider the following (benchmark) example (from [25]) treated by almost all action theories, which formalizes an electrical circuit with two serial switches.

Example 1 *There is a circuit with two switches and a lamp. If both switches are on, the lamp is alight. One of the switches being off causes the lamp not to be alight. There are two actions of toggling each of the switches. The domain description is the following (for $i = 1, 2$):*

Π: $\Box(\neg sw_i \rightarrow (do(tg_i) > sw_i))$ $\quad$ $\Box(sw_i \rightarrow (do(tg_i) > \neg sw_i))$
$\quad \Box(sw_1 \wedge sw_2 > light)$ $\quad$ $\Box(\neg sw_i > \neg light)$
$\quad \Box(\neg(do(tg_1) > \neg do(tg_2)))$ $\quad$ $\Box(\neg(do(tg_2) > \neg do(tg_1)))$
Obs: $\quad do(\epsilon) > (\neg sw_1 \wedge \neg sw_2 \wedge \neg light)$
$Frame_0 = \{(f, a) : a \in \Delta_0, f \in \mathcal{F}\}$.

The first two rules in Π describe the immediate effects of the action of toggling a switch. The third and forth rules are causal laws which describe the dependencies of the light on the status of the switches. The last two laws are constraints saying that the two actions tg_1 and tg_2 do not interfere. All fluents are supposed to be persistent and the actions tg_1 and tg_2 are independent. As we will see, from the above domain description we can derive $do(tg_1) > \neg light$, $do(tg_1) > do(tg_2) > light$ and $do(tg_1) \wedge do(tg_2) > (sw_1 \wedge sw_2 \wedge light)$ (as actions $do(tg_1)$ and $do(tg_i)$ are independent).
Let us see the derivation of the concurrent execution of tg_1 and tg_2.

(1) $(do(tg_1) > sw_1) \wedge \neg(do(tg_1) > \neg do(tg_2))$
$\quad \rightarrow (do(tg_1) \wedge do(tg_2) > sw_1) \quad (CV)$
(2) $(do(tg_2) > sw_2) \wedge \neg(do(tg_2) > \neg do(tg_1))$
$\quad \rightarrow (do(tg_2) \wedge do(tg_1) > sw_2) \quad (CV)$
(3) $(do(tg_1) \wedge do(tg_2) > sw_1) \wedge (do(tg_1) \wedge do(tg_2) > sw_2)$
$\quad \rightarrow (do(tg_1) \wedge do(tg_2) > sw_1 \wedge sw_2)$ from $sw_1 \wedge sw_2 \rightarrow sw_1 \wedge sw_2$ by (RCK)
(4) $\neg sw_1 \wedge \neg sw_2 \rightarrow (do(tg_1) \wedge do(tg_2) > sw_1 \wedge sw_2)$ from (1), (2), (3) and the first two laws in Π

(5) $(do(tg_1) \wedge do(tg_2) > (sw_1 \wedge sw_2 > light)$ from the third law in Π by (MOD)
(6) $\neg sw_1 \wedge \neg sw_2 \rightarrow (do(tg_1) \wedge do(tg_2) > light)$ from (5), (4) by (CE) and $(CLASS)$

Observe that the above derivation is monotonic and it does not make use of any persistency assumption. Note also that we could have avoided introducing $\neg light$ in the initial state, as it can be derived, for instance, from $\neg sw_1$: from $do(\epsilon) > \neg sw_1$ and the forth action law we can derive $do(\epsilon) > \neg light$ by (CE).

Axiom (CA) makes it possible to deduce consequences of actions even when it is not deterministically known which action occurs.

Example 2 *If the temperature is low, then going to swim causes you to get a cold. If you have no umbrella, then raining causes you to get cold. We have the following domain description:*

Π: $\Box(cold \rightarrow (do(swim) > get\_cold))$
$\Box(no\_umbrella \rightarrow (do(rain) > get\_cold))$
Obs: $do(\epsilon) > (cold \wedge no\_umbrella)$
$Frame_0 = \{(f,a) : a \in \Delta_0, f \in \mathcal{F}\}$

From this theory, we can derive $cold \wedge no\_umbrella \rightarrow (do(swim) \vee do(rain) > get\_cold)$ by:

(1) $(do(swim) > get\_cold) \wedge (do(rain) > get\_cold) \rightarrow (do(swim) \vee do(rain) > get\_cold)$ (CA)
(2) $cold \wedge no\_umbrella \rightarrow (do(swim) \vee do(rain) > get\_cold)$ from (1) and the laws in Π by $(CLASS)$ (cut)

Note that this (monotonic) derivation also holds in the domain frame, as we did not use the observations. Taking into account Obs, we obtain from (2) (and Obs)

$$do(\epsilon) > (do(swim) \vee do(rain) > get\_cold).$$

The following example has been discussed by Halpern and Pearl in [17].

Example 3 *Two arsonists drop lit matches in two different parts of a dry forest and each of them causes the trees to start burning. There are two scenarios. In the first either match by itself suffices to burn down the whole forest; in the second scenario, both matches are necessary to burn down the forest. If only one match were lit, the fire would die down. Our formalization of these two scenarios is the following:*

1. Scenario $\Box(do(lit_1) > start\_burn)$ $\quad$ $\Box(do(lit_2) > start\_burn)$
$\Box(start\_burn > burn\_down)$

2. Scenario $\Box(do(lit_1) > start\_burn)$ $\quad$ $\Box(do(lit_2) > start\_burn)$
$\Box(do(lit_1) \wedge do(lit_2) > (start\_burn > burn\_down))$
$\Box\neg(do(lit_1) > \neg do(lit_2))$ $\quad$ $\Box\neg(do(lit_2) > \neg do(lit_1))$

Both scenarios contain the first two causal laws stating that dropping down lit matches causes the forest to start burning. In both scenario, we can derive $do(lit_1) \vee do(lit_2) > start\_burn$ using axiom (CA).
In the first scenario we can deduce

(1) $do(lit_1) > (start\_burn > burn\_down)$ by (MOD)
(2) $do(lit_2) > (start\_burn > burn\_down)$ by (MOD)
(3) $do(lit_1) > burn\_down$ from (1) and by (CE)
(4) $do(lit_2) > burn\_down$ from (2) and by (CE)

In the second scenario, we can deduce $do(lit_1) \wedge do(lit_2) > burn\_down$ by

(5) $\neg(do(lit_1) > \neg do(lit_2))$ independency law
(6) $do(lit_1) > start\_burn)$ action law
(7) $(do(lit_1) > start\_burn) \wedge \neg(do(lit_1) > \neg do(lit_2)) \rightarrow (do(lit_1) \wedge do(lit_2) > start\_burn)$ instance of (CV)
(8) $do(lit_1) \wedge do(lit_2) > start\_burn$ from (7), (6) and (5)
(9) $do(lit_1) \wedge do(lit_2) > (start\_burn > burn\_down)$ causal law
(10) $do(lit_1) \wedge do(lit_2) > burn\_down$ from (9) and (8) by (CE)

In the second scenario $do(lit_1) > burn\_down$ cannot be deduced because $\Box(start\_burn > burn\_down)$ is not a causal law in scenario 2, where it is replaced by the nested causal law. Start burning is not a sufficient cause here to burn down the forest. It is a sufficient cause instead that both arsonists acted, which is expressed by the causal law $\Box(do(lit_1) \wedge do(lit_2) > (start\_burn > burn\_down))$ of scenario 2.

This example illustrates the use of nested causation laws which for this example is crucial. In both scenarios, to drop a lit match has the effect to start a forest fire. In the first scenario, starting a forest fire causes the forest to burn down. In the second scenario starting this causal law is itself conditioned by the two arsonists having dropped down their lit matches.

The following example, taken from [10], involves causal laws with preconditions.

Example 4 *Consider the following scenario, where a number of blocks are in a sequence: when the first block, a, is pushed from the place p_1 to the place p_2, all other blocks move also to the next place. Let $p_1, \ldots p_n$ be places and a, b, c be blocks, and let $push(x, p)$ be the action which consists in pushing the block x from the place p to the next place $next(p)$.*[3]

Π: $\Box(at(x,p) \rightarrow (do(push(x,p)) > at(x, next(p))))$
$\Box(at(y,r) \rightarrow (at(z,r) > at(y, next(r))))$, *for* $z \neq y$
$\Box(at(x,p) > \neg at(x,q))$, *for* $p \neq q$
Obs: $do(\epsilon) > (at(a,p_1) \wedge at(b,p_2) \wedge at(c,p_3))$
$Frame_0 = \{(f,a) : a \in \Delta_0, f \in \mathcal{F}\}$

[3] $next(p)$ is used as an abbreviation. Given the places $p_1, p_2, \ldots$, $next(p_1)$ stands for p_2, $next(p_2)$ stands for p_3, and so on.

The first (action) law says that, if block x is in p, pushing x from the place p moves it to the next place $next(p)$. The second (causal) law says that if block y is at r then moving block z to position r causes y to move to a next position. The third laws causes block x not to be at q if it is at p (different from q).

Given the initial state, it holds that $do(push(a,p_1)) > (at(a,p_2) \wedge at(b,p_3) \wedge at(c,p_4))$: pushing block a from position p_1 to p_2 pushes block b from position p_2 to p_3 and block c from position p_3 to p_4.

Example 5 *There is a bowl of soup. Assuming that initially the soup is not spilled, it is expected that, whenever Mary tries to lift the bowl with one hand, she spills the soup. When she uses both hands, she does not spill the soup.*

$\Box(do(lift\_l) > up\_left)$ $\qquad \Box(do(lift\_r) > up\_right)$
$\Box(up\_left \wedge \neg up\_right > spilled)$ $\qquad \Box(\neg up\_left \wedge up\_right > spilled)$
$\Box(\neg(do(lift\_l) > \neg do(lift\_r)))$ $\qquad \Box(\neg(do(lift\_r) > \neg do(lift\_l)))$
Obs: $do(\epsilon) > (\neg up\_left \wedge \neg up\_right \wedge \neg spilled)$
$Frame_0 = \{(f,a) : a \in \Delta_0, f \in \mathcal{F}\}$.

As actions $lift\_l$ and $lift\_r$ are independent, the action $lift\_l$ has always the effect of lifting the left hand side of the bowl, also when it is executed in parallel to $lift\_r$. Hence, in a scenario in which both actions are executed in parallel we get:

$$(i) \;\; (do(lift\_l) \wedge do(lift\_r)) > (up\_left \wedge up\_right)$$

by applying the first two action laws and (CV) together with the third law saying that the two actions are independent. As spilled is not caused by action laws or causal laws, by assuming the persistency of $\neg spilled$ from the initial state[4] we get:

$$(ii) \;\; (do(lift\_l) \wedge do(lift\_r)) > (up\_left \wedge up\_right \wedge \neg spilled).$$

In a different scenario, when the action $do(lift\_l)$ is executed alone, its execution causes up_left by the first causal law. We can then assume the persistency of $\neg up\_right$ from the initial state, and apply the third (causal) law to get:

$$(iii) \;\; do(lift\_l) > (up\_left \wedge \neg up\_right \wedge spilled)$$

The formulas (ii) and (iii) hold in two different extensions, which are relative to two different scenarios, corresponding to different courses of actions. While it is clear that all the monotonical consequences of a domain description hold in all extensions (like, for instance, formula (i)), this is not true for the persistency assumptions that, as we will see in the next sections, are relative to a given scenario. For instance, formula (iii) does not hold in the scenario where the actions $do(lift\_l)$ and $do(lift\_r)$ are executed concurrently, as in this scenario we cannot make the assumption that $\neg up\_right$ persists from the initial state after the execution of action $do(lift\_l)$ alone. In the next section we will define a notion of extension which is relative to a given action history.

[4]For a detailed description of persistency of frame fluents we refer to the next section, where we introduce the notion of extensions of domain descriptions.

3.2 Extensions for a domain description

In discussing the examples above we have often described fluents persisting from one state to the next one, after an action has been performed. In this section we will provide a non-monotonic construction for our causal action logic AC to deal with persistency of fluents. The non-monotonic solution to the frame problem we adopt here is similar to the one adopted in [12] and in [10], but, as a difference with the proposals above, here we define the notion of extension "relative to an action sequence", that is, relative to the history of actions which have been executed. As we will see, this provides an easier way for dealing with persistency in presence of concurrent actions with respect to the solution proposed in [9].

