Dynamic Causal Calculus

Alexander Bochman
Computer Science Department,
Holon Institute of Technology,
Israel

Abstract

We introduce dynamic causal calculus, a nonmonotonic formalism that can be viewed as a direct logical counterpart of the action description language C+ from (Giunchiglia et al. 2004). We formulate a nonmonotonic semantics of the associated causal language, and compare this semantics with the indirect, two-stage semantics for C+, given in (Giunchiglia et al. 2004). It will be shown, in particular, that the suggested semantics allows us to alleviate syntactic distinctions between propositional atoms, maintained by C+, as well as type restrictions imposed on its causal laws. We will describe also a logical formalism of dynamic causal inference that constitutes a complete description of the logic that is adequate for this dynamic calculus.

Introduction

Unlike general descriptions of temporal dynamics in temporal logics and related logical formalisms, action theories in AI have to deal primarily with two quite specific reasoning tasks, namely the *prediction* task (what are the results of a given sequence of actions from an initial state) and the *planning* task (what sequence of actions could lead from an initial state to a target goal state). These reasoning tasks immediately lead to a triple of famous problems, known as the frame, ramification and qualification problems (see (Shanahan 1997)). It was realized quite early that classical logic and its temporal/dynamic extensions, taken by themselves, encounter difficulties in resolving these problems. More precisely, it has become clear that these problems have an essentially *nonmonotonic* character, so their proper solution requires augmenting purely logical, monotonic reasoning with an appropriate mechanism for making nonmonotonic conclusions[1].

In recent years a dominant approach to solving these problems has been based, in one form or another, on what can be broadly termed causal reasoning. Given a set of action and (causal) rules describing the domain, the causal approach employs a distinction between facts that hold in a situation versus facts that are caused (or explained) by other facts and the rules. In this context, the corresponding causal closure assumption (see. e.g., (Reiter 2001)) can be viewed as a particular form of an old philosophical principle of *universal causation*, which amounts to the requirement that all facts that hold in a situation should be either caused by other occurrent facts, or else preserve their truth-values in time (due to the accompanying *inertia* assumption). A direct incorporation of such causal assertions into the language of the situation calculus has been proposed in (Lin 1995; 1996), and has been shown to provide a natural account of both the frame and ramification problems. Subsequently, a general formal framework for this kind of causal reasoning, called a causal calculus, has been suggested in (McCain and Turner 1997).

An elaborate implementation of the above causal principles in reasoning about actions has been given in (Giunchiglia et al. 2004). The formalism of (Giunchiglia et al. 2004), however, is a multi-sorted and multi-layered representation framework. As its top layer, it employs a causal action description language C+ that provides high-level descriptions of action domains in terms of three kinds of propositional atoms (actions, simple fluents and statically determined fluents) and three different kinds of causal laws (static laws, action dynamic laws and fluent dynamic laws). Domain descriptions in this language are then instantiated by assigning temporal stamps to propositions, and incorporating the resulting descriptions into an atemporal causal calculus of (McCain and Turner 1997). The models of the resulting causal theories are viewed then as intended models of the source, higher-level action descriptions.

In this study[2] we will attempt to single out and 'streamline' the logical framework behind the language C+. To this end, we will introduce a dynamic generalization of the original causal calculus, which will be formulated, ultimately, in terms of a single basic kind of *dynamic* causal rules. This dynamic calculus will provide a direct and uniform logical description for the language C+. In addition, we will describe also a logical (monotonic) system of dynamic causal inference that will constitute a concise logical framework for causal reasoning in dynamic domains.

[1]Even the well-known "monotonic" solution to the frame problem, described in (Reiter 1991), can be seen as a result of compiling this nonmonotonic information into the successor state equivalences.

[2]Preliminary version of this paper has appeared in (Bochman 2013).

The plan of the paper is as follows. After an overview of the original (atemporal) causal calculus and its use in the framework of (Giunchiglia et al. 2004), we will introduce a dynamic causal calculus as a direct reformulation of the action description language C+. We will describe a modified nonmonotonic semantic for this dynamic causal language that is based on a comprehensive principle of universal causation, and show that this modification allows us, in particular, to alleviate most of the syntactic restrictions imposed in C+ on the form of its causal rules, as well as syntactic distinctions between the propositions of the language. Finally, we will describe a logical formalism of dynamic causal inference that constitutes an underlying logic for causal reasoning in action domains.

The Causal Calculus

Throughout this study, we will assume that our basic language is a classical propositional language with the usual connectives and constants $\{\wedge, \vee, \neg, \rightarrow, \mathbf{t}, \mathbf{f}\}$. $\vDash$ and Th will stand, respectively, for the classical entailment and the associated logical closure operator. In what follows, we will also identify propositional interpretations ('worlds') with the sets of propositional formulas that hold in them.

A *causal rule* is a rule of the form $A \Rightarrow B$, where A and B are classical propositions. We will informally interpret such rules as saying plainly *"A causes B"*.[3]

By a *causal theory* we will mean an arbitrary set of causal rules. For a set u of propositions and a causal theory Δ, we will denote by $\Delta(u)$ the set of all propositions that are caused by u in Δ, that is,

$$\Delta(u) = \{B \mid A \Rightarrow B \in \Delta, \text{ for some } A \in u\}$$

Then the nonmonotonic semantics of a causal theory can be defined as follows.

Definition 1. A world (= propositional interpretation) α is an *exact model* of a causal theory Δ if it is a unique model of $\Delta(\alpha)$. The set of exact models forms a *nonmonotonic semantics* of Δ.

The above semantics of causal theories coincides, in effect, with the semantics for such theories, described in (McCain and Turner 1997) and (Giunchiglia et al. 2004). It can also be verified that exact models of a causal theory are precisely the worlds that satisfy the following condition:

$$\alpha = \mathrm{Th}(\Delta(\alpha)).$$

Accordingly, exact worlds are not only closed with respect to the causal rules, but also such that any proposition that holds in them is caused (that is, explained) ultimately by other propositions.

