Chronological Ignorance

Time, Nonmonotonicity, Necessity and Causal Theories

Yoav Shoham

Yale University
Computer Science Department

Abstract. Concerned with the problem of reasoning efficiently about change within a formal system, we identify the *initiation* problem. The solution to it which we offer, called the logic of *chronological ignorance*, combines temporal logic, nonmonotonic logic, and the modal logic of necessity. We identify a class of theories, called *causal theories*, which have elegant model-theoretic and complexity properties in the new logic.

1 Introduction: the prediction task

The work overviewed here falls into the class of attempts to formalize aspects of commonsense reasoning. The particular task considered is the *prediction* task. When we see someone pulling the trigger of a gun we brace ourselves, predicting that a loud noise will follow. I would like to be able to emulate this process on a computer. In other words, I am interested in being able to reason efficiently, naturally, and rigorously about the behavior of a system, given a description of it and of the relevant rules of "lawful change." This is related to, but distinct from, the work done in qualitative physics [1], since I am interested in a precise *logic*. The stress is on maintaining formal semantics throughout the process, so that the denotation of our symbols always remains clear.

This research was motivated by the need to reason about complex and continuous processes, such as billiard balls rolling and colliding with one another, or liquids heating until they boil. In [11] I indeed discuss such scenarios, but in this paper I will necessarily have to simplify the discussion. First, I will view time as being discrete and linear. Second, I will interpret propositions over time *points* rather than time *intervals*. Neither assumption is one I believe in (see, e.g., [12]). However, the essential concepts — *chronological ignorance* and *causal theories* — can be explained already in this constrained framework.

Consider the following very simple scenario, to which I will make reference throughout the paper. In it a gun is loaded at $t = 1$ and fired at $t = 5$. Furthermore, our knowledge of guns tells that if a loaded gun is fired at $t = i$ then there is a loud noise at $t = i+1$, provided no "weird" circumstances obtain: there is air to carry the sound, the gun has a firing pin, the bullets are made out of lead and not marshmallows, and this list can be continued arbitrarily. Are we justified in concluding that there will be a loud noise at time $t = 6$?

The answer is of course no, and there are two reasons for that. First, there is the question of whether the gun is loaded at time $t = 5$. It was loaded at $t = 1$, but how long did that last? We would like to say that it lasted until the firing, or more generally, that it lasted for "as long as possible" (that is, the interval of "being loaded" cannot be extended without violating some physical law and other facts that happen to be true). How do we capture in our logic the property of persisting "for as long as possible"? Second, even if we managed to show that the gun was loaded at time $t = 5$, we would still not be able to show that none of the "weird" circumstances hold at $t = 5$; that is not entailed by our statements.

I will term the first problem (that of assigning "inertia" to propositions) the *persistence problem*, and the second problem (that of excluding unusual circumstances) the *initiation problem*. These two problems are related to the infamous *frame problem* [5], but transcend the particular framework of the situation calculus [5]; they arise whenever one uses "local rules of change." I explore this issue further both in [10] and in the full version of this paper. Here I will make do with an informal description of the problem.

In this paper I outline a solution to the initiation problem. Intuitively, it is the problem of having to specify many mundane details – such as the gun having a firing pin, there being air, the bullets being made out of lead, and so on – in order to make a single prediction (in our case - that a noise will follow the shooting).

The basic solution is to allow nonmonotonic inferences. We conclude that a noise will follow the shooting, but retract that if we learn that the scenario takes place in a vacuum. Several nonmonotonic logics have appeared in the literature. I will discuss them briefly in the last section, but at this point let me just say that *none* of them have the right properties for our purposes. The rest of the paper gives the details of the appropriate nonmonotonic logic, the logic of *chronological ignorance*. Beside a solution to the particular problem, the paper offers two more general contributions. First, it offers a uniform and flexible way of constructing nonmonotonic logics (either classical or modal). Second, it suggests a semantical account of causation; in the full version of the paper, and in [11], I make that account explicit.

2 Chronological Ignorance

There are in principle two ways to resolve the initiation problem. The first is a syntactic one: we treat the assertions that "the gun was loaded and then fired" merely as *shorthand*, as an abbreviation of a much richer set of assertions, including the fact that in was not fired in the interim, that there is air, that the gun has a firing pin, and so on. This way of explaining away the initiation problem requires that we actually provide a translation rule, which expands an arbitrary abbreviated theory into the full-blown one.