We deal with the frame problem by introducing a set of persistency laws, which can be assumed in each extension. Persistency laws are essentially frame axioms. They are used, in addition to the formulas in Π, to determine the next state when an action is performed. As a difference with the formulas in Π, persistency laws are *defeasible*. They are regarded as assumptions to be maximized. Changes in the world are minimized by maximizing these assumptions. Moreover, persistency laws have to be assumed if this does not lead to inconsistencies.

Let $ca_1, \ldots, ca_n$ be (possibly) concurrent actions of the form $do(a_1) \wedge \ldots \wedge do(a_m)$ (for $m = 1$ we have an atomic action). We introduce a set of *persistency laws* of the form $ca_1 > \ldots > ca_{n-1} > (l \rightarrow (ca_n > l))$ for every sequence of (concurrent) actions $ca_1, \ldots, ca_{n-1}$ and for every fluent literal $l \in Lit$ which is a frame fluent with respect to the (concurrent) action ca_n (according to the definition of $Frame$ in the last subsection), that is, for every fluent literal l which is frame for *every elementary action* in ca_n. The persistency law says, that, "if l holds in the state obtained by executing the sequence of actions $ca_1, \ldots, ca_{n-1}$, then l persists after executing action ca_n in that state" [5].

Our notion of extension will require to introduce two different kinds of assumptions. The first kind of assumptions, as we have seen, are persistency assumptions. Given a set $Frame$ of frame fluents, the set of persistency assumptions $WP_{ca_1,\ldots,ca_n}$ is defined as follows:

$$WP_{ca_1,\ldots,ca_n} = \{ca_1 > \ldots > ca_{j-1} > (l \rightarrow (ca_j > l))) : \ (|l|, ca_j) \in Frame, \ 1 < j \leq n, \ l \in Lit, \ ca_1 = do(\epsilon)\}.$$

Note that the set of persistency assumptions has been defined relative to a sequence of (concurrent) actions, that is, a state.

In addition to persistency assumptions, in defining our extensions, we introduce another kind of assumptions, which are needed to deal with non frame fluents. If a fluent $f \in \mathcal{F}$ is not persistent with respect to a concurrent action ca then, in the state obtained after executing ca, the value of f is arbitrary, it may hold or not. Hence, we introduce assumptions which allow to assume, in any state, that f holds (or does not hold) for every non-frame fluent f, as well as assumptions for all fluents in the

[5] Notice that introducing persistency laws of the form $\Box(l \rightarrow (ca > l))$ wouldn't be enough to deal with the persistency of literals at each different state. In fact, l may persist when executing action ca in one state, while it may not persist when ca is executed in a different state, as the effects of action ca in the two states may be different

initial state. Given a set $Frame$ of frame fluents, we define the set of assumptions $Ass_{ca_1,\ldots,ca_n}$ relative to a sequence $ca_1, \ldots, ca_n$ of concurrent actions as follows:

$$Ass_{ca_1,\ldots,ca_n} = \{ca_1 > \ldots > ca_j > l : \; (|l|, ca_j) \notin Frame, 1 < j \leq n, \; l \in Lit, \; ca_1 = do(\epsilon)\} \cup \{do(\epsilon) > l : \; l \in Lit\}$$

We represent a generic assumption in this set by $ca_1 > \ldots > ca_j > l$, which includes assumptions on the initial state.

Observe that the fluents in $\mathcal{F}_{\Delta_0}$ are not subject to persistency nor they have to be assumed to be true or false in any state. In fact, assuming $do(a)$ true in a state forces the action a to be executed in that state and its effect to be caused. We do not want action execution to be non-deterministically forced or forced by persistency.

We can now define our notion of extension, for domain frames $(\Pi, Frame_0)$, and for domain descriptions $(\Pi, Frame_0, Obs)$. An extension E of a domain frame is defined *relative to a state*, which can be identified by the sequence of actions $ca_1, \ldots, ca_n$ leading to that state. It is obtained by augmenting Π by as many as possible persistency laws, while preserving the consistency of states.

Definition 3 An *extension* of a domain frame $D = (\Pi, Frame_0)$ relative to the action sequence $ca_1, \ldots, ca_n$ is a set $E = Th(\Pi \cup WP' \cup F)$, such that $WP' \subseteq WP_{ca_1,\ldots,ca_n}$, $F \subseteq Ass_{ca_1,\ldots,ca_n}$ and

a) for $1 < j \leq n$, if $ca_1 > \ldots > ca_{j-1} > (l \to (ca_j > l)) \in WP_{ca_1,\ldots,ca_n}$ then: $ca_1 > \ldots > ca_{j-1} > (l \to (ca_j > l)) \in WP' \iff ca_1 > \ldots > ca_j > \neg l \notin E$

b) for $1 \leq j \leq n$ if $ca_1 > \ldots > ca_j > l \in Ass_{ca_1,\ldots,ca_n}$ then $ca_1 > \ldots > ca_j > l \in F \iff ca_1 > \ldots > ca_j > \neg l \notin E$.

The $\Rightarrow$-part of condition a) is a consistency condition, which guarantees that a persistency axiom $ca_1 > \ldots > ca_{j-1} > (l \to (ca_j > l))$ cannot be assumed in WP' if $\neg l$ can be deduced as an immediate or indirect effect of the action ca_j. We say that the formula $ca_1 > \ldots > ca_j > \neg l$ blocks the persistency axiom. The $\Leftarrow$-part of condition a) is a maximality condition which forces the persistency axiom to be assumed in WP', if the formula $ca_1 > \ldots > ca_j > \neg l$ is not proved. Condition b) forces each state of an extension to be complete: for all finite sequences of actions $ca_1, \ldots, ca_j$ each non persistent fluent must be assumed to be true or false in the state obtained after executing them. In particular, since the sequence of actions may contain the empty action $do(\epsilon)$ alone (for $j = 1$), the initial state has to be complete in a given extension E. This is essential for dealing with domain descriptions in which the initial state is incompletely specified and with postdiction. The conditions above have a clear similarity with the applicability conditions for a default rule in an extension. We refer to [10] for a detailed description of the relationship between a similar notion of extension and default extensions.

Observe that our persistency law correspond to the positive and negative *frame axioms* in the situation calculus. The two frame axioms $F(x, s) \wedge \neg\gamma_F^+(x, a, s) \to F(x, do(a, s))$ and $\neg F(x, s) \wedge \neg\gamma_F^-(x, a, s) \to \neg F(x, do(a, s)))$ describe the persistency of fluent F from state s to the next state $do(a, s)$. As a difference, we do not have

$neg\gamma_F^+(x,a,s)$ (respectively, $\neg\gamma_F^-(x,a,s)$) in the antecedent of persistency laws, as we regard them as default rules and we adopt a default construction, rather than a completion based construction through the use of successor state axioms. Moreover, while in situation calculus situations are represented by terms $do(a_1, do(a_2, (\ldots, do(a_n)\ldots)))$, and replaced by variables in the frame axioms, here the state to which a persistency law applies is made explicit through the actions in the antecedent of the conditional implication. For this reason we cannot have a compact representation of persistency laws, while in the situation calculus a compact representation of frame axioms is given by replacing actions and states by variables. Observe, that given an extension E of a domain description relative to the action sequence $ca_1, \ldots, ca_n$, the number of laws in the sets $WP_{ca_1,\ldots,ca_n}$ and $Ass_{ca_1,\ldots,ca_n}$ is $n \times 2 \times$ *number of fluents*.

Definition 4 E is an *extension for a domain description* $(\Pi, Frame_0, Obs)$ *relative to the action sequence* $ca_1, \ldots, ca_n$ if it is an extension for the domain frame $(\Pi, Frame_0)$ relative to the action sequence $ca_1, \ldots, ca_n$ and $E \vdash Obs$.

Notice that first we have defined extensions of a domain frame $(\Pi, Frame_0)$; then we have used the observations in Obs to filter out those extensions which do not satisfy them.

As a difference with the notion of extension proposed in [9, 10], here an extension only describes a single course of actions, a *history*, and assumptions are localized to that sequence of actions. This allows us to deal with concurrent actions without introducing two different modalities for actions (called *open* and *closed* action modalities in [9]) in order to prevent the (AND) law $(do(a) > C) \rightarrow (do(a) \wedge do(b) > C)$ to be applied to non-monotonic consequences of actions, derived by means of persistency assumptions. This point will be explained in more detail below, when discussing example 5.

Let us consider again the example 1. Relative to the action sequence $do(\epsilon), do(tg_1), do(tg_2)$ we get one extension E containing the frame laws

(a) $do(\epsilon) > (\neg light \rightarrow (do(tg_1) > \neg light))$,
(b) $do(\epsilon) > (\neg sw_2 \rightarrow (do(tg_1) > \neg sw_2))$,
(c) $do(\epsilon) > (do(tg_1) > (sw_1 \rightarrow (do(tg_2) > sw_1)))$,

in which the following sentences hold:

(1) $do(\epsilon) > (do(tg_1) > \neg light)$,
(2) $do(\epsilon) > (do(tg_1) > (do(tg_2) > light))$,

E contains also

(3) $do(\epsilon) > (do(tg_1) \wedge do(tg_2)) > light$

as was shown in the last subsection. Moreover, from (1), we can derive

(4) $do(\epsilon) > (do(tg_1) \wedge do(tg_2)) > \neg light$

by (CV) and the independency of $do(tg_1)$ and $do(tg_2)$. Then, (3) and (4) together entail

$do(\epsilon) > (do(tg_1) \wedge do(tg_2)) > \bot$

which means that in this extension, which is relative to the action sequence $do(\epsilon)$, $do(tg_1), do(tg_2)$, the alternative sequence $do(\epsilon), do(tg_1) \wedge do(tg_2)$ is not possible: it leads to an "inconsistent" state, i.e. when executing $do(tg_1)$ alone, the execution of $do(tg_1) \wedge do(tg_2)$ is not possible.

An extension E relative to $ca_1, \ldots, ca_n$ determines an initial state and a transition function among the states obtained by executing actions $ca_1, \ldots, ca_n$. In particular, the *state* reachable through an action sequence $ca_1, \ldots, ca_j$ ($1 \leq j \leq n$) in E can be defined as :

$$S^E_{ca_1,\ldots,ca_j} = \{l \,:\, E \vdash ca_1 > \ldots > ca_j > l\},$$

where $S^E_{do(\epsilon)}$ represents the initial state. Due to condition (b) of definition 3, we can prove that each state $S^E_{ca_1;\ldots;ca_j}$ is *complete* with respect to the fluents in $\mathcal{F}$: for each fluent $f \in \mathcal{F}$, it contains either f or $\neg f$. Moreover, it can be shown that the state obtained after execution of the sequence of actions $ca_1, \ldots, ca_n$, is only determined by the assumptions made from the initial state up to that state.

Referring to example 1, the extension E above relative to the action sequence $do(\epsilon)$, $do(tg_1)$, $do(tg_2)$ determines the

states:

$$S^E_{do(\epsilon)} = \{\neg sw_1, \neg sw_2, \neg light\},$$
$$S^E_{do(\epsilon),do(tg_1)} = \{sw_1, \neg sw_2, \neg light\},$$
$$S^E_{do(\epsilon),do(tg_1),do(tg_2)} = \{sw_1, sw_2, light\}.$$

Observe that for the domain description in example 1 we do not obtain the unexpected extension where $do(\epsilon) > do(tg_1) > do(tg_2) > (\neg sw_1 \wedge sw_2 \wedge \neg light)$ holds: our theory prevents that toggling sw_2 in the state $\{sw_1, \neg sw_2, \neg light\}$ mysteriously changes the position of sw_1 and lets $\neg light$ persist. To avoid this extension it is essential that causal rules are directional (see [2, 29, 26, 42, 10]). Indeed, the causal rules in Π are different from the constraint $\Box(sw_1 \wedge sw_2 \rightarrow light)$ and, in particular, they do not entail the formula $sw_2 \wedge \neg light \rightarrow \neg sw_1$. As observed in [26] and [42], though this formula must be clearly true in any state, it should not be applied for making causal inferences. In our formalism, contraposition of causal implication is ruled out because the conditional $>$ does not satisfy (MP): from $\Box(\alpha > \beta)$ we can neither conclude $\alpha \rightarrow \beta$ nor $\neg\beta \rightarrow \neg\alpha$. On the other hand, it is easy to see that, using (CE), in any state of any extension, if $\alpha > \beta$ holds, and α holds, β also holds.