An overview of the Language C+

The underlying propositional language of the action description language C+, introduced in (Giunchiglia et al. 2004), is somewhat more general than the standard classical propositional language in that it is based on a multi-valued propositional signature that consists of a set of constants, along with a function Dom assigning every constant c a nonempty finite set $Dom(c)$ of values. Propositional atoms in this signature are expressions of the form $c = v$, where c is a constant, while v is one of its possible values. Still, propositional formulas in this language are defined as usual combinations of atoms with the help of the ordinary classical connectives. Moreover, *Boolean* constants are defined in this setting as a special kind of constants whose domain is the set $\{\mathbf{f}, \mathbf{t}\}$ of truth values.

An *action description* in C+ is defined as a set of causal laws. There are, however, three kinds of causal laws in C+, and the differences between them are based, ultimately, on a distinction between three kinds of constants, and thereby three kinds of atoms, stipulated by the theory. First, propositional atoms are partitioned into *action* atoms and *fluent* atoms, while the latter are further partitioned into *simple* and *statically determined* fluents. A *fluent formula* is defined as a formula such that all constants occurring in it are fluent constants, whereas an *action formula* is a formula that contains at least one action constant and no fluent constants.

Granted the above syntactic distinctions among propositional atoms, the following three kinds of causal laws are defined in C+:

- *Static laws* are expressions of the form

  $$\textbf{caused } F \textbf{ if } G,$$

  where F and G are fluent formulas;
- *Action dynamic laws* are expressions of the form

  $$\textbf{caused } F \textbf{ if } G,$$

  where F is an action formula and G is a formula;
- *Fluent dynamic laws* are expressions of the form

  $$\textbf{caused } F \textbf{ if } G \textbf{ after } H,$$

  where F is a fluent formula that does not contain statically determined constants, G is a fluent formula, and H is an arbitrary formula.

Static laws are used in C+ to talk about causal dependencies between fluents in the same state, while action dynamic laws are purported to express causal dependencies between concurrently executed actions. In accordance with their very name, statically determined fluent constants are allowed in the heads of static laws, but not in the heads of dynamic laws. As we will see in what follows, the necessity of introducing a separate syntactic sort of statically determined fluents stems from a particular definition of a state of a transition system, which has been used in interpreting action descriptions in C+.

The language C+ employs also a number of abbreviations and names for special kinds of causal laws. Two such abbreviations play an especially important role in the descriptions of action domains that serve to illustrate the general theory. If c is a simple fluent constant, then

$$\textbf{inertial } c$$

[3](Giunchiglia et al. 2004) adopted a more cautious informal reading of such rules, namely *"If A holds, then B is caused"*.

stands for the fluent dynamic laws

$$\textbf{caused } c = v \textbf{ if } c = v \textbf{ after } c = v$$

for all $v \in Dom(c)$. These laws provide an encoding of the *inertia assumption* for (simple) fluent atoms. Similarly, if c is an action constant, the expression

$$\textbf{exogenous } c$$

stands for the action dynamic laws

$$\textbf{caused } c = v \textbf{ if } c = v$$

for all $v \in Dom(c)$. These laws make action atoms exogenous in an action domain, which exempt them, in effect, from explanation.

Interpretations and models of action descriptions in C+ are defined indirectly by translating them into plain causal theories. To begin with, for every natural number m, an action description D is transformed into an atemporal causal theory D_m as follows. First, 'time stamps' i: for $i \in \{0, \dots, m\}$ are inserted in front of every occurrence of every atom in propositional formulas. Then each static law is translated into the following set of causal rules, for every $i \leq m$:

$$i : A \Rightarrow i : B,$$

where $i : F$ is the result of inserting i: in front of every occurrence of every atom in a formula F. Similarly, any action dynamic law is translated into a set of causal rules of the same form, but only for $i < m$.

Finally, any fluent dynamic law is translated into the following set of causal rules:

$$(i : C) \wedge (i + 1 : A) \Rightarrow i + 1 : B,$$

for every $i < m$.

As a concluding step, in order to deal with the initial states, the following causal rules are added to the resulting causal theory:

$$0 : l \Rightarrow 0 : l$$

for every *simple* fluent atom l. These rules make simple fluent atoms exogenous (self-explained) in the initial state. As a result, we obtain an ordinary causal theory, and the exact models of this theory are considered to be the models of the original action description in C+. Such models can be visualized as histories of length m of the source dynamic domain. More precisely, these are histories in which the initial state is 'self-explainable', but every subsequent state is already causally explained by the preceding state and actions taken in it.

(Giunchiglia et al. 2004) contains also a more general semantic construction, according to which an action description D in C+ describes, in effect, a *transition model* (i.e., a set of states with a set of transitions among them) in which states are the models of the 'smallest' (static) causal theory D_0 (which is a theory D_m for $m = 0$)[4], while transitions correspond precisely to the models of the minimal dynamic causal theory D_1 (that is, a theory D_m for $m = 1$). It has been shown in (Giunchiglia et al. 2004, Proposition 8) that, for any $m > 0$, models of a causal theory D_m are exactly histories (paths) of length m in this transition model.

[4]Note that, by the definition of the translation, D_0 includes no dynamic laws, and only '0-stamped' static causal laws.

On statically determined fluents

The above construction of a canonical transition model for an action description implicitly relied on the property (stated as Proposition 7 in (Giunchiglia et al. 2004)) that for any transition $\langle s, e, t \rangle$ in the above sense (that is, for any model of D_1), both s and t are states (models of D_0). As has been observed in the paper, the validity of this property depends essentially on the fact that the heads of fluent dynamic laws were not allowed to contain statically determined fluent constants, which explained the very need for a *syntactic* (type) separation between simple and statically determined fluent constants in the framework of the general theory described in (Giunchiglia et al. 2004).

It should be noted, however, that, taken by itself, the distinction between simple and statically determined fluents is not as clear-cut as it may seem. Consider, for example, a fluent predicate *Safe* (for a baby), discussed first in (Myers and Smith 1988). As was observed in (Giunchiglia and Lifschitz 1995), there might be different kinds of safety in this sense. An object in question may be a heavy hammer, so it is safe for a baby only if it is far from the reach, say on a table. This is a clear example of a statically determined fluent. But the object may be a doll, in which case its safety is clearly an ordinary, simple inertial fluent. Furthermore, we can easily imagine an action like *Isolate* that wrap up a potentially non-safe hammer with a material that makes it fully safe for a baby. On an account of (Giunchiglia et al. 2004) this would require, however, a syntactic shift for this fluent from statically determined to simple one.