I will take another route, the semantic one. The semantic solution of the initiation problem is an instruction to interpret the theory differently than it usually is. In other words, it provides a new meaning to the assertions.

2.1 Definitions: TETL, c.m.i. models

In order to discuss semantics, we first have to fix a language for representing temporal information. As was mentioned in the introduction, I will use a toy language in this paper.

Definition 1. The *Toy Temporal Logic (TTL)* is defined as follows. Assume a set of *primitive propositions* Φ, and a set of *time-point symbols* Θ. *Atomic formulas* are those of the form $TRUE(t,p)$, where $p \in \Phi$ and $t \in \Theta$. The set of all formulas is the boolean closure of the atomic ones, that is, their closure under '$\neg$' and '$\wedge$' (quantification over time points would be a simple addition, but I will not even allow *that* here). The semantics of formulas is obvious: the meaning of a primitive proposition is a set of time points, and for any interpretation $V = \langle V1, V2 \rangle$, $V \models TRUE(t,p)$ iff $V1(t) \in V2(p)$. The meaning of more complex formulas is defined inductively in the usual way. As was also mentioned in the introduction, we assume a fixed interpretation of time, namely that of the integers, and so we can use the syntactic t and the semantic $V1(t)$ interchangeably.

Given more space I would be able to motivate the following introduction of a modal operator. Since I don't, I'll have to refer the reader to the full paper, assuring him at this point that the transition is not thoughtless extravagance on my part, but rather a calculated and advantageous step.

Definition 2. *Toy Epistemic Temporal Logic (TETL)* is TTL augmented by the modal operator K. We intend the usual meaning for this operator, as it as been used in recent years: Kx is read 'x is known', and, assuming the now-standard Kripke semantics, will say that Kx is true in a world exactly if x is true in all worlds accessible from that particular world. We furthermore assume an S5 system, so that possible worlds form equivalence classes.[1] For more details on the modal logic of knowledge see, e.g., [2].

Definition 3. *Atomic knowledge sentences* are those of the form $K(TRUE(t,p))$ or of the form $K(\neg TRUE(t,p))$. We use $K(t,p)$ as abbreviation for $K(TRUE(t,p))$, and $K(t,\neg p)$ as abbreviation for $K(\neg TRUE(t,p))$.

We are now in a position to define the notion of chronological ignorance.

Definition 4. A (Kripke) structure M_1 is *chronologically more ignorant* than a structure M_2 if there exists a time t_0 such that

1. the two structures agree on all atomic knowledge sentences $K(t,\varphi)$ such that $t<t_0$ (wrt the global interpretation of time; I'll omit this comment in the future),
2. For any atomic knowledge sentence $x = K(t_0,\varphi)$, if $M_1 \models x$ then also $M_2 \models x$.
3. There exists an atomic knowledge sentence $x = K(t_0,\varphi)$ such that $M_2 \models x$ but $M_1 \not\models x$.

Definition 5. A structure M is a *chronologically maximally ignorant (c.m.i.)* model of a formula Φ if M

[1] In this paper I will interpret the modal operator epistemically. In fact, in this context I have an alternative and, I believe, better interpretation of the modality. Going into that, however, will be too lengthy, and I will reserve that discussion to the full paper.

is a model of Φ and if there is no other model of Φ that is chronologically more ignorant than M. Notice that chronological maximal ignorance is nonmonotonic; a c.m.i. model of $\Phi_1 \wedge \Phi_2$ need not be, and usually is not, a c.m.i. model of Φ_1.

2.2 The shooting scenario revisited

Armed with these definitions, let us reexamine the shooting scenario. First, we formulate the theory in TETL. There is more than one way this can be done; I choose the following axiom and axiom schemata for reasons that will become apparent soon.

1. *K(1,loaded)*
2. *K(5,fire)*
3. *K(i,loaded)* $\wedge$ *K(j,fire)* $\wedge$
$\bigwedge_{i<l<j}$ $\neg$*K(l,fire)* $\wedge$
$\neg$*K(j,vacuum)* $\wedge$
$\neg$*K(j,nofiringpin)* $\wedge$
$\neg$*K(j,marshmallow-bullets)* $\wedge$
$\ldots\neg K\ldots$other "weird" conditions
$\supset$
K(j+1,noise), for all $i<j$

Axioms 1 and 2 can be thought of the *boundary conditions* of the scenario. Axiom schema 3 represents "physics," in this case consisting of a single *causal rule.* It says that firing after loading caused a noise, unless certain conditions obtain which "disable" this particular rule.