Our solution to the frame problem is an abductive solution and is different from the solution proposed for EPDL in [5]. There persistency laws of the form $l \rightarrow [a]l$ are added explicitly at every state. In EPDL, persistency laws are not global

to an extension but they have to be added state by state, according to which action is expected. In our theory, the frame problem is solved globally by minimizing changes modulo causation. As a further difference, in [5] unexpected solutions can be obtained by adding persistency laws as above to the domain description. As observed by Zhang and Foo (see [5], example 4.1) in the circuit example above the state $S_1 = \{sw_1, \neg sw_2, \neg light\}$ has two optional next states under action $toggle_2$, namely $S'_2 = \{sw_1, sw_2, light\}$ and $S''_2 = \{\neg sw_1, sw_2, \neg light\}$. The second one is unexpected.

This behavior is a side effect of (MP), which holds for EPDL and allows the material implication to be derived from the causal implication. To overcome this problem, Zhang and Foo propose an alternative approach to define the next-state function which makes use of a fixpoint property in the style of McCain and Turner's fixpoint property [29]. Their definition employs the causal operator for determining whether the indirect effects of the action are caused by its immediate effects together with the unchanged part of the state, according to the causal laws. It has to be observed, that this definition of the next state function does not require any integrated use of causal laws and action laws in the theory. In fact, "if the direct effects of an action have been given, $EPDL^-$ (that is, the logic obtained from EPDL when the set of action symbols is empty) is enough to determine how effects of actions are propagated by causal laws" [5]. On the contrary, our solution to the frame problem in the conditional logic CA relies on an integrated use of action laws and causal laws to derive conclusions about action effects.

A domain description may have no extensions. Consider the following example also mentioned by [29]:

Π: $\Box(do(a) > p)$ $\Box(q > \neg p)$
Obs: $do(\epsilon) > (q \wedge \neg p)$
$Frame_0 = \{(f, a) : a \in \Delta_0, f \in \mathcal{F}\}$.

If $q \wedge \neg p$ holds in the initial state, performing action a makes p true, but this cannot block the persistency of q since $\neg q$ cannot be derived from p since the causal rule is not contrapositive. However, assuming that q persists after the action a, leads to $do(a) > q$ and since $q > \neg p$, by (CE), we derive $do(a) > \neg p$ from which we get together with $do(a) > p$, $do(a) > \bot$. Hence, $do(a) > \neg q$ (as any formula can be derived from an inconsistency). Therefore, q cannot persist and the domain description above has no extension.

A domain description may have extensions containing inconsistent states, when $\bot \in S_{a1,\ldots,a_n}$. In fact, it may be that case that the set of laws and constraints in Π are themselves inconsistent or they (monotonically) derive the inconsistency after the execution of a sequence of actions. It may happen, for instance, that the concurrent execution of two actions declared as being independent may nevertheless produce an inconsistent state. Consider the following example:

Example 6 Consider a swinging door and two actions $push\_in$ and $push\_out$ the first one opening the door by pushing from out-side to open it and the second by pushing it in the opposite direction. We get the following formalization:

Π: $\Box(do(push\_in) > open\_in)$ $\Box(do(push\_out) > open\_out)$
$\Box(open\_in > \neg open\_out)$ $\Box(open\_out > \neg open\_in)$
$\Box(\neg(do(push\_in) > \neg do(push\_out)))$ $\Box(\neg(do(push\_out) > \neg do(push\_in)))$
$Frame_0 = \{(f, a) : a \in \Delta_0, f \in \mathcal{F}\}$.

We have assumed that the two actions are independent. But when trying to perform them concurrently, an inconsistent state is obtained, because there is a conflict between their effects. All the extensions of the theory contain the formulas:

(1) $do(push\_in) > open\_in$,
(2) $do(push\_out) > open\_out$,
(3) $do(push\_in) \wedge do(push\_out) > \bot$

where all the formulas above are derived monotonically for Π.

It should be pointed out that for the concurrent action step $do(push\_in) \wedge do(push\_out)$ leading to an inconsistent state no persistency law is applied: since everything is true in that state nothing has to be obtained by a persistency law.

The contradictory actions $push\_in$ and $push\_out$ are independent because they can be performed independently, even if their results are contradictory. Their concurrent execution produces an inconsistent state. Observe that this is a natural solution, as pushing the door in both directions blocks the door. Hence, the concurrent occurrence of both actions, $push\_in$ and $push\_out$ does not yield a resulting state.

We argue that, in some cases, the outcome of an inconsistent state (or the absence of the resulting state) may hint at some implicit qualification which are missing, or it may suggest that the two actions are actually dependent.

Let us now consider again example 5. We refer to this example to explain how our notion of extension deals with persistency in presence of concurrent actions. Let us first consider the extension E_1 relative to the action sequence $do(\epsilon), do(lift\_l)$ containing the persistency law:

$$do(\epsilon) > (\neg up\_right \rightarrow (do(lift\_l) > \neg up\_right)).$$

E_1 determines the states:

$$S^{E_1}_{do(\epsilon)} = \{\neg up\_left, \neg up\_right, \neg spilled\},$$
$$S^{E_1}_{do(\epsilon),do(lift\_l)} = \{up\_left, \neg up\_right, spilled\}.$$

In E_1 we have

(i) $do(\epsilon) > (do(lift\_l) > (up\_left \wedge \neg up\_right \wedge spilled))$

The extension E_2 relative to the action sequence $do(\epsilon), do(lift\_l) \wedge do(lift\_r)$ contains the persistency law:

$$do(\epsilon) > (\neg spilled \rightarrow (do(lift\_l) \wedge do(lift\_r) > \neg spilled))$$

E_2 determines the states:

$$S^{E_2}_{do(\epsilon)} = \{\neg up\_left, \neg up\_right, \neg spilled\},$$
$$S^{E_2}_{do(\epsilon),do(lift\_l) \wedge do(lift\_r)} = \{up\_left, up\_right, \neg spilled\},$$

In E_2 we have

$$do(\epsilon)) > (do(lift\_l) \wedge do(lift\_r)) > (up\_left \wedge up\_right \wedge \neg spilled).$$

Observe that the formula $do(\epsilon) > do(lift\_l) \wedge do(lift\_r) > up\_left \wedge up\_right$, which is monotonically derivable from the action laws using (CV), holds in both extensions. As the formula $do(\epsilon) > do(lift\_l) \wedge do(lift\_r) > \neg up\_right$ holds in E_1 (it can be inferred from (i) by (CV) and action independence), we can conclude that $do(\epsilon) > do(lift\_l) \wedge do(lift\_r) > \bot$ holds in E_1; that is, when we are reasoning about a course of actions in which $do(lift\_l)$ is executed alone (as in E_1, which is relative to the action sequence $do(\epsilon), do(lift\_l)$), the concurrent action $do(lift\_l) \wedge do(lift\_r)$

is not executable. In fact, action $do(lift\_r)$ causes up_right as its immediate effect, which is inconsistent with the fact that $\neg up\_right$ persists.

From this example it emerges that, when we want to evaluate a conditional formula $ca_1 > \ldots > ca_j > l$ in a domain description, in order to to check whether a fact l is caused by the sequence of actions $ca_1, \ldots, ca_j$, we should refer to those extensions of the domain description which are relative to that course of actions, that is, relative to a sequence of actions starting with the actions $ca_1, \ldots, ca_j$ in the antecedents of the conditional formula. Only such extensions are relevant to the course of actions described by the conditional.

It must be noted that, given a course of action $ca_1, \ldots, ca_n$ and an extension E relative to it, the values of fluents at the different states $S_{ca_1,\ldots,ca_j}$ (for $j \leq n$) are determined by evaluating the conditionals of the form $ca_1 > \ldots > ca_j > l$ in the extension. However, arbitrary (conditional) formula can of course be evaluated at a state in the extension, and this provides further information about the state: which laws hold at the state, which actions are (or are not) executable in the state, and so on.

The solution to the problem of separating different sequences of action occurrence proposed in [9] was to introduce two different modalities for distinguishing between the behavior of actions when they are executed in isolation ("closed action") and when they are executed in parallel with other actions ("open actions"). While persistency applies to closed actions, the (AND) rule applies to open actions. The behavior of the action theory emerges from the interplay between these two kinds of actions. As a difference, the solution adopted in AC refers explicitly to the current (linear) course of actions and defines the notion of extension relative to it. In this way, the persistency assumptions which can be taken in an extension are relative to the actions which have occurred.

4 Related Work

In this section we compare our approach to actions and causation with the solutions presented in the literature. Starting from the observation that causality cannot be represented by the classical implication, several different ways for representing actions and causality have been proposed in the literature. We can distinguish among the following approaches:

- Causality is formulated in the framework of a classical language, by introducing a special new non-logical predicate, as for instance the *Caused* predicate introduced by Lin [26, 27].

- Causality is considered as an inference relation on classical formulas, as for instance by McCain and Turner in [29].

- Causal relations are modelled by introducing a new causal operator in the language. McCain and Turner in [30, 44] introduce a new causal operator $\Rightarrow$. Thielscher [42] presents a STRIPS-like approach augmented by causal rules which are directional implications. The conditional approach we propose in this paper also falls into this class.

- Causality is defined through causal modalities: either a unary modal operator (for instance, Ⓒ, where Ⓒϕ means that ϕ is caused) or a set of modalities. The approach with a unary operator has been followed by Geffner [7], by Turner [44] and by Giordano, Martelli and Schwind [9, 10]. A multi-modal logic, actually an extension of dynamic logic, for causal reasoning has, instead, been proposed by Zhang and Foo [45]: formulas of the form $[\alpha]A$ (α causes A) allow both immediate and indirect effects of actions to be expressed. Also the work in [11, 12], which defines a theory for reasoning about actions in a linear time temporal logic, falls within this approach.

Most of the above mentioned proposals, though not all of them, develop a non-monotonic approach to the formalization of actions and causation. Hence, a further aspects on which they differ is the kind of non-monotonic formalism upon which they rely. In the following we will compare our approach to those mentioned above, by outlining the different properties of the causality relation.

The first systematic solution to the ramification problem has been proposed by Lifschitz in [25], by introducing a distinction between *frame* and *non-frame* fluents. However, as observed by Lin [26] in some cases (see, for instance, the suitcase example [26]) this distinction is not enough to prevent unwanted contrapositions of causal dependencies.

To overcome this problem, Lin in [26, 27] introduces a predicate *Caused(f,v,s)*, meaning that fluent f is caused to have the truth value v in the situation s. Lin's proposal is based on Situation Calculus [37]. The Situation Calculus is one of the most popular formalism for reasoning about actions and it has provided the very first account on action and causality. In its original formulation, it does not include the formalization of causal dependencies between fluents. In Lin's proposal, action laws are expressed by formulas of the form

$$Poss(a, s) \ \rightarrow (Holds(\phi, s) \rightarrow Caused(F, v, do(a, s))),$$

where F is a fluent name. The predicate $Caused$ is used for formulating causal relationships. For example, the first causal law in the circuit example 1 can be formalized as follows: $Holds(sw_1, s) \wedge Holds(sw_2, s) \rightarrow Caused(light, true, s)$. The *Caused* predicate is used in order to control the persistency of fluents: only fluents which are caused are allowed to change value (and for this reason the *Caused* predicate has to occur also in action laws) and only fluents which are *not* caused are allowed to persist. Circumscription is used to minimize the predicate Caused. In [38] Schwind has shown that the notion of causality in Lin's theory, differently from ours, satisfies the property of monotonicity (from "f CAUSES g" we can derive "f$\wedge$ h CAUSES g"), as well as transitivity ("f CAUSES g" $\wedge$ "g CAUSES h" $\rightarrow$ "f CAUSES h"). As discussed above we get weak monotonicity, from (CV), only in certain cases, when dealing with independent actions. Transitivity does not hold in our logic AC: the causal implication $>$ is not transitive, ($A > B \wedge B > C \rightarrow A > C$ is not derivable). Nevertheless, we have a weaker form of transitivity: from $\vdash ca > B$ and $\vdash \Box(B > C)$ we can derive $\vdash ca > C$ using (NEC), (MOD) and (CE). As our action logic, Lin's approach does not satisfy contraposition neither reflexivity (ID), but it takes into account reasoning by cases (CA). This is due to the fact that a causal rule "A CAUSES B" is represented in Lin's system by classical implication together with a unary causality predicate, and classical implication allows for reasoning by cases, but also for monotonicity, while Lin's predicate $Caused(F, v, do(a, s))$ prevents contraposition: $Caused(F, true, do(a, s))$ is not equivalent to $\neg Caused(F, false, do(a, s))$.