The distinction between simple and statically determined fluents is only a very special case of a broader distinction between inertial and non-inertial fluents. Such distinctions between fluents play an important role in reasoning about actions and change in AI, so they should not be obliterated, or neglected. Still, they need not be defined as syntactic distinctions between types of fluents. Rather, they could be viewed as a by-product of the actual ways these fluents are described, or defined, in (causal) rules. Thus, we can still maintain that a particular fluent is *statically determined* in a given action description if it happens to appear only in heads of its static rules. But this characterization could change, for example, with an addition of new dynamic rules which involve this fluent in their heads. Such a 'dynamic' characterization would be unfeasible, however, in the framework of (Giunchiglia et al. 2004) because statically determined fluents play an essential role in the definition of the very notion of a state of a transition model that determines, in turn, the semantics of C+. Accordingly, in order to make the language of C+ syntactically uniform, we should change also its semantic description. This will be done in the next section where we will describe an alternative semantics for the dynamic causal calculus that will provide a logical representation for the action descriptions in C+.

As a final note here, we should mention that there are also other, perhaps even more serious, problems with the representation of statically determined fluents in C+. Thus, many fluents of this kind are naturally defined using recursive (Prolog-like) rules. It turns out, however, that such definitions are not expressible in C+, though they are naturally expressible in an alternative action description language, namely the language B (see (Gelfond and Lifschitz 1998)). As a matter of fact, representation of such recursively determined fluents creates similar difficulties also for the dynamic causal calculus, described in this study, since the latter closely follows the general semantic framework of C+[5].

Dynamic Causal Calculus

In this section we are going to describe a dynamic causal calculus that is intended to provide a direct logical reformulation of the action description language C+ and its semantics. As a first step, we will use a more convenient, 'logic-oriented' notation, namely, we will rewrite a dynamic law **caused** B **if** A **after** C as a causal rule of the form

$$C.A \Rightarrow B.$$

We will call such rules *dynamic causal rules*. Moreover, extending our previous re-interpretation of atemporal causal rules, we will assign a more 'active' informal reading to such rules, namely *"After C, A causes B"*. By the intended interpretation, a rule $C.A \Rightarrow B$ describes a dynamic transition from a state that satisfies proposition C to a subsequent state in which B is caused by A.

Remark. As a matter of fact, the above dynamic causal rules are somewhat ambiguous from a syntactic point of view. On a more abstract level, such a rule can be viewed simply as an instantiation of a primitive ternary propositional operator. There are, however, at least two other, more articulated possibilities[6]. Thus, a dynamic causal rule $C.A \Rightarrow B$ could be viewed as a plain causal rule $A \Rightarrow B$ that is conditioned by a preceding context C. In fact, this 'parsing' agrees with the informal reading of such rules, given above. Furthermore, we will see that this understanding of dynamic causal rules as conditional static rules provides a natural justification for the postulates of the associated dynamic causal inference that will be given below. Still, a different parsing possibility consists in viewing such rules as binary causal rules, though with complex premises consisting of pairs of propositions (C, A). Again, we will see later that this reading can also be given a formal support, due to a possible translation of such rules as propositions of the form $(C \circ A) \to B$ in arrow logic.

As a second step in our reformulation, we will uniformly view both static and action dynamic laws of the form **caused** B **if** A as plain (static) causal rules $A \Rightarrow B$. Moreover, in the version of the dynamic causal calculus that we will present in this study, we will make one further step and identify such rules with a special kind of dynamic rules of the form $\mathbf{t}.A \Rightarrow B$, where $\mathbf{t}$ is the truth constant. In other words, we will adopt the following definition:

$$A \Rightarrow B \ \equiv_{df} \ \mathbf{t}.A \Rightarrow B.$$

According to this definition, static causal rules are precisely rules that are valid after any legitimate transition. The consequences and variations created by this identification will be discussed below.

The above reduction of static causal rules to a special kind of dynamic rules makes the dynamic causal rules the only kind of rules of the dynamic calculus. In accordance with this, a *dynamic causal theory* will be defined below simply as a set of dynamic causal rules.

Now we are going to provide a direct description of the nonmonotonic semantics of dynamic causal theories. The guiding principle behind this nonmonotonic semantics will be a thorough enforcement of the *principle of universal causation*, according to which every state of a dynamic model should be explained (i.e., caused) by a preceding state and the causal rules of the domain.

Given a dynamic causal theory Δ and worlds α, β, we will denote by $\Delta(\alpha.\beta)$ the set

$$\{C \mid A.B \Rightarrow C \in \Delta \text{ for some } A \in \alpha, B \in \beta.\}$$

For any given state β, the set $\Delta(\alpha.\beta)$ is precisely the set of propositions in β that are caused due to a transition from α.

Definition 2.(i) A pair (α, β) of worlds will be called an *exact transition* with respect to a dynamic causal theory Δ if β is the unique model of $\Delta(\alpha.\beta)$, that is

$$\beta = \mathrm{Th}(\Delta(\alpha.\beta)).$$

(ii) An *exact transition model* of a dynamic causal theory Δ is a set of worlds $\mathfrak{I}$ such that, for any $\beta \in \mathfrak{I}$ there is $\alpha \in \mathfrak{I}$ such that (α, β) is a an exact transition wrt Δ.

An exact transition is a transition between two states in which the resulting state is fully explained (caused), given the preceding state and the causal laws of the domain. It is important to note that if (α, β) is an exact transition, then the output world β is always closed with respect to the static causal rules (for our definition of the latter), namely if $A \Rightarrow B$ (that is, $\mathbf{t}.A \Rightarrow B$) belongs to Δ and $A \in \beta$, then $B \in \beta$. Note now that any world of an exact transition model is also an output of some exact transition. Accordingly, we immediately obtain the following

Lemma 1. *Any world of an exact transition model of a dynamic causal theory Δ is closed with respect to the static causal rules of Δ.*

It can be easily verified that the union of two exact transition models is also an exact transition model. Consequently, if a dynamic causal theory has at least one exact transition model, it has a unique maximal such model. We will call the latter the *canonical exact transition model* of a dynamic causal theory.

In the next section we are going to compare the above dynamic causal calculus with the theory presented in (Giunchiglia et al. 2004).

[5]Recently, (Lee, Lifschitz, and Yang 2013) have suggested a generalization of the language C+, called a language BC, that incorporates expressive capabilities of the language B. Instead of a causal interpretation, however, the authors have provided only formal translations of this language directly into the framework of logic programming.

[6]Compare with a similar compositional analysis of general and relevant conditionals, discussed in (Beall et al. 2012).

Comparisons with C+

In accordance with the reformulation procedure sketched in the preceding section, any action description in the action description language C+ can be immediately transformed into a dynamic causal theory in our sense. Despite the obvious similarities, however, this reformulation reveals also some basic differences between the two formalisms. To begin with, the language of the above dynamic causal calculus is thoroughly *uniform* in that it does not presuppose any a priori, *syntactic* distinctions between its propositional atoms, and it employs only a single form of causal rules, instead of three kinds of causal laws used in C+. On the other hand, the above notion of an exact transition model is somewhat different, and apparently more restricted, than the corresponding notion of a model for C+. In what follows we will discuss these differences in more details.

Fluents versus Actions

Many action theories in AI, as well as some general dynamic formalisms, maintain rigid semantic and syntactic distinctions between *fluent* propositions that describe particular states or situations and *actions* that describe transitions between states. Thus, the situation calculus (Reiter 2001) treats actions essentially as modifiers, or functions on situations. Similarly, Propositional Dynamic Logic (see (Harel, Kozen, and Tiuryn 2000)) interprets actions as binary relations on states.

However, in contrast to the above formalisms, it can be immediately observed that there are no *inherent* syntactic differences between fluent and action atoms in C+; both kinds are translated as (temporally indexed) propositional atoms in the underlying causal calculus. Of course, there are still important differences between these two kinds of propositions: (simple) fluents are governed by the inertia principle, while action atoms are normally treated in C+ as exogenous (see above). However, these distinctions are semantical, and they can be secured by incorporating causal rules that characterize, respectively, inertia and exogeneity for these atoms directly into the corresponding action descriptions. Once we add such rules to the action descriptions, there is no need to maintain a syntactic distinction between fluents and actions in C+, and there is no need to maintain a separate category of action dynamic laws in addition to static laws.

Remark. The treatment of actions in C+ makes the latter much similar to *temporal* formalisms for analyzing general computation processes, such as Linear Temporal Logic (LTL). In the latter, computational processes are represented as plain temporal sequences of states (which are characterized by fluent propositions), and they do not use explicit action descriptions. Still, actions are usually encoded in these formalisms by using associated 'action fluents' that appropriately constrain such temporal sequences (see, e.g., (Calvanese, De Giacomo, and Vardi 2002; De Giacomo and Vardi 2013)).

State Constraints and the Principle of Universal Causation

An exact transition model of a dynamic causal theory was defined above as a set of states in which every state is caused as a result of some exact transition. As a result, any state of such a model will be closed with respect to the static laws on our reformulation of the latter. Moreover, it can be verified that any such state will be a state in the sense of (Giunchiglia et al. 2004). Consequently, we will obtain that any exact model in our sense will correspond to a model (transition system) in the sense of (Giunchiglia et al. 2004):

Theorem 2. *If Δ is an action description in C+ and Δ_C its corresponding dynamic causal theory, then any exact transition model of Δ_C is a transition model of Δ.*

Still, our semantics for dynamic causal theories is more restrictive, since it requires that any state of the model, *including the initial state*, should be an output of some transition, whereas (Giunchiglia et al. 2004) adopted more relaxed requirements on initial states. As we have seen earlier, the construction of a causal theory for a given action description in C+ necessarily involves an addition of causal rules that make simple fluents exogenous (self-explainable) in the initial state. As a result, the framework (Giunchiglia et al. 2004) exempts to some extent initial states from the need of explanation, though it still requires that such states should be closed with respect to the static laws (which are completely separated from dynamic ones) and, moreover, that any statically determined fluent literal that holds in an initial state should still be explained (caused) by the static laws.

Remark. As a matter of fact, our definition of an exact transition almost coincides with the corresponding definition of a *causally explained transition*, given in (Giunchiglia and Lifschitz 1998) for a more restricted action description language C, a predecessor of C+[7]. In fact, the only difference between the two definitions is that (Giunchiglia and Lifschitz 1998) required further that both the initial and resulting states of such a transition should be closed with respect to the static causal laws. On our construction, this additional requirement is accounted for, respectively, as a by-product of our definition of static causal rules on the one hand (for the resulting states), and an exact transition model on the other hand (for the initial states).

In order to make a proper assessment of the above discrepancy, we should distinguish two aspects of the difference, conceptual and practical one. On the conceptual side, we believe that an ultimate reason for imposing even the above minimal restrictions on initial states in C+ stems from a broader requirement that such states should be somehow accessible in accordance with the laws of the domain. In other words, any state of a dynamic system should be consistently viewed as a result of some legitimate transition (including possible 'loops' in this state). Speaking more generally, we contend that static laws and constraints should be viewed as constraints that are effective after every legitimate transition, and vice versa, any constraint that happens to hold after any

[7] I am grateful to Vladimir Lifschitz for pointing this out (personal correspondence).

possible transition should be considered as a static law of the domain.

The suggested semantics of exact transition models allows us to treat static causal rules as a special case of general dynamic causal rules. In addition, it allows us to remove the syntactic distinction between simple and statically determined fluents. Furthermore, the principle of universal causation, used as a conceptual basis of this semantics, implies also some important consequences concerning general principles of reasoning in dynamic domains.

Generally speaking, the principle of universal causation provides logical foundations for *abductive reasoning*, namely for backward reasoning from effects to their causes (see (Bochman 2007)). This kind of reasoning constitutes an essential part of our commonsense reasoning, especially in reasoning about actions and change, and it occupies also an important part of reasoning in current action theories in AI. For example, the well-known regression method ((Reiter 2001)) can be viewed as a systematic implementation of abductive reasoning in the situation calculus.

Due to its causal foundations, abductive inference is also an essential, though implicit, part of the representation framework of C+; this is because every state of a transition model for an action description, *except the initial one*, is explained as an output of some causal transition. The associated abductive explanation may sanction, in particular, some further static constraints for such states, constraints that arise as a by-product of the fact that the state in question is a result of a particular action with some further effects. By the same token, however, the initial states in C+ are exempted from the abductive explanation of this kind, which, at least in some regular domains, may lead to a loss of important information about these states.