Javier Pinto has also shown how to treat causality and concurrency in the framework of situation calculus [34, 36, 35]. He does not modify the language of situation calculus in order to treat causation, but he includes it rather as an abstract notion. For concurrency, he proposes to represent concurrent actions as sets of atomic actions and to modify the successor state axioms consequently. In [35], he treats several types of concurrent actions. *Precondition interaction* corresponds to what we treated in example 4, *effect cancellation* concerns the bowl of soup example 5.

In [29] McCain and Turner define a causal theory in which causal rules are represented by inference rules. Given a state S and an action a, the next state function $Res^4_C(\mathit{Eff}, S)$, which provides the set of states which can be obtained by executing a in S, is defined through a fixpoint construction. McCain and Turner's causal theory does not satisfy *reasoning by case* (see [10]) while it satisfies another property, which was called *cumulative transitivity* (if "f CAUSES g" and "$f \wedge g$ CAUSES h" then "f CAUSES h") (see [45]). On the other hand, the logic AC does not satisfy cumulative transitivity. In [10] it was shown that, given a consistent state S, all the states computed by McCain and Turner's next state function could be obtained by the action theory in [10]. This is not true of AC, because we do not have *cumulative transitivity*. The vice-versa does not hold neither, because McCain and Turner's theory does not satisfy the (OR) property (axiom (CA) of our logic).

In [30] the same authors present a slightly different formalism for causality where causal laws are expressions of the form $\phi \Rightarrow \psi$ allowing for a formalization of both action immediate and indirect effects. For instance, the first action law and the first constraint of the circuit example above could be expressed in this formalism by the laws: $tg(i)_t \wedge \neg sw(i)_t \Rightarrow sw(i)_{t+1}$ and $sw(1)_t \wedge sw(2)_t \Rightarrow light_t$. Given a causal theory D, an interpretation I is *causally explained* according to D if I is the unique model of D^I, where D^I is the set of all heads of all laws in D whose bodies are satisfied by I. It can be shown that this notion of causality is transitive and takes into account reasoning by cases. Moreover, if a causal theory D contains a causal law $\phi \Rightarrow \psi$, then the material implication $\phi \rightarrow \psi$ holds in all the causally explained models of D (see [38]). It must be noticed, that though (MP) holds in this theory, it does not produce unwanted solutions. This is due to the fact that, in essence, causal laws are interpreted as default rules (see [43]), and therefore a contrapositive use of causal laws is not possible. Similar considerations can be made for the logic of universal causation introduced in [44], which extends McCain and Turner' causal theory by introducing in the language a modal operator C ("caused") in order to make a distinction between propositions that are caused and propositions that obtain, and for the nonmonotonic causal theories proposed by Giunchiglia et al. in [15] which deal with non-deterministic and concurrent actions. This causal theory of actions, which has been applied to several challenge problems in the theory of commonsense knowledge, also takes its origin in the non-monotonic causal logic introduced in [30].

The causal action theory presented by Thielscher in [42] is based on a STRIPS-like approach augmented by causal rules which are directional implications. The causal rules have the form l **causes** l' if $\bigwedge_i l_i$, where l, l' and l_i are literals, and they are automatically generated from a set of classical domain constraints, given an influence relation between fluents. The action theory in [42] does not allow (as ours) contraposition, reflexivity nor transitivity (see [38]). Some of the properties discussed above like reasoning by cases (OR) or monotonicity cannot be expressed, as conjunctions and disjunction are not admitted as preconditions (l above) of causal rules.

It has been established in [42] that all the resulting states computed in the fixpoint characterization in [29] can be obtained through Thielscher's notion of *successor* state, while the converse does not hold: there are successor states which do not correspond to any fixpoint. In Thielscher's approach a stable state can be reached through a sequence of unstable states, and the value taken by fluents in these unstable states (though different from their final value) may affect the value of some other fluents. This is not possible in our approach which is not able to reason about the unstable states. Hence, there are successor states in Thielscher's approach which do not correspond to any extension of our action theory.

The causal action logic introduced by Giordano, Martelli and Schwind in [10] is based on a modal language in which, modalities $[a]$ represent actions and the modality $\copyright$ is defined to represent causal dependency between fluents. $\copyright$ is a unary modality and "ϕ causes ψ" is expressed by the classical implication $\phi \rightarrow \copyright\psi$ (if ϕ holds then ψ is caused), whereas "action a causes ϕ under precondition ψ" is expressed by $\psi \rightarrow [a]\phi$. Interaction axioms, which rule the interactions among the different modalities, allow to infer $[a]\psi$ from $[a]\phi$ provided that action a causes ϕ and ϕ causes ψ. In [9] the language in [10] is extended to deal with concurrent actions.

Concerning the properties of causality, monotonicity holds for this notion of causality, as from $\phi \rightarrow \copyright\psi$ we can derive $\phi \wedge \pi \rightarrow \copyright\psi$ by propositional reasoning. Also, for the same reason, the property of disjunctive antecedents holds (if $\phi \vee \pi \rightarrow \copyright\psi$ then $\phi \rightarrow \copyright\psi$ and $\pi \rightarrow \copyright\psi$). Monotonicity does not hold in AC. Reasoning by case (disjunctive antecedents) is possible in AC due to the axiom CA. Cumulative transitivity holds for the causal operator in [9, 11], while it does not hold for AC.

As we have observed in section 3.2, a further difference with the proposals in [9, 10] is in the definition of extensions. In the logic AC, we have defined the notion of extension "relative to an action sequence". This allows a simplification in the treatment of concurrent actions. In [9], we needed two kinds of modalities, for *open* and *closed* actions in order to distinguish between concurrent and non-concurrent occurrence of an action.

In [12] a theory for reasoning about actions has been presented based on a linear time temporal logic, DLTL, in which regular programs of propositional dynamic logic can be used for indexing temporal modalities. In this theory a causal dependency among fluents "ϕ causes ψ" can be represented by the formula $\Box(\bigwedge_{a\in\Sigma}([a]\phi \rightarrow [a]\psi))$, meaning that for all actions a, if ϕ holds after the execution of a, then ψ also holds after its execution. This representation makes the causal laws directional (contraposition of causal laws is not allowed). Moreover, (MP) does not hold in this theory. As a difference with the logic AC, reasoning about cases (CA) is not allowed and causal implication is monotonic: from "ϕ causes ψ" it follows that "$\phi \wedge \alpha$ causes ψ".

Zhang and Foo in [45] present an extended propositional dynamic logic (EPDL) for causal reasoning. In EPDL the causal dependencies between actions and their effects are expressed through formulas of the form $[\alpha]A$, where α is a primitive or compound action and A is a property. Indirect effects of actions are expressed by allowing propositions as modalities. For instance, the first causal law in the circuit example 1 can be expressed by the formula $[sw_1 \wedge sw_2]light$. We have already outlined in the previous chapters the main differences of this approach with respect to our proposal from the point of view of the properties of causality and of the non-monotonic construction. Concerning the properties of causality, a major difference with EPDL is that the axiom (MP), which cannot be derived in our action logic AC, is derivable in EPDL. Though in EPDL contraposition of causal laws is not allowed (so that $[\neg C]\neg B$ cannot be derived from $[B]C$), as a consequence of (MP), it holds that: if $[A]\neg C$ and $[B]C$ are causal laws in Σ then $\vdash^{\Sigma} [A]\neg B$ (the derivation makes use of the inference rules (EN),(CW), (EK) and the axiom for the test action). The above inference makes some contrapositive use of the causal law $[B]C$.

Zhang and Foo also introduce a hierarchy of causal logics, to meet different requirements of causal reasoning. More precisely, they consider four logics EPDL$_1, ..,$ EPDL$_4$, obtained from EPDL through the incremental addition (to the axioms of EPDL) of the four axioms: (AND) $[\phi]A \rightarrow [\phi \wedge \psi]A$; (OR) $[\phi]A \wedge [\psi]A \rightarrow [\phi \vee \psi]A$; (Chn)$[\phi \wedge \psi]A \rightarrow [\phi][\psi]A$; and (PsC) $A \rightarrow [\phi]A$. They study the properties of these different logics. In AC only (OR) holds, which is the property of reasoning by cases, in our logic called (CA), which is the standard name of this axiom in conditional logics. The other axioms are not derivable in AC. In particular, AC does not contain (AND), but contains instead (CV), which is weaker than (AND): (CV) is a logical consequence of (AND). (CV) allows to apply (AND) only for independent actions or preconditions. Moreover, AC contains, as an inference rule, (RCK) which is implied by (AND) in EPDL and is one of the reasons mentioned by Zhang and Foo for accepting the (AND) property (reasons for refuting it are mentioned by the same authors and have also been mentioned in the introduction).

Another difference with our present approach is due to the different modelisation of actions. In EPDL, syntactically, actions are not formulas. This makes it impossible to combine assertions about actions with assertions about causality: it is not possible to express for example that action a and fact B cause the effect C and concurrent actions cannot be defined.

Besides the above proposals for integrating actions and causation, the concept of causality has been deeply explored by other researchers in the artificial intelligence community. In particular, other important approaches to reasoning about causality evolved from the area of Bayesian networks. Judea Pearl has defined a theory of causal reasoning based on the language of structural equations [32, 33]. According to Pearl, a causal model is given by two sets of variables U and V (exogenous and endogenous variables) and a set of functions, one for every endogenous variable X, associating to every vector of all the other variables in U and $V \setminus \{X\}$ a value in the set of possible values of X. These functions define structural equations relating the values of the variables of the system. Given a causal model, one can define a sub-model according to a vector X of endogenous variables and a vector $\overrightarrow{x}$ of values for these variables and a vector $\overrightarrow{u}$ of values of the exogenous variables by instantiating X to $\overrightarrow{x}$ and all exogenous variables to $\overrightarrow{u}$ inside all the structural equations defined by the functions. A sub-model describes a possible counterfactual situation. This submodel describes what would happen if the variables X are set to $\overrightarrow{x}$.

A *basic causal formula*, as defined in [17], has the form $[X_1 \leftarrow x_1, \ldots, X_1 \leftarrow x_1]\phi$ (where ϕ is a boolean combination of primitive events, i.e. formulas of the form $Y = y$). It says that ϕ holds in the counterfactual world that would arise if X_i is set to x_i, for all $i = 1, \ldots, n$.

Although in Pearl's theory causal implication is not represented as a logical connective, a basic causal formula is very close to a counterfactual and its evaluation in a causal model M (in the context $\overrightarrow{u}$) is obtained by "minimally changing" the model M by setting all variables X_i to x_i (and U_j to u_j) and then verifying if ϕ holds, as the result of the simultaneous instantiation of equations in M.

In [6], the authors study the causal interpretation of counterfactual sentences using the structural equation model for causality. They compare causal models to Lewis's logic for counterfactual sentences [24]. In the causal model the meaning of a Lewis's statement $A > B$ is "If we force a set of variables to have the values A, a second set of variables will have the values B", where A stands for the set of variables $\overrightarrow{x}$ and B for the set of variables $\overrightarrow{y}$.

Galles and Pearl show that causal implication as defined by causal models satisfies all the axioms and rules of Lewis'conditional logic. In particular, their system includes axioms (ID), (MP) and (CS) which, as we have already discussed in section 1,do not hold in our logic AC.

Galles and Pearl observe that, when restricting to recursive models, the causal model framework does not require stronger axioms for counterfactuals than those present in Lewis's logic, while in the non recursive case, further axioms would be needed to account for the property of "reversibility". In our conditional logic, on the other hand, we do not even introduce axioms (ID), (MP) and (CS), which hold for recursive models.

Starting from the notion of causality based on the language of structural equations Halpern and Pearl define a different notion of causality, which they call *actual causality* as well as a notion of causal explanation [17]. As the authors observe, for this notion of actual causality, which is defined to be reflexive, one might want to avoid reflexivity (to avoid that $X = x$ is a cause for itself). As they suggest, reflexivity can be avoided by requiring that $X = x \wedge \neg\phi$ be consistent for $X = x$ to be a cause of ϕ.

5 Conclusion

We have presented a new conditional logical approach to reason about actions and causality which uses a single implication $>$ for causal consequence. This new approach is based on conditional logic and includes standard axioms and inference rules of Lewis' conditional logic VCU. Action execution and causal implication are represented uniformly. This makes it possible to integrate reasoning about mutual action dependence or independence into the language of the logic itself. This possibility distinguishes our approach from many other approaches, for example that in [18] or [42] who formulate dependencies or influences outside the logic. Our action language can handle (co-operating, independent and conflicting) concurrent actions in a natural way without adding extra formal devices, and we believe that the language can be naturally extended to handle other boolean expressions concerning action performance. This paper extends the work presented in [13], with respect to which we have slightly changed the axiom system, to make it as essential as possible in modelling causality. Moreover, we have restricted the language to conditional formulas with propositional antecedents. As we have explained, conditional logic can be characterized semantically by several types of semantics (ordered models, spheres or selection functions). We choose selection function semantics because this is the most general semantic system for conditional logics. We think however that correspondence with the other semantic systems can be shown along the same lines as by Grahne [16].

There are several issues which have still to be addressed. First of all, we need to develop a proof theory for this conditional logic, to make the approach usable in practice. For that we want to use analytic tableaux. In [8] a tableaux system was developed for a series of conditional logics built from a system called (CE) which includes the minimal conditional logic , (ID) and (CA) together with more specific axioms. We think that it is possible to formulate tableaux rules in the spirit of [8] for our logic AC. Moreover, we think that tableaux are especially adapted for handling abductive inference as used in this present paper. Extensions as we defined them here can be obtained as tableaux branches.

Another important issue is to determine whether this logic can be made tractable under suitable restrictions of the language. On the one hand, we want to consider syntactic restrictions as, for instance, a restriction to clausal formulas. On the other hand, we can think of putting some restriction to rule the interplay between the "do(a)"-literals and the other literals in conditional formulas.

Another interesting issue to be tackled is that of exploring a spectrum of different notions of causality, which can be obtained by changing some of the postulates or by the addition of new postulates, as it has been done by Zhang and Foo [45] for EPDL. Different notions of causality have also been studied by other authors in philosophy [1] and in artificial intelligence [33, 17].

A most important issue concerns the integration of time (or of "real time") into our logic. In our present proposal, we have only a notion of state. We think that for representation issues states are a very important notion of qualitative representation of actions but it could be very useful to have additionally a notion of time without abandoning states. This would allow to include duration of actions, beginning and end of actions, delayed effects and other concepts related especially to time. Including temporal reasoning would probably necessitate to work with first-order conditional logics and to represent time point and intervals by natural or real numbers. Most work on conditional logic today is on propositional conditional logic, only few researchers have worked yet on first-order conditional logic [3]. We think that this extension of our logic will be not trivial, but is completely necessary for many real applications.

We would like to thank the anonymous referees for their comments which helped to improve this paper.

References

[1] Jonathan Bennet. *Event Causation: The Counterfactual Analysis*, pages 217 – 233. In Sosa and Tooley [41], 1993.

[2] G. Brewka and J. Hertzberg. How to do things with worlds: on formalizing actions and plans. *Journal of Logic and Computation*, 3(5), 1993.

[3] James P. Delgrande. On first-order conditional logics. *Artificial Intelligence*, 105(1-2), 1998.

[4] Nonman Foo, Dongmo Zhang, Yan Zhang, Samir Chopra, and Bao Vo. Encoding solutions of the frame problem in dynamic logic. In T. Eiter, W. Faber, and M. Truszczynski, editors, *Logic Programming and Nonmonotonic Reasoning*, volume 2173 of *LNAI*, pages 240–253. Springer, 2001.

[5] Norman Foo and Dongmo Zhang. Dealing with the ramification Problem in extended Propositional Dynamic logic. In F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyaschev editors, *Advances in Modal Logic: Volume 3*, World Scientific, pp. 173-191, 2002.

[6] David Galles and Judea Pearl. An axiomatic characterization of causal counterfactuals. *Foundation of Science*, 3(1):151–182, 1998.

[7] Hector Geffner. *Default Reasoning: Causal and Conditional Theories*. MIT Press, 1992.

[8] Laura Giordano, Valentina Gliozzi, Nicola Olivetti, and Camilla Schwind. Tableaux calculi for preference-based conditional logics. In Marta Cialdea Mayer and Fiora Pirri, editors, *Theorem Proving with Analytic Tableaux and Related Methods*, pages 81–101, Rome, Italy, 2003. Springer Verlag, LNAI 2796.

[9] Laura Giordano, Alberto Martelli, and Camilla Schwind. Dealing with concurrent actions in modal action logic. In H. Prade, editor, *Proceedings of the 13th European Conference on Artificial Intelligence*, pages 537–541. John Wiley & Sons, Ltd, 1998.

[10] L. Giordano, A. Martelli and C. Schwind. Ramification and Causality in a Modal Action Logic. *Journal of Logic and Computation*, vol. 10 No. 5, pp. 625-662, 2000.

[11] Laura Giordano, Alberto Martelli, and Camilla Schwind. Reasoning about actions in dynamic linear time temporal logic. In *Proceedings of the 3rd International Conference on Formal and Applied Practical Reasoning*, London, United Kingdom, 2000. Imperial College.

[12] Laura Giordano, Alberto Martelli, and Camilla Schwind. Reasoning about actions in dynamic linear time temporal logic. *Logic Journal of the Interest Group in Pure and Applied Logics IGPL*, 9(2):289–303, 2001.

[13] Laura Giordano and Camilla Schwind. Towards a Towards a Conditional Logic of Action and Causation. In *Logics in Artificial Intelligence, 8th European Conference JELIA'2002*, Lamezia, Italy, Springer LNAI 2424, pages 382-394, 2002.

[14] E. Giunchiglia and V. Lifschitz. Dependent fluents. In IJCAI'95 [20], pages 1964–1969.

[15] Enrico Giunchiglia, Joohyung Lee, Vladimir Lifschitz, Norman McCain, and Hudson Turner. Nonmonotonic causal theories. *International Journal of Intelligent Systems*, 2002. http://www.cs.utexas.edu/users/tag/cc/.

[16] Gösta Grahne. Updates and counterfactuals. *Journal of Logic and Computation*, 8(1):87–117, 1998.

[17] Joseph Halpern and Judea Pearl. Causes and explanations: A structural model approach, part i. In IJCAI'01 [19], pages 27–34.

[18] Andreas Herzig Marcos A. Castilho, Olivier Gasquet. Formalizing action and change in modal logic i: the frame problem. *Journal of Logic and Computation*, 9(5):701–735, 1999.

[19] *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, Seattle, USA, 2001. Morgan Kaufmann.

[20] *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1995. Morgan Kaufmann.

[21] Frank Jackson. A causal theory of counterfactuals. *Australian Journal of Philosophy*, 55(1):3–21, 1977.

[22] David Lewis. Causation. *Journal of Philosophy*, 70:556–567, 1973. Reprinted in [41].

[23] David Lewis. *Counterfactuals*. Basil Blackwell Ltd, 1973.

[24] D. Lewis, Counterfactuals and Comparative Possibility, in W.L. Harper, R. Stalnaker, and G. Pearce *Ifs*, Reidel, Dordrecht, Holland, (1981).

[25] V. Lifschitz. Actions in the frame of situations. *Artificial Intelligence*, 8:183–212, 1990.

[26] Fangzhen Lin. Embracing causality in specifying the indirect effects of actions. In IJCAI'95 [20], pages 1985–1991.

[27] Fangzhen Lin. Embracing causality in specifying the indeterminate effects of actions. In *Proceedings of the 13th National Conference on Artificial Intelligence*, pages 1985–1991. AAAI Press / The MIT Press, August 1996.

[28] David Makinson. Five faces of minimality. Studia Logica 52, pages 339-379, 1993.

[29] N. McCain and H. Turner. A causal theory of ramifications and qualifications. In IJCAI'95 [20], pages 1978–1984.

[30] Norman McCain and Hudson Turner. Causal theories of action and change. In *Proceedings of the 14th National Conference on Artificial Intelligence*, pages 460–465. AAAI Press / The MIT Press, July 1997.

[31] Donald Nute. *Topics in Conditional Logic*. Reidel Publishing Company, Dordrecht, 1980.

[32] Judea Pearl. Reasoning with cause and effect. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, Stockholm, Sweden, 1999. Morgan Kaufmann.

[33] Judea Pearl. *Causality: Model, Reasoning and Inference*. Cambridge University Press, 2000.

[34] Javier Pinto. Causality in theories of action. In R. Miller and M. Shanahan, editors, *Fourth Symposium on Logical Formalizations of Commonsense Reasoning*, January 1998.

[35] Javier Pinto. Concurrency and action interaction. Technical report, 2000. submitted.

[36] Javier A. Pinto. Concurrent actions and interacting effects. In Anthony G. Cohn, Lenhart Schubert, and Stuart C. Shapiro, editors, *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning (KR-98)*, pages 292–303, San Francisco, June 2–5 1998. Morgan Kaufmann Publishers.

[37] Raymond Reiter. *Knowledge in Action*. The MIT Press, 2001.

[38] C. Schwind. Causality in action theories. *Electronic Articles in Computer and Information Science*, 3(A):27–50, 1999.

[39] Kurt Segerberg. Notes on conditional logic. *Studia Logica*, 48:157–168, 1989.

[40] J. R. Shoenfield. *Mathematical Logic*. Addison - Wesley Publishing Company, Mass., 1967.

[41] E. Sosa and M. Tooley, editors. *Causation*. Oxford Readings in Philosophy. Oxford University Press, 1993.

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

[43] H. Turner. Representing Actions in Logic Programs and default theories: a situation calculus approach, *J. Logic Programming*, 31, pages 245–298, 1997.

[44] Hudson Turner. A logic of universal causation. *Artificial Intelligence*, 113:87–123, 1999.

[45] Dongmo Zhang and Norman Foo. EPDL: A logic for causal reasoning. In IJCAI'01 [19], pages 131–136.

A APPENDIX-Completeness

The completeness is shown by the construction of a canonical model. We construct a model such that for any consistent formula A (i.e. any formula A, such that $\not\vdash \neg A$), there is a world in this model satisfying A. Moreover, we show that the semantic properties of AC (S-CV, S-OR, etc.) hold in the canonical model. We use induction over the height of formulas which is the number of connectives (classical, modal and conditional) occurring in a formula.

Definition 5

1. *A set of formulas* Γ *is called* inconsistent *iff there is a finite subset of* Γ, $\{F_1, \ldots F_n\}$ *such that* $\vdash \neg F_1 \vee \neg F_2 \vee \ldots \neg F_n$. Γ *is called* consistent *if* Γ *is not inconsistent. If an (in)consistent* Γ *contains only one formula* F*, we say that* F *is (in)consistent.*

2. *A set of formulas* Γ *is called* maximal consistent *iff it is consistent and if for any formula* F *not in* Γ, $\Gamma \cup \{F\}$ *is inconsistent.*

We will use properties of maximal consistent formula sets the proof of which can be found in most text books of formal logic (see e.g. [40].

Lemma 1 *Let* w *be a maximal consistent set of formulas and* A *and* B *formulas of AC. Then* w *has the following properties:*

1. If $\vdash A \rightarrow B$ and $A \in w$, then $B \in w$

2. If from $A \in w$ we infer $B \in w$, then $A \rightarrow B \in w$.

3. $A \wedge B \in w$ iff $A \in w$ and $B \in w$

4. $A \notin w$ iff $\neg A \in w$

Given a maximal consistent formula set w, we set

$$w^A = \{B : A > B \in w\}$$
$$w^\Box = \{A : \Box A \in w\}$$

$w^\Box$ is consistent since by axiom (T) and Lemma 1, $w^\Box \subseteq w$, and w is consistent. w^A may be inconsistent.
The canonical model CM is defined by $CM =< W, f, R, [[\ ]] >$ where

1. W is the set of all maximal consistent formula sets of AC.
   We set $\|A\| = \{w : w \in W$ and $A \in w\}$ for any formula A.

2. For formula A and $w \in W$, $f(A, w) = \{w' : w' \in W$ and $w^A \subseteq w'\}$. Note that $f(A, w) = \emptyset$ whenever w^A is inconsistent.