Persistent Action Domains

In case we don't accept the above understanding of static laws, we should maintain two separate kinds of state constraints, dynamic and purely static ones, as is actually done in (Giunchiglia et al. 2004)), as well as in many other action formalisms. This separation would allow us to include a broader class of transition models as admissible models for action descriptions, namely models that involve (initial) states that need not be a result of some transition, though they still satisfy the static laws of the domain. Moreover, we can easily construct some artificial, though logically consistent, action domains for which such a distinction would be necessary, namely action domains in which there are distinguished initial states that lack some property that holds for any state that results from a transition. Still, to find 'real-life' examples of such domains is really difficult (beyond the famous 'Big Bang' exception in Physics). This naturally brings us to the practical side of the difference between the two formalisms, namely to the question whether, and if so, how much we miss in restricting the semantics of action descriptions to exact transition models in our sense.

It turns out that for a fairly broad class of action descriptions in the language C+ (including all the examples given in (Giunchiglia et al. 2004)), we can *guarantee* in advance that any transition model of C+ that satisfies a given action description can be extended to an exact transition model in our sense. For such action descriptions, our dynamic causal calculus provides the same answers to the queries as the original theory of (Giunchiglia et al. 2004)). In what follows, we will illustrate this correspondence for a broad class of what we will call persistent action domains.

Informally speaking, a persistent action domain is a dynamic domain that does not involve involuntary, 'natural' actions that lead to unavoidable, necessary changes of some state. For such action domains, any legitimate state either remains persistent in the absence of any further actions upon it (alias after an action *Wait*), or else it can be forced to persist by using suitable (voluntary) actions.

Formally, by a *persistent action domain* we will mean any action description D in the language C+ such that for any state s of D (that is, for any model of D_0), there is a consistent transition (i.e., a model of D_1) from s to s.

It turns out to be surprisingly difficult to express the above semantic property of transition models in terms of some syntactic restrictions on action descriptions in C+. Still, it is easy to verify the validity of this persistence property for many descriptions used in action theories. Thus, in the central Monkey and Bananas example from (Giunchiglia et al. 2004) every simple fluent is inertial, so any state of the transition model can be shown to persist in the absence of actions. On the other hand, for the case of non-inertial fluents that tend to change by themselves, the corresponding action descriptions usually contain actions that make them persist. Thus, the pendulum domain from (Giunchiglia et al. 2004) involves an action *Hold* that keeps the position of the pendulum (which tends to sway otherwise).

Now, in persistent action domains, any state of a transition model can be consistently viewed as a result of some transition, which immediately leads to the following

Theorem 3. *If Δ is a persistent action description in C+ and Δ_C its corresponding dynamic causal theory, then the canonical transition model of Δ coincides with the exact transition model of Δ_C.*

As a matter of fact, the above correspondence between C+ and the dynamic causal calculus can be extended, though in a somewhat weaker form, even beyond persistent action domains. To this end, we should note that practically all queries that are usually formulated in action formalisms are 'future-oriented', namely they ask whether there is a path in a transition model from a given initial state that satisfies certain further requirements (such as whether it ends in a target goal state). Now, in many cases we can extend the source action description in C+ with some auxiliary actions that will allow us again to reconstruct any given initial state as an output of some exact transition (though in the extended action description). Moreover, this can be done without changing the future of the original states of the transition model, but only by augmenting their 'past'. As an immediate consequence, we will obtain that, for such action domains, the dynamic causal calculus will provide the same answer to such queries as the original action descriptions in C+.

Causal Inference

The framework described in the preceding section is a typical example of a nonmonotonic formalism; conclusions that can be obtained on the basis of the exact transition semantics can change non-monotonically if we add some further facts or causal rules to the original dynamic causal theory. Still, as with other formalisms for nonmonotonic reasoning (see (Bochman 2011)), the causal rules of the dynamic causal calculus presuppose a certain underlying logic that agrees with the above nonmonotonic semantics. Such a logic will provide us with a formal description of the associated dynamic causal inference.

Causal Inference Relations

The original causal calculus of (McCain and Turner 1997) has been defined only semantically, but (Bochman 2004) has described a logical formalism of causal inference relations, which has been shown to provide a complete formalization of a logical (monotonic) reasoning in causal theories. From a logical point of view, causal inference relations were defined as sets of causal rules that were required to satisfy almost all the usual postulates of classical inference, except Reflexivity $A \Rightarrow A$. The latter feature has turned out to be essential for an adequate representation of causal reasoning.

Definition 3. A *causal inference relation* is a relation $\Rightarrow$ on the set of propositions satisfying the following conditions:

(Strengthening) If $A \vDash B$ and $B \Rightarrow C$, then $A \Rightarrow C$;

(Weakening) If $A \Rightarrow B$ and $B \vDash C$, then $A \Rightarrow C$;

(And) If $A \Rightarrow B$ and $A \Rightarrow C$, then $A \Rightarrow B \wedge C$;

(Or) If $A \Rightarrow C$ and $B \Rightarrow C$, then $A \vee B \Rightarrow C$;

(Cut) If $A \Rightarrow B$ and $A \wedge B \Rightarrow C$, then $A \Rightarrow C$;

(Truth) $\mathbf{t} \Rightarrow \mathbf{t}$;

(Falsity) $\mathbf{f} \Rightarrow \mathbf{f}$.

The rule Or permits reasoning by cases; this feature can be seen as one of the main advantages of causal reasoning as compared with, say, default logic. It indicates that the causal logic is an *objective* (extensional) logical system, a system of reasoning about the world. In this respect, it is similar to classical logic, and distinct from modal (intensional) formalisms that deal primarily with beliefs and knowledge.

Yet another important feature of causal inference stems from the validity of the following rule:

(Coherence) If $A \Rightarrow \neg A$, then $A \Rightarrow \mathbf{f}$.

The above rule says that if a proposition causes propositions that are incompatible with it, then it is causally inconsistent. This feature indicates, in effect, that the above notion of causal inference is *atemporal*. For example, the rule $p \wedge q \Rightarrow \neg q$ cannot be understood as saying that p and q jointly cause $\neg q$ (afterwards) in a temporal sense; instead, by Coherence it implies $p \wedge q \Rightarrow \mathbf{f}$, which means, in effect, that $p \wedge q$ cannot hold. Just as in classical logic, however, a representation of temporal domains in this formalism can be obtained by adding explicit temporal arguments to propositions; this is what has been actually done in the action description framework of (Giunchiglia et al. 2004).

A possible worlds semantics

A semantic interpretation of causal inference relations can be given in terms of ordinary possible worlds (or Kripke) models (W, R, V), where W is a set of possible worlds, R a binary accessibility relation on W, and V a function assigning each world a propositional interpretation. Intuitively, $R\alpha\beta$ means that α is an initial state, and β a possible output state of a causal process. A Kripke model is *quasi-reflexive* if it satisfies the condition that if $R\alpha\beta$, then $R\alpha\alpha$.

Definition 4. A rule $A \Rightarrow B$ is *valid* in a Kripke model (W, R, V) if, for any worlds α, β such that $R\alpha\beta$, if A holds in α, then B holds in β.

By a set of causal rules *determined* by a Kripke model we will mean the set of rules that is valid in it. It can be verified that such a set satisfies all the postulates of causal inference. Moreover, as for other modal formalisms, a suitable construction of a canonical semantics allows us to obtain the corresponding completeness result:

Proposition 4. *A set of causal rules forms a causal inference relation if and only if it is determined by some quasi-reflexive Kripke model.*

As a by-product, the above semantics immediately sanctions a simple modal representation of causal rules. Namely, let $\Box$ be the usual modal operator definable in a possible worlds model: $\Box A$ holds in α iff A holds in all β such that $R\alpha\beta$. Then the validity of $A \Rightarrow B$ in a possible worlds model is equivalent to validity of the formula $A \to \Box B$. Consequently, causal rules are representable by modal formulas of the latter form. As a matter of fact, this modal representation has actually been used in many approaches to formalizing causation in action theories (see, e.g., (Geffner 1990; Turner 1999; Giordano, Martelli, and Schwind 2000; Zhang and Foo 2001)).

The nonmonotonic semantics of causal inference

Causal inference relations are just a special kind of causal theories, so they also possess a nonmonotonic semantics. Moreover, due to the logical properties of causal inference, the description of this nonmonotonic semantics can be simplified as follows.

To begin with, we extend causal rules to rules having arbitrary sets of propositions as premises: given a causal inference relation $\Rightarrow$ and an arbitrary set u of propositions, $u \Rightarrow A$ will be taken to hold if, for some finite $a \subseteq u$, $\bigwedge a \Rightarrow A$ belongs to $\Rightarrow$. $\mathcal{C}(u)$ will denote the set of propositions caused by u, that is $\mathcal{C}(u) = \{A \mid u \Rightarrow A\}$. Then a world α is an *exact world* of a causal inference relation if and only if

$$\alpha = \mathcal{C}(\alpha).$$

Given an arbitrary causal theory Δ, we will denote by $\Rightarrow_\Delta$ the least causal inference relation that includes Δ. It has been shown in (Bochman 2003) that Δ has the same nonmonotonic semantics as $\Rightarrow_\Delta$, which means that the rules of causal inference are adequate for causal reasoning with respect to the nonmonotonic semantics of causal theories. Moreover, it has been shown that causal inference relations constitute in this respect a maximal such logic (see (Bochman 2004) for details).

Dynamic Causal Inference

By a *dynamic causal inference relation* we will mean a set of dynamic causal rules of the form $A.B \Rightarrow C$ that satisfies the conditions described below.

The first group of postulates states that a set of dynamic causal rules with a fixed first premise (D) should satisfy the postulates of an 'ordinary' causal inference:

(Strengthening) If $A \vDash B$ and $D.B \Rightarrow C$, then $D.A \Rightarrow C$;

(Weakening) If $D.A \Rightarrow B$ and $B \vDash C$, then $D.A \Rightarrow C$;

(And) If $D.A \Rightarrow B$ and $D.A \Rightarrow C$, then $D.A \Rightarrow B \wedge C$;

(Or) If $D.A \Rightarrow C$ and $D.B \Rightarrow C$, then $D.A \vee B \Rightarrow C$;

(Cut) If $D.A \Rightarrow B$ and $D.A \wedge B \Rightarrow C$, then $D.A \Rightarrow C$;

(Truth) $\mathbf{t}.\mathbf{t} \Rightarrow \mathbf{t}$;

(Falsity) $\mathbf{t}.\mathbf{f} \Rightarrow \mathbf{f}$.

In view of the above postulates, dynamic causal rules $C.A \Rightarrow B$ can be seen as ordinary, binary causal rules $A \Rightarrow B$ that are conditioned by the preceding context C.

The next two postulates describe the logical properties of this preceding context in dynamic causal rules:

(Left-Str) If $A \vDash B$ and $B.D \Rightarrow C$, then $A.D \Rightarrow C$;

(Left-Or) If $A.D \Rightarrow C$ and $B.D \Rightarrow C$, then $A \vee B.D \Rightarrow C$.

The combined effect of the above pair of postulates is that the associated semantic interpretation of dynamic causal inference (described in the next section) will be again a kind of a possible world semantics, in which both the two premises and conclusion of a dynamic causal rule are evaluated with respect to worlds (complete states).

Finally, the last postulate is a formal expression of the requirement that any state that can be an input of some transition, is also an output state of at least one transition:

(Transition) If $\mathbf{t}.A \Rightarrow \mathbf{f}$, then $A.\mathbf{t} \Rightarrow \mathbf{f}$.

Recall that we have decided to identify static causal laws $A \Rightarrow B$ with dynamic causal laws of the form $\mathbf{t}.A \Rightarrow B$. Then the above postulate can be rewritten as

(Transition1) If $A \Rightarrow \mathbf{f}$, then $A.\mathbf{t} \Rightarrow \mathbf{f}$.

On this reformulation, the above postulate stipulates, in effect, that any input state of a consistent transition should be (statically) causally consistent. Combined with the other postulates, this will immediately imply that both the input and output state of a transition should be closed with respect to the valid static laws.

In the next section we will describe a possible-worlds semantics for the above logical formalism of dynamic inference.

A possible worlds semantics

A possible worlds semantics of dynamic causal relations can be obtained by generalizing an accessibility relation on possible worlds to ternary relations.