3. For $w, w' \in W$, $R(w, w')$ iff $w^\Box \subseteq w'\}$.

4. for any atom $p \in ATM$, $[[p]] = \{w : p \in w\}$ and for formulas A and B, we have

$$[[A \wedge B]] = [[A]] \cap [[B]]$$
$$[[\neg A]] = W - [[A]]$$
$$[[A > B]] = \{w : \ f(A, w) \subseteq [[B]]\}$$
$$[[\Box A]] = \{w : \ R(w) \subseteq [[A]]\}$$

If w^A is inconsistent for some formula A, $f(A, w) = \emptyset$. $w^\Box$ is always consistent, because T holds for $\Box$.

We first will show that for any formula A, $\|A\| = [[A]]$. This is proven by induction on the height of formulas. First we need the following two lemmas concerning conditionals and modal formulas.

Lemma 2 *For any conditional formula $A > B \in w$ we have $A > B \in w$ iff for all $w' \in f(A, w)$, $B \in w'$.*

Proof : If w^A is inconsistent, $f(A, w)$ is empty and the lemma trivially holds. If not, for the "$\Rightarrow$" direction, let $A > B \in w$ and $w' \in f(A, w)$. Then $B \in w^A$ and $w^A \subseteq w'$, from which follows that $B \in w'$.
"$\Leftarrow$": We first observe that $w^A \cup \{\neg B\}$ is an inconsistent formula set. Suppose for the contrary, that $w^A \cup \{\neg B\}$ is consistent. Then it is included in a maximal consistent formula set $w' \in W$, $w^A \cup \{\neg B\} \subseteq w'$. But then $w^A \subseteq w'$, which means that $w' \in f(A, w)$. From this follows by our precondition that $B \in w'$. This is a contradiction to $\neg B \in w'$, since w' is consistent. Since $w^A \cup \{\neg B\}$ is inconsistent, there are formulas $\{F_1, \ldots F_n\} \subseteq w^A$ such that $\vdash \neg F_1 \vee \neg F_2 \vee \ldots \vee B$. By the rules of propositional calculus and rule RCK, we conclude $\vdash (A > F_1) \wedge (A > F_2) \wedge \ldots (A > F_n) \rightarrow (A > B)$. Since $F_i \in w^A$, $A > F_i \in w$ for $1 \leq i \leq n$, hence $A > B \in w$ since w is maximal consistent (by lemma 1). Q.E.D.

Lemma 3 *For any modal formula $\Box A$, we have $\Box A \in w$ iff for all $w' \in R(w)$, $A \in w'$.*

Proof : The first half follows immediately from the definition of the state transition relation R: if $\Box A \in w$ then $A \in w^\Box$. Let be w' with $R(w, w')$. Then $w^\Box \subseteq w'$, hence $A \in w'$. For the second half, suppose that $A \in w'$ for all w' such that $R(w, w')$. Then, we first observe that $w^\Box \cup \{\neg A\}$ is an inconsistent formula set. Suppose for the contrary, that $w^\Box \cup \{\neg A\}$ is consistent. Then it is included in a maximal consistent formula set $w'' \in W$, i.e. $w^\Box \cup \{\neg A\} \subseteq w''$. But then $w^\Box \subseteq w''$, which means that $R(w, w'')$. From this follows by our precondition that $A \in w''$. This is a contradiction to $\neg A \in w''$, since w'' is consistent. Since $w^\Box \cup \{\neg A\}$ is inconsistent (and $w^\Box$ is consistent, there are formulas $\{F_1, \ldots F_n\} \subseteq w^\Box$ such that $\vdash \neg F_1 \vee \neg F_2 \vee \ldots \vee A$. By the rules of propositional calculus, necessitation (NEC) and (K), we conclude $\vdash \Box F_1 \wedge \Box F_2 \wedge \ldots \Box F_n \rightarrow \Box A$. But $\Box F_i \in w$ for $1 \leq i \leq n$, hence $\Box A \in w$ by lemma 1.1, since w is maximal consistent. Q.E.D.

Lemma 4 *$\|F\| = [[F]]$ for arbitrary formula F.*

Proof : This is shown by induction on the height n of F.

- *For $n = 0$, F is a propositional variable (including an action predicate $do(a)$), then the lemma follows immediately from the definition of the model.*

- *Let $F = \neg A$, then $w \in \|\neg A\|$ iff $\neg A \in w$ which is equivalent to $A \notin w$ by lemma 1, 4; which is equivalent to $w \notin \|A\|$, hence by induction hypothesis, $w \notin [[A]]$ which means equivalently that $w \in [[\neg A]]$ by the definition of the valuation function in the model.*

- *Let $F = A \wedge B$, then $w \in \|A \wedge B\|$ iff $A \wedge B \in w$. By lemma 1, 3., this is equivalent to $A \in w$ and $B \in w$ or equivalently $w \in \|A\|$ and $w \in \|B\|$. By induction hypothesis $\|A\| = [[A]]$ and $\|B\| = [[B]]$, hence $w \in [[A]]$ and $w \in [[B]]$, i.e. $w \in [[A]] \cap [[B]] = [[A \wedge B]]$ by the definition of the valuation function $[[]]$ of the canonical model. Hence, we have that $\|F\| = [[F]]$.*

- *Let $F = \Box A$. By induction hypothesis, we have $\|A\| = [[A]]$ since A is a subformula of F. Let be $w \in \|\Box A\|$. By the definition of $\|\Box A\|$, this is equivalent to $\Box A \in w$. By lemma 3, this is the case iff for all $w' \in R(w)$, $A \in w'$. By the definition of $\|A\|$, this is equivalent to $\forall w' \in R(w)$, $w' \in \|A\|$. Since by induction hypothesis, $\|A\| = [[A]]$, this is equivalent to $\forall w' \in R(w)$, $w' \in [[A]]$. And this is the case iff $w \in [[\Box A]]$.*

- *Let $F = A > B$. By induction hypothesis $\|A\| = [[A]]$ and $\|B\| = [[B]]$, since A and B are subformulas of F. Let be $w \in \|A > B\|$. By the definition of $\|A > B\|$, this is equivalent to $A > B \in w$. By lemma 2, this is the case iff for all $w' \in f(A, w)$, $B \in w'$. By the definition of $\|B\|$, this is equivalent to $\forall w' \in f(A, w)$, $w' \in \|B\|$. By induction hypothesis, $\|B\| = [[B]]$ and we get $\forall w' \in f(A, w)$, $w' \in [[B]]$. And this is the case iff $f(A, w) \subseteq [[B]]$ which means that $w \in [[A > B]]$.*

Q.E.D.

It remains to show that the canonical model CM has the properties required by our logic AC, provided the corresponding axioms belong to the logic $(S{-}CV, S{-}CV, \ldots)$.

- S-RCEA if $[[A]] = [[B]]$ then $f(A, w) = f(B, w)$
  If $[[A]] = [[B]]$, then $w \in [[A]]$ iff $w \in [[B]]$. By lemma 1, 2, this implies $A \leftrightarrow B \in w$. By RCEA, it follows that $A > C \leftrightarrow B > C \in w$, from which we get $f(A, w) = f(B, w)$.

- (S-CV) if $f(A, w) \cap [[C]] \neq \emptyset$ then $f(A \wedge C, w) \subseteq f(A, w)$
  Let be $w' \in f(A \wedge C, w)$ iff $w^{A \wedge C} \subseteq w'$. By precondition, we have $f(A, w) \cap [[C]] \neq \emptyset$, i.e. $f(A, w) \notin (W \setminus [[C]])$, but $W \setminus [[C]] = [[\neg C]]$, from which follows that $\neg(A > \neg C) \in w$. This yields using axiom CV, $(A > B) \rightarrow (A \wedge C > B) \in w$. From this we conclude $\{B : A > B \in w\} \subseteq \{B : A \wedge C > B \in w\}$ which means that $w^A \subseteq w^{A \wedge C}$. Hence we get $w^A \subseteq w'$, i.e. $w' \in f(A, w)$.

- (S-CA) $f(A \vee B, w) \subseteq f(A, w) \cup f(B, w)$
  Suppose for the contrary that there is $w_1 \in W$ such that $w_1 \notin f(A, w)$ and $w_1 \notin f(B, w)$. Then there are formulas F_1 and F_2 such that $A > F_1 \in w$ and $F_1 \notin w_1$ and $B > F_2 \in w$ and $F_2 \notin w_1$ by the definition of the selection function of the canonical model. Since w_1 is maximal consistent, we have that $\neg F_1 \in w_1$ and

$\neg F_2 \in w_1$. By RCK and the maximality of w_1 , we get $A > F_1 \vee F_2 \in w_1$ and $B > F_1 \vee F_2 \in w_1$. By axiom CA this yields $A \vee B > F_1 \vee F_2 \in w_1$, from which follows that $F_1 \vee F_2 \in w^{A \vee B}$. Hence we cannot have $w^{A \vee B} \subseteq w_1$ because this would contradict $\neg F_1 \in w_1$ and $\neg F_2 \in w_1$ (maximality of w_1). Therefore $w_1 \notin f(A \vee B, w)$.

- (S-CE) if $f(ca, w) \subseteq [[B]]$ then

  $$ValProp(f(ca, w)) \subseteq ValProp(f(B, f(ca, w)))$$

  where $f(B, f(ca, w))$ represent the set of worlds $\{z \in f(B, x) \quad : \quad x \in f(ca, w)\}$.
  By the precondition, we have $ca > B \in w$; by axiom CE, we then get $(ca > (B > C)) \rightarrow (ca > C) \in w$. But this means that $\{F : B > F \in w^{ca}\} \subseteq \{F : ca > F \in w\}$, where $\{F : ca > F \in w\} = w^{ca}$. To prove our thesis we prove that all the propositional formulas which hold in all the worlds $w' \in f(B, f(ca, w))$ also hold in all the worlds $w'' \in f(ca, w)$. If $\alpha \in \mathcal{L}$ and, for all $w' \in f(B, f(ca, w))$, $\alpha \in w'$, then it must be that $\alpha \in \{F : B > F \in w^{ca}\} \cap \mathcal{L}$. Hence, by the inclusion above, $\alpha \in \{F : ca > F \in w\} \cap \mathcal{L}$. This means that, $\alpha \in \mathcal{L}$ and, for all $w' \in f(ca, w)$, $\alpha \in w'$, which proves our thesis.

- (S-MOD) $f(ca, w) \subseteq \{w' : R(w, w')\}$
  We first show that for any $a \in \Delta_0$, $w^\Box \subseteq w^{ca}$. Let be $A \in w^\Box$. Then $\Box A \in w$. By axiom MOD this gives $ca > A \in w$, since w is maximal consistent; and this is equivalent to $A \in w^{ca}$. Now, let be $w' \in f(ca, w)$. This is the case iff $w^{ca} \subseteq w'$ by the definition of the canonical model. Since $w^\Box \subseteq w^{ca}$, we get $w^\Box \subseteq w'$, which is equivalent to $R(w, w')$, by the definition of the model.

- (S-4) and (S-T) follow straightforwardly from the definitions of R by the canonical model using the S4 axioms and completeness properties of the model states.

Proof of the completeness theorem: Let A be a formula not derivable in AC. Then $\not\vdash A$, i.e. $\{\neg A\}$ is consistent. Then there is a maximal consistent set of formulas w such that $\neg A \in w$, i.e. $w \in \|\neg A\|$. By lemma 4, we get that $w \in [[\neg A]]$ which means that the canonical model CM satisfies $\neg A$, i.e. $CM, w \not\models A$. Q.E.D.

B APPENDIX-Decidability

We prove that:

Theorem 2 *The logic AC is decidable.*

We show that for any formula F, if there is a $AC-$model M and a world w_0 of M such that $(M, w_0) \models F$, then there is also a *finite* $AC-$model M^* and a world w'_0 in it such that $(M^*, w'_0) \models F$. This property is called the *finite model property*. This finite model property, together with the recursiveness of the axiomatization of the logic entails that the logic is *decidable*.

Indeed, for any formula F if it is a theorem of the logic, it will eventually be derived from the axioms and derivation rules. If it is *not* a theorem of the logic, by considering all the finite $AC-$models we shall eventually find a *finite* $AC-$model M^* that falsifies it.

Let F be a formula, M a AC-model and w_0 a world in M such that $M, w_0 \models F$. We show that we can build a finite $AC-$model M^* containing w_0 such that $M^*, w_0 \models F$.