A *causal possible world model* of dynamic causal inference is a triple (W, R, V), where W is a set of possible worlds, R a ternary accessibility relation on W, and V a function assigning each world a propositional interpretation. The accessibility relation will be required to satisfy the following two conditions:

(Quasi-reflexivity) If $R\alpha\beta\gamma$, then $R\alpha\beta\beta$.

(Transition) If $R\alpha\beta\beta$, then $R\delta\alpha\alpha$, for some $\delta \in W$.

Definition 5. A rule $A.B \Rightarrow C$ is *valid* in a model (W, R, V) if, for any worlds α, β, γ such that $R\alpha\beta\gamma$, if A holds in α and B holds in β, then C holds in γ.

Given the above definition of validity, it is easy to verify the following

Lemma 5. *The set of dynamic causal rules valid in a causal possible world model forms a dynamic causal inference relation.*

Moreover, using a suitable construction of a canonical semantics for a dynamic causal inference relation, the following completeness result can be established:

Theorem 6. *A set of dynamic causal rules forms a dynamic causal inference relation if and only if it is determined by a causal possible world model.*

Proof. (A sketch) Due to the connection between dynamic causal rules and the original, atemporal causal rules, the proof is a relatively straightforward generalization of the corresponding completeness proof for causal inference relations, given in (Bochman 2004, Theorem 7.4). More precisely, given a dynamic causal relation $\Rightarrow$, we can construct the corresponding canonical model (W, R_c) by taking W to be the set of all maximal consistent sets of propositions, and defining R_c as follows[8]:

$$R_c\alpha\beta\gamma \;\equiv\; C(\alpha.\beta) \subseteq \beta \cap \gamma.$$

Notice that this definition directly implies quasi-reflexivity of R_c. Moreover, the use of the Transition postulate allows us to prove the transition property of R_c. Finally, it can be shown that $A.B \Rightarrow C$ holds for the source dynamic causal relation if and only if it is valid in (W, R_c). $\square$

Remark. One of the interesting consequences of the above semantic characterization of dynamic causal inference is that, similarly to a straightforward modal translation of ordinary causal rules as formulas of the form $A \to \Box B$, dynamic causal rules can be represented as formulas of *arrow logic* (see, e.g., (Venema 1997)). As a matter of fact, one of the principal motivations behind arrow logic has also consisted in providing an abstract description of dynamic (transition) models (cf. (van Benthem 1994)). Moreover, semantic interpretation of arrow logic is also based on a possible world semantics with a ternary accessibility relation, and it can be easily verified that, by the above semantic description, a dynamic causal rule $A.B \Rightarrow C$ turns out to be equivalent to a formula

$$A \circ B \to C$$

of arrow logic, where $\circ$ is a binary 'arrow conjunction' operator having the following semantic interpretation: $A \circ B$ holds in a world α if and only if there are worlds β, γ such that $R\beta\gamma\alpha$, A holds in β and B holds in γ.

[8] See the next section for a definition of $C(\alpha.\beta)$.

Correspondences

Recall that a dynamic causal theory is an arbitrary set of dynamic causal rules. For any dynamic causal theory Δ there exists a least dynamic causal inference relation that includes Δ. We will denote it by $\Rightarrow_\Delta$. Clearly, $\Rightarrow_\Delta$ is the set of all dynamic causal rules that can be derived from Δ using the postulates of dynamic causal inference.

As before, we will extend the notation of dynamic causal rules to sets of propositional formulas in premises: for sets u, v of propositional formulas, $u.v \Rightarrow A$ will be taken to hold if $\bigwedge a. \bigwedge b \Rightarrow A$, for some finite $a \subseteq u, b \subseteq v$. In addition, $\mathcal{C}(u.v)$ will denote the set of propositions $\{A \mid u.v \Rightarrow A\}$.

Due to the logical properties of a dynamic causal inference relation, the definition of an exact transition can now be simplified, namely a pair of worlds (α, β) will be an *exact transition* with respect to a dynamic causal inference relation if and only if

$$\beta = \mathcal{C}(\alpha.\beta).$$

Then the following key result of this study shows, in effect, that the logic of causal dynamic inference is adequate for reasoning with respect to the exact semantics of dynamic causal theories, since it preserves the latter.

Theorem 7. *The exact transition models of a dynamic causal theory Δ coincide with the exact transition models of $\Rightarrow_\Delta$.*

Proof sketch. It can be shown that if C_Δ is the provability operator of $\Rightarrow_\Delta$, then, for any worlds α, β, if $C_\Delta(\alpha.\beta)$ is a consistent set, then it coincides with $\mathrm{Th}(\Delta(\alpha.\beta))$. Consequently, $\alpha = C_\Delta(\alpha.\beta)$ iff $\alpha = \mathrm{Th}(\Delta(\alpha.\beta))$, and therefore exact transitions of Δ will coincide with exact transitions of $\Rightarrow_\Delta$. Hence the result. □

Moreover, it can be shown that the logic of causal dynamic inference constitutes a maximal logic that is adequate for reasoning with exact causal models.

Definition 6. Two dynamic causal theories Δ and Γ are *strongly equivalent* if, for any set Φ of causal rules, $\Delta \cup \Phi$ has the same exact transition models as $\Gamma \cup \Phi$.

Strongly equivalent theories are 'equivalent forever', that is, they are interchangeable in any larger causal theory without changing the nonmonotonic semantics. Consequently, strong equivalence can be seen as an equivalence with respect to the background monotonic logic of causal theories. And the next result shows that this logic is precisely the logic of dynamic causal inference.

Theorem 8. *Dynamic causal theories Δ and Γ are strongly equivalent if and only if $\Rightarrow_\Delta = \Rightarrow_\Gamma$.*

The above result states that dynamic causal theories are strongly equivalent if and only if each of them can be obtained from the other using the postulates of dynamic causal inference.

Summary and Perspectives

The primary objective of this study consisted in showing that causal reasoning in dynamic action domains can be given a direct and concise logical representation. Moreover, being combined with the wealth of representation capabilities of such a reasoning, demonstrated in (Giunchiglia et al. 2004), the results of this study strongly indicate that a theory of dynamic causal inference can be viewed as a self-subsistent logical theory that provides an adequate and comprehensive representation framework for reasoning in dynamic domains. The study creates also obvious incentives for broader questions about the role and scope of causation in commonsense reasoning, as well as in knowledge representation in AI.