Intuitively, the new finite model M^* is built from M, w_0 by considering only the portion of M that is relevant to determine the truth value of F in w_0. Furthermore, a sort of filtration is applied to the model so obtained: in M we define two worlds to be *equivalent* when they agree on the evaluation of all the subformulas of F. For each set of worlds selected by the selection function, we consider only one representing element for each equivalence class.

Let $nl(\alpha)$ be the maximum nesting level of the $\Box$ and the $>$ connectives in α. We assume that n is the maximum nesting level of the $\Box$ and the $>$ connectives in F ($nl(F) = n$). Let Var_F be the set of the propositional variables appearing in F, $\mathcal{L}_F$ the boolean closure of $Var_F \cup \{\top\}$ and $Subf_>(F)$ be the set of all subformulas of F. We can now define the closure $\mathcal{L}_F^>$ of a formula F, the set of formulas we will use to define the equivalence relation on worlds, by the following conditions:

- $Subf_>(F) \subseteq \mathcal{L}_F^>$;
- if $\Box\alpha \in \mathcal{L}_F^>$ then $\Box\Box\alpha \in \mathcal{L}_F^>$;
- if $A \in \mathcal{L}_F$ and $\alpha \in \mathcal{L}_F^>$ (with $nl(\alpha) = n$) then $A > \alpha \in \mathcal{L}_F^>$;
- if $\alpha \in \mathcal{L}_F$ and α is a non negated formula then $\neg\alpha \in \mathcal{L}_F^>$.

Nothing else is contained in $\mathcal{L}_F^>$.

$\mathcal{L}_F^>$ corresponds to what is usually called the *Fischer-Ladner closure* of F. However, for this conditional logic we need to define it for the different levels of nesting of the conditional and modal formulas up to n.

We define, for all $i = 0, \dots, n$,

$$\mathcal{L}_{F,i}^> = \{\alpha \in \mathcal{L}_F^> \ : \ nl(\alpha) \leq i\}.$$

For example, $\Box(\neg A > \neg\Box(C > D))$ (with $A, C, D \in \mathcal{L}_F$) is a formula with 4 levels of nesting, which belongs to $\mathcal{L}_{F,i}^>$, for $i \geq 4$. Moreover $\Box\Box(\neg A > \neg\Box(C > D))$

belongs $\mathcal{L}^{>}_{F,i}$, for $i \geq 5$, and $(\neg)B_1 > \ldots > (\neg)(B_6 > C)$ with $B_1, \ldots, B_6, C \in \mathcal{L}_F$ belongs to $\mathcal{L}^{>}_{F,i}$, for $i \geq 6$.

As a property of the sets $\mathcal{L}^{>}_{F,i}$, we have that: (1) $\mathcal{L}^{>}_{F,i} \subseteq \mathcal{L}^{>}_{F,i+1}$, for all $i = 1, \ldots, n$, and (2) $\mathcal{L}^{>}_{F,n}$ contains F and all its subformulas.

The set of formulas $\mathcal{L}^{>}_{F}$ is clearly finite. The size of $Subf_{>}(F)$ is linear in $|F|$. Observe that, given a number $O(|F|)$ of propositional variables, we can distinguish among $O(2^{|F|})$ different propositional evaluations, and hence, $O(2^{|F|})$ different propositional formulas. Hence, the size of $\mathcal{L}_{\mathcal{F}}$ is $O(2^{|F|})$. As the maximum number n of levels of nesting in any conditional formula in $\mathcal{L}^{>}_{F}$ is n, the size of $\mathcal{L}^{>}_{F}$ is exponential in $|F|$, namely $|\mathcal{L}^{>}_{F}| = O(2^{|F|})$.

We define an equivalence relation $\equiv_F$ over the set of worlds W by stipulating that two worlds are equivalent if they evaluate in the same way all the formulas in $\mathcal{L}^{>}_{F}$. Thus:

$$w \equiv_F w' \text{ if and only if for any formula } A \in \mathcal{L}^{>}_{F}, \quad w \models A \text{ if and only if } w' \models A.$$

The equivalence class of the world w in $W/\equiv_F$ will be denoted by $[w]$, and identified with a representative element in the class. For any set S of possible worlds in W, we denote by $S/\equiv_F$ the set of the equivalence classes of S according to $\equiv_F$.

We build the model $M^* = \langle W^*, f^*, [[\ ]]^* \rangle$ as follows:

The set of worlds W^* of M^* is the set of the equivalence classes:

$$W^* = \{[w] : w \in W\}$$

For properly defining the selection function f^*, we introduce finite sequence W_0, $W_1, \ldots, W_n$ of sets of worlds:

$W_0 = \{[w_0]\}$;
$W_1 = W_0 \cup \{[w'] : w' \in f(A, w_0),\ \text{for} A \in \mathcal{L}_F\} \cup \{[w'] : w_0 R w'\}$;
$\ldots$
$W_i = W_{i-1} \cup \{[w'] : w' \in f(A, w),\ \text{for} A \in \mathcal{L}_F,\ [w] \in W_{i-1}\} \cup \{[w'] :$
$wRw',\ [w] \in W_{i-1}\}$;
$\ldots$
$W_n = W_{n-1} \cup \{[w'] : w' \in f(A, w),\ \text{for} A \in \mathcal{L}_F,\ [w] \in W_{n-1}\} \cup \{[w'] :$
$wRw',\ [w] \in W_{n-1}\}$

W_n contains the equivalence classes $[w]$ of the worlds w reachable form w_0 in at most n steps through the accessibility relation R and the selection function f.

The valuation function $[[\ \ ]]^*$ is defined as:

$$[[p]]^* = \{[w] : w \in [[p]]\}$$

for all $p \in Var_F$[6]. It extends to more complex formulas in a standard way as:

$[[\top]]^* = W^*$ $[[\bot]]^* = \emptyset$
$[[A \wedge B]]^* = [[A]]^* \cap [[B]]^*$

[6]Observe that all worlds in the same equivalence class have the same propositional valuation

$[[\neg A]]^* = W^* - [[A]]^*$
$[[\Box A]]^* = \{[w] : [w]R^*[w']$ and $[w'] \in [[A]]^*\}$
$[[A > B]]^* = \{[w] : f^*(A, [w]) \subseteq [[B]]^*\}$,

where f^* and R^* are defined as follows: For $n > 0$:

$$f^*(A, [w]) = \begin{cases} \{[w'] : w' \in f(A, w_R)\} & \text{if } [w] \in W_n \\ \emptyset & \text{if } [w] \in W^* - W_n \end{cases}$$

where w_R is the representative element for the class $[w]$ and $A \in \mathcal{L}_F$, and:

for $[w] \in W_n$, R^* is the transitive closure of the relation R', where $R' = \{([w], [w']) : wRw'\}$
for $[w] \in W^* - W_n$, $R^* = \{([w], [w])\}$.

For $n = 0$, for all $[w] \in W^*$, $A \in \mathcal{L}_F$:

$$f^*(A, [w]) = \emptyset \text{ and } R^* = \{([w], [w])\}.$$

Observe that the definition of f^* at a world $[w]$ is based on the value of f on the representative element w of the class $[w]$. The definition of f^* depends on the choice of the representative element of the class $[w]$. However, it can be proved that: for $w_1, w_2 \in W$, if $w_1 \equiv_F w_2$ then the sets $f(A, w_1)$ and $f(A, w_2)$ can be regarded as being equivalent concerning the evaluation of formulas in the set $\mathcal{L}^>_{F,n-1}$, i.e. formulas containing nested conditionals and $\Box$ with at most $n-1$ levels of nesting. More precisely, let $\equiv_{F,n-1}$ be the equivalence relation obtained by replacing in the definition $\mathcal{L}^>_F$ with $\mathcal{L}^>_{F,n-1}$. If $w_1 \equiv_F w_2$ then for all $w'_1 \in f(A, w_1)$ there exists $w'_2 \in f(A, w_2)$ such that $w'_1 \equiv_{F,n-1} w'_2$. A similar property holds for the accessibility relation R^*: if $w_1 \equiv_F w_2$ then for all w'_1 such that $w_1Rw'_1$ there exists w'_2 such that $w_2Rw'_2$ and $w'_1 \equiv_{F,n-1} w'_2$. Observe that (as a difference with f^*) the definition of the accessibility relation R^* does not depend on the representative elements of the equivalence classes: if wRw' holds then $[w]R^*[w']$ holds.

We can prove that $M^*, w_0 \models F$. To this purpose, we first prove the two following lemmas that show which is the correspondence between the evaluation of formulas in the original model M and their evaluation in M^*. From the definition of $[[\ ]]^*$, it immediately follows that:

Lemma 5 For all formulas $G \in \mathcal{L}_F$, for all $w \in W$, $[w] \in [[G]]^*$ if and only if $w \in [[G]]$.

This property the property follows straightforwardly from the definition of $[[\ ]]^*$ and the induction on the structure of G. For the worlds $[w] \in W_n$, it extends to modal and conditional subformulas of F as follows:

Lemma 6 For all subformulas G of F with nesting level $nl(G) \leq k (\leq n)$, for all $[w_k] \in W_n$ such that $[w_0], [w_1], \ldots, [w_k]$ is a sequence of worlds reachable from $[w_0]$ in k steps (i.e. such that $[w_i]R^*[w_{i+1}]$ or $f^*([w_i], [w_{i+1}])$, for all $i = 0, k-1$):

$$[w_k] \in [[G]]^* \text{ if and only if } w_k \in [[G]].$$

Proof By double induction on k and on the structure of G.

If $k = n$ then nl(G)=0 (G is a boolean combination of propositional variables) and the property follows straightforwardly by Lemma 5.

Assume the property holds for $k + 1$. We have to prove it for k. Take G such that $nl(G) \leq k$.

We proceed by considering the different cases for G. If G is $G_1 \wedge G_2$,or $\neg G$ or another boolean combination of formulas (with at most k levels of nesting) we apply induction on the structure of G.

Let $G = C > D$, with $nl(C > D) = k > 0$. We have to prove that $[w_k] \in [[C > D]]^*$ if and only if $w_k \in [[C > D]]$.

In one direction, assume $w_k \in [[C > D]]$, i.e. for all $w' \in f(C, w_k)$, $w' \in [[D]]$. We have to prove $[w_k] \in [[C > D]]^*$, i.e. for all $[w_{k+1}] \in f^*(C, [w_k])$, $[w_{k+1}] \in [[D]]^*$. If $[w_{k+1}] \in f^*(C, [w_k])$, then there is a world $w_{kR} \in W$, which is the representative element of the equivalence class $[w_k]$, such that $w'_{k+1} \in f(C, w_{kR})$, and $w'_{k+1} \equiv_F w_{k+1}$. As $w_{kR} \equiv_F w_k$, then there is a world $w' \in f(C, w_k)$ such that $w' \equiv_{F,n-1} w'_{k+1}$. Hence, $w' \equiv_{F,n-1} w_{k+1}$. As $D \in \mathcal{L}^{>}_{F,n-1}$, $nl(D) \leq k - 1$ and, from the hypothesis, $w' \in [[D]]$, we have $w_{k+1} \in [[D]]$.

Observe that $[w_0], [w_1], \dots, [w_k], [w_{k+1}]$ is a sequence of worlds reachable from $[w_0]$ in $k + 1$ steps in M^* and $nl(D) \leq k - 1$. Hence, by inductive hypothesis, $[w_{k+1}] \in [[D]]^*$.

In the other direction, assume $[w_k] \in [[C > D]]^*$, i.e. for all $[w_{k+1}] \in f^*(C, [w_k])$, $[w_{k+1}] \in [[D]]^*$. We have to prove that $w_k \in [[C > D]]$, i.e. for all $w' \in f(C, w_k)$, $w' \in [[D]]$. Let w_{kR} be the representative element of the class $[w_k]$. As $w_{kR} \equiv_F w_k$, if $w' \in f(C, w_k)$, then there is a world $w_{k+1} \in f(C, w_{kR})$ such that $w' \equiv_{F,n-1} w_{k+1}$. By construction of M^*, $[w_{k+1}] \in f^*(C, [w_k])$ and, from the hypothesis, $[w_{k+1}] \in [[D]]^*$. From the fact that $nl(D) \leq k - 1$ and $[w_{k+1}]$ reachable from $[w_0]$ in $k + 1$ steps in M^*, we conclude by inductive hypothesis that $w_{k+1} \in [[D]]$. Moreover, as $w' \equiv_{F,n-1} w_{k+1}$ and $nl(D) \leq k - 1$ we can conclude that $w' \in [[D]]$.

Let $G = \Box A$. With $nl(\Box A) = k > 0$. We have to prove that $[w_k] \in [[\Box A]]^*$ if and only if $w_k \in [[\Box A]]$.

In one direction, assume $w_k \in [[\Box A]]$, i.e. for all w' s.t. $w_k R w'$, $w' \in [[A]]$. We have to prove $[w_k] \in [[\Box A]]^*$, i.e. for all $[w_{k+1}]$ s.t. $[w_k]R^*[w_{k+1}]$, $[w_{k+1}] \in [[A]]^*$. Let $[w']$ be such that $[w_k]R^*[w']$. There are two cases, that $([w_k][w']) \in R'$ and that $([w_k][w']) \notin R'$ and it is introduced in R^* by applying the transitive closure to R'.

In the first case, let $w' = w_{k+1}$ and $([w_k], [w_{k+1}]) \in R'$. Then there is a world $w'_k \in W$, which is in the equivalence class $[w_k]$ ($w'_k \equiv_F w_k$), such that $w'_k R w'_{k+1}$, and $w'_{k+1} \equiv_F w_{k+1}$. As $\Box A$ is a subformula of F, $\Box A \in \mathcal{L}^{>}_F$: from the hypothesis $w_k \in [[\Box A]]$, we get $w'_k \in [[\Box A]]$. Hence, $w'_{k+1} \in [[A]]$. Since, $w'_{k+1} \equiv_F w_{k+1}$, $w_{k+1} \in [[A]]$. Observe that $[w_0], [w_1], \dots, [w_k], [w_{k+1}]$ is a sequence of worlds reachable from $[w_0]$ in $k + 1$ steps in M^* and $nl(A) \leq k - 1$. Hence, by inductive hypothesis, $[w_{k+1}] \in [[A]]^*$.

In the second case, assume that $([w_k], [w_{k+n}]) \in R^*$, and that there is a sequence of worlds

$$[w_k], [w_{k+1}], [w_{k+2}], \dots, [w_{k+n-1}][w_{k+n}],$$

such that $[w_{k+i}]R^*[w_{k+i+1}]$, $(i = 0, n-1)$. Then there is a world $w'_k \in [w_k]$, a world $w''_{k+n} \in [w_{k+n}]$ and, for all $i = 1, n-1$ there are two worlds $w'_{k+i}, w''_{k+i} \in [w_{k+i}]$ such that, for all $i = 0, n-1$, $w'_{k+i} R w''_{k+i+1}$ and, for all $i = 1, n-1$, $w'_{k+i} \equiv_F w''_{k+i}$.

As $w_k \in [[\Box A]]$ and $\Box A \in \mathcal{L}_F^>$, as $w_k \equiv_F w'_k$, $w'_k \in [[\Box A]]$. By transitivity, $w'_k \in [[\Box\Box A]]$. As $\Box\Box A \in \mathcal{L}_F^>$, $w''_{k+1} \in [[\Box A]]$. But, from the equivalence, $w'_{k+1} \equiv_F w''_{k+1}$, $w'_{k+1} \in [[\Box A]]$. By a similar reasoning pattern, we can conclude, $w'_{k+2} \in [[\Box A]]$ $\dots$ $w'_{k+n} \in [[\Box A]]$. Hence, by reflexivity, $w'_{k+n} \in [[A]]$. From the equivalence $w'_{k+n} \equiv_F w_{k+n}$ we conclude $w_{k+n} \in [[A]]$.

As there is a sequence $[w_0], [w_1], \dots, [w_k], [w_{k+n}]$ of worlds reachable from $[w_0]$ through f^* and R^*, and $nl(A) \leq k-1$, by inductive hypothesis, $[w_{k+n}] \in [[A]]^*$.

In the other direction, assume $[w_k] \in [[\Box A]]^*$. We have to prove $w_k \in [[\Box A]]$, i.e. for all w_{k+1} s.t. $w_k R w_{k+1}$, $w_{k+1} \in [[A]]$.

Let $w_{k+1} \in W$ s.t. $w_k R w_{k+1}$. Then by construction, $[w_k]R^*[w_{k+1}]$ and, from the hypothesis, $[w_{k+1}] \in [[A]]^*$. As $[w_0], [w_1], \dots, [w_k], [w_{k+1}]$ is a sequence of worlds reachable from $[w_0]$ in $k+1$ steps in M^* and $nl(A) \leq k-1$, by inductive hypothesis, $w_{k+1} \in [[A]]$, which concludes the proof.

Q.E.D.

As an immediate consequence of the lemmas, we have the following Corollary.

Corollary 1 $M^*, w_0 \models F$.

We can now prove that M^* is a AC model, and that it is finite.

Theorem 3 *The model $M^* = \langle W_n, f^*, [[]]^* \rangle$ is an $AC-$structure.*

Proof We show that the selection function f^* satisfies the conditions $(S-RCEA)-(S-REFL)$.

In proving it we will consider two cases: that $nl(F) > 0$ and $nl(F) = 0$.

Let us prove the semantic properties for $nl(F) > 0$.

$(S-RCEA)$ Assume that, for $A, B \in \mathcal{L}_F$, $[[A]]^* = [[B]]^*$. We want to show that $f^*(A, [w]) = f^*(B, [w])$. From the hypothesis we have: $[[A]] = [[B]]$. In fact, $w \in [[A]]$ iff (by Lemma 5) $[w] \in [[A]]^*$ iff (from the hypothesis) $[w] \in [[B]]^*$ iff $[w] \in [[B]]$ (again by Lemma 5). Hence, from the fact that $(S-RCEA)$ holds for M, for $[w] \in W_n$,

$$f^*(A, [w]) = \{[w'] : w' \in f(A, w_R)\} = \{[w'] : w' \in f(B, w_R)\} = f^*(B, [w]).$$

For $[w] \in W^* - W_n$, $f^*(A, [w]) = \emptyset = f^*(B, [w])$.

$(S-CV)$ We have to show that if $f^*(A, [w]) \cap [[C]]^* \neq \emptyset$ then $f^*(A \wedge C, [w]) \subseteq f^*(A, [w])$.

Let $[w] \in W_n$. Assume that $f^*(A, [w]) \cap [[C]]^* \neq \emptyset$. Then, as $f^*(A, [w]) = \{[w'] : w' \in f(A, w_R)\}$ (where w_R is the representing element of the class $[w]$), there must be a $w' \in f(A, w_R)$ such that $[w'] \in [[C]]^*$. As $C \in \mathcal{L}_F$, by Lemma 5 $w' \in [[C]]$. Hence, $f(A, w_R) \cap [[C]] \neq \emptyset$. As $(S-CV)$ holds for M, $f(A \wedge C, w_R) \subseteq f(A, w_R)$. Therefore: $f^*(A \wedge C, [w]) = \{[w'] : w' \in f(A \wedge C, w_R)\} = \{[w'] : w' \in f(A, w_R)\} = f^*(A, [w])$.

Let $[w] \in W^* - W_n$. Then, $f^*(A, [w]) = \emptyset$, and the thesis holds trivially.

$(S-CA)$ We have to show that $f^*(A \vee B, [w]) \subseteq f^*(A, [w]) \cup f^*(B, [w])$.

Let $[w] \in W_n$. Then $f^*(A \vee B, [w]) = \{[w'] : w' \in f(A \vee B, w_R)\}$
$\subseteq \{[w'] : w' \in f(A, w_R)\} \cup \{[w'] : w' \in f(A, w_R)\}$ (as (S-CA) holds for M)
$= f^*(A, [w]) \cup f^*(B, [w])$.

Let $[w] \in W^* - W_n$. Then, $f^*(A \vee B, [w]) = f^*(A, [w]) = f^*(B, [w]) = \emptyset$, and the thesis holds trivially.

(S-CE) We have to prove that: if $f^*(ca, [w]) \subseteq [[B]]^*$ then

$$ValProp(f^*(ca, [w])) \subseteq ValProp(f^*(B, f(ca, [w]))).$$

Let $[w] \in W_n$. Assume that $f^*(ca, [w]) \subseteq [[B]]^*$. From this it follows that $f(ca, w) \subseteq [[B]]$. In fact, let $w' \in f(ca, w)$ and let w_R be the representative element of $[w]$. Then there is a world $w'' \in f(ca, w_R)$ with $w'' \equiv_{F,n-1} w'$. Hence, by construction, $[w''] \in f(ca, [w_R])$ and, from the hypothesis, $[w''] \in [[B]]^*$. As $B \in \mathcal{L}_F$, by Lemma 5, $w'' \in [[B]]$ and, from $w'' \equiv_{F,n-1} w'$, we get $w' \in [[B]]$.

As a consequence of the fact that $f(ca, w) \subseteq [[B]]$ and the fact that (S-CE) holds for the model M, we have that $ValProp(f(ca, w)) \subseteq ValProp(f(B, f(ca, w)))$. We can now prove that $ValProp(f^*(ca, [w])) \subseteq ValProp(f^*(B, f(ca, [w])))$.

Let $[w'] \in f^*(ca, [w])$. Then there is $w_1 \in f(ca, w_R)$ such that $w_1 \equiv_F w'$. From (CE) in M, there must be two worlds w_2 and w_3 such that $w_2 \in f(ca, w_R)$ and $w_3 \in f(B, w_2)$ so that w_1 and w_3 have the same propositional valuation: $ValProp(w_1) = ValProp(w_3)$. By construction $[w_2] \in f(ca, [w])$ and, if we let w_2R to be the representative element of $[w_2]$, there is a world $w_3' \in f(B, w_2R)$ with $w_3' \equiv_{F,n-1} w_3$. Clearly, $ValProp(w_3') = ValProp(w_3)$ and, therefore, $ValProp(w_3') = ValProp(w_1)$. Moreover, by construction, $[w_3'] \in f^*(B, [w_2])$, so that $[w_3'] \in f^*(B, f(ca, [w]))$.

Let $[w] \in W^* - W_n$. Then, $f^*(ca, [w]) = \emptyset$, and the thesis holds trivially.

(S-MOD) if $[w'] \in f^*(ca, [w])$ then $R^*([w], [w'])$.

Let $[w] \in W_n$. If $[w'] \in f^*(ca, [w])$ then, by construction, $w' \in f(ca, w_R)$ (where w_R is the representing element of the class $[w]$). Hence, as (MOD) holds for M, $R(w_R, w')$, and by construction $R^*([w], [w'])$.

Let $[w] \in W^* - W_n$. Then $f^*(ca, [w]) = \emptyset$, and the thesis holds trivially.

(S-TRANS) if $R^*([w], [w'])$ and $R^*([w'], [w''])$ then $R^*([w], [w''])$, for all $[w]$, $[w']$, $[w''] \in W^*$.

Let $[w] \in W_n$. By construction, since R^* is defined as the transitive closure of R'.

Let $[w] \in W^* - W_{n-1}$. Then $R^* = ([w], [w])$ and the thesis holds trivially.

(S-REFL) $R^*([w], [w])$, for all $[w] \in W^*$.

Let $[w] \in W_{n-1}$. Since R is reflexive, $R(w, w)$. Hence, by construction,$R^*([w], [w])$.

Let $[w] \in W^* - W_{n-1}$. Then $R^* = ([w], [w])$ and the thesis holds trivially.

For $nl(F) = 0$ the semantic properties hold and their proofs are the same as for the case $nl(F) > 0$, for $[w] \in W^* - W_n$. Q.E.D.

Theorem 4 *The model M^* is finite.*

Proof By construction, since $n < \omega$ and W^* is finite. Q.E.D.

The number of worlds in W^* cannot be more than the number of truth assignments to the formulas in $\mathcal{L}_F^{\geq}$. As the number of such formulas is $O(2^{|F|})$, then the number of worlds in W^* is at most double exponential in $|F|$, i.e. $|W^*| = O(2^{2^{|F|}})$. This provides a non-deterministic algorithm to decide the satisfiability of a formula in double exponential time: first non-deterministically construct a model of size double exponential in $|F|$; then verify that it is a AC-model of F. This verification requires a time which is linear in the size of the model and hence it requires double exponential time.