Causation has always been one of the most discussed concepts in the philosophy of science. It is intimately related to practically all notions that are essential both for a commonsense and scientific view of the world, such as laws, counterfactuals, explanation and abduction. On the other hand, causation and related notions have shown to be extremely elusive and problematic concepts. Efforts of many philosophers and logicians in the past have been focused on a formal, logical explication of these notions, but the task has turned out to be surprisingly difficult. Furthermore, starting with David Hume, an influential line of philosophical thought has argued, in effect, that causation should be expelled from the language of Science and Logic.

In recent years, however, we witness a revival of interest in the concept of causation, accompanied with new, more practical, insights about its role in our reasoning. Most prominent in this respect is Pearl's theory of causal reasoning (Pearl 2000) and its applications in statistics, economics, cognitive and social sciences.

An important alternative source of the new understanding of causation and its role in our reasoning comes from Artificial Intelligence, especially from theories of action and change. In these theories causation is a working concept allowing us to single out intended models of commonsense action descriptions. They make especially vivid the fact that causal reasoning, that is, asking why and seeking explanations, is germane to our reasoning about the world. However, these theories have also made evident that, though causal reasoning includes an important logical part, it is *not reducible* to a plain logical derivation in some ingenious causal logic. Instead, causal reasoning should be viewed as an important case of general nonmonotonic (assumption-based) reasoning. Accordingly, the tools and formalisms of nonmonotonic reasoning should hopefully provide us with a more adequate understanding of the concept of causation itself.

References

Beall, J.; Brady, R.; Dunn, J.; Hazen, A.; Mares, E.; Meyer, R.; Priest, G.; Restall, G.; Ripley, D.; Slaney, J.; and Sylvan, R. 2012. On the ternary relation and conditionality. *Journal of Philosophical Logic* 41(3):595–612.

Bochman, A. 2003. A logic for causal reasoning. In *IJCAI-03, Proceedings of the Eighteenth International Joint Con-*

ference on Artificial Intelligence, Acapulco, Mexico, August 9-15, 2003, 141–146. Acapulco: Morgan Kaufmann.

Bochman, A. 2004. A causal approach to nonmonotonic reasoning. *Artificial Intelligence* 160:105–143.

Bochman, A. 2007. A causal theory of abduction. *Journal of Logic and Computation* 17:851–869.

Bochman, A. 2011. Logic in nonmonotonic reasoning. In Brewka, G.; Marek, V. W.; and Truszczynski, M., eds., *Nonmonotonic Reasoning. Essays Celebrating its 30th Anniversary*. College Publ. 25–61.

Bochman, A. 2013. Causation in flux. In *Commonsense-2013, 11th International Symposium on Logical Formalization of Commonsense Reasoning*.

Calvanese, D.; De Giacomo, G.; and Vardi, M. Y. 2002. Reasoning about actions and planning in LTL action theories. In *Proc. Eights Int. Conf. on Principles of Knowledge Representation and Reasoning (KR-02), Toulouse, France*, 593–602. Morgan Kaufmann.

De Giacomo, G., and Vardi, M. Y. 2013. Linear temporal logic and linear dynamic logic on finite traces. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*.

Geffner, H. 1990. Causal theories for nonmonotonic reasoning. In *Proceedings of National Conference on Artificial Intelligence (AAAI)*, 524–530. AAAI Press.

Gelfond, M., and Lifschitz, V. 1998. Action languages. *Electronic Transactions on Artificial Intelligence* 3:195–210.

Giordano, L.; Martelli, A.; and Schwind, C. 2000. Ramification and causality in a modal action logic. *J. Log. Comput.* 10(5):625–662.

Giunchiglia, E., and Lifschitz, V. 1995. Dependent fluents. In *Proceeedings IJCAI-1995*, 1964–1969.

Giunchiglia, E., and Lifschitz, V. 1998. An action language based on causal explanation: preliminary report. In *Proceedings of National Conference on Artificial Intelligence (AAAI)*, 623–630. AAAI Press.

Giunchiglia, E.; Lee, J.; Lifschitz, V.; McCain, N.; and Turner, H. 2004. Nonmonotonic causal theories. *Artificial Intelligence* 153:49–104.

Harel, D.; Kozen, D.; and Tiuryn, J. 2000. *Dynamic Logic*. MIT Press.

Lee, J.; Lifschitz, V.; and Yang, F. 2013. Action language BC: Preliminary report. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*.

Lin, F. 1995. Embracing causality in specifying the inderect effect of actions. In *Proc. Int. Joint Conf. on Artificial Intelligence, IJCAI-95*, 1985–1991. Montreal: Morgan Kaufmann.

Lin, F. 1996. Embracing causality in specifying the indeterminate effects of actions. In *Proceedings AAAI-96*, 670–676.

McCain, N., and Turner, H. 1997. Causal theories of action and change. In *Proceedings AAAI-97*, 460–465.

Myers, K. L., and Smith, D. E. 1988. The persistence of derived information. In *Proceedings AAAI-88*, 496–500.

Pearl, J. 2000. *Causality*. Cambridge UP.

Reiter, R. 1991. The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression. In Lifschitz, V., ed., *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*. Academic Press. 318–420.

Reiter, R. 2001. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamic Systems*. MIT Press.

Shanahan, M. P. 1997. *Solving the Frame Problem*. The MIT Press.

Turner, H. 1999. A logic of universal causation. *Artificial Intelligence* 113:87–123.

van Benthem, J. 1994. A note on dynamic arrow logic. In van Eijck, J., and Visser, A., eds., *Logic and Information Flow*. Cambridge, MA, USA: MIT Press. 15–29.

Venema, Y. 1997. A crash course in arrow logic. In Marx, M.; Pólos, L.; and Masuch, M., eds., *Arrow Logic and Multi-Modal Logic*. Stanford, CA, USA: Center for the Study of Language and Information. 3–34.

Zhang, D., and Foo, N. 2001. EPDL: A logic for causal reasoning. In *Proc Int. Joint Conf. on Artificial Intelligence, IJCAI-01*, 131–136. Seattle: Morgan Kaufmann.