What do c.m.i. models of this theory look like? There are many different such models, but they have one thing in common: they all satisfy the same atomic knowledge sentences. These are the sentences *K(1,loaded)*, *K(5,fire)* and *K(6,noise)* — exactly the ones we would have liked. In fact, this is no coincidence. In the next subsection I will identify a class of theories all of which have this property.

Notice a certain tradeoff that is taking place here. Consider, for example, the conjunct $\neg$*K(j,vacuum)* in the causal rule. We could replace this conjunct by a conjunct *K(j,$\neg$vacuum)*, but the result would be slightly different. In the theory as we have it above, we need not say anything about there being air in order to be able to infer that there will be a noise after the firing. On the other hand, if there is *no* air we had better state that fact explicitly in the initial conditions, otherwise we will erroneously conclude that there will be a loud noise. If we changed the formulation as we have just described, we would be in the exact opposite situation.

The principle underlying our logic may be called *the ostrich principle*, or the *what-you-don't-know-won't-hurt-you* principle. If *K(t,φ)* appears on the l.h.s. of a causal rule then we have in effect set the default of φ to be *false*, since if we say nothing about φ then *K(t,φ)* is false. Notice, however, that we have set the default to false only as far as *this* particular rule is concerned. On the other hand, if $\neg$*K(t,$\neg\varphi$)* appears on the l.h.s. of a disabling rule then we have in effect set the default of φ to be *true*, as far as this particular causal rule is concerned. Which alternative is better depends on what happens more often. If shooting scenarios rarely take place in a vacuum, as indeed is the case in everyday life, then we are better off sticking with the original theory; we will not need to mention the atmospheric conditions, except in those unusual occasions when things indeed take place in a vacuum.

3 Causal theories

In general, a theory might have many different c.m.i. models, or none at all. However, it was demonstrated that in at least one case, the shooting example, all c.m.i. models are essentially the same, and furthermore that it is exactly the model we intend for our theory. In this section I pin down the discussion further by giving general conditions under which we can expect chronological ignorance to be useful. Intuitively speaking, the reason the concept was useful in the shooting example is because events in that domain only had influence on the *future*: loading and firing after t_0 could not affect any noises *before* t_0. This property is common to all theories which we intuitively think of as *causal*; causes must precede their effects.

3.1 Defining causal theories

Definition 6. A *toy causal theory* is a collection of sentences in TETL, which can be divided into two subcollections (in the following, [$\neg$] means that the negation sign may or may not appear, and *p*, with or without a subscript, is a primitive proposition):

1. *"Boundary conditions:"* a collection of sentences of the form *K(t,[$\neg$]p)*.
2. *"Causal rules:"* a collection of sentences of the form $\Phi \wedge \Theta \supset K(t_i,[\neg]p)$, where Φ is a nonempty conjunction of sentences $K(t_j,[\neg]p_j)$ such that $t_j<t_i$, and Θ is a (possibly empty) conjunction of sentences $\neg K(t_j,[\neg]p_j)$ such that $t_j<t_i$.

These toy causal theories embody a few simplifications beyond the ones made in the underlying temporal logic. First, since we are interpreting formulas over time points and not time intervals, "causes" (i.e., the conjuncts of Φ) cannot overlap in time with their "effect". Furthermore, we even prohibit simultaneity of cause and effect (since we demand $t_j < t_i$ in the causal rules). This is clearly too limiting, and general causal theories do not have these limitations. They are discussed in [11].

3.2 The simplicity of causal theories

When discussing the simple model-theoretic properties of the shooting scenario, I claimed that those properties were no coincidence. I will now make the statement more concrete.

Theorem 1. The atomic knowledge sentences satisfied by a c.m.i. model of a toy causal theory consist of the boundary conditions and a subset of the r.h.s.'s of the causal rules.

Corollary 2. All c.m.i. models of a finite toy causal theory satisfy a finite number of atomic knowledge sentences.

Theorem 3. (The 'unique' c.m.i. model property.) All the c.m.i. models of a toy causal theory satisfy the same atomic knowledge sentences.

The simple c.m.i. model-theory of causal theories makes them also very easy to reason about. The result given here refers to toy causal theories, but extends easily to general causal theories. The general argument is that in order to enumerate the atomic knowledge sentences, all you need to do is "sweep forward in time." Since you'd like to know as little as possible for as long as you can, as you move forward in time you add only that knowledge that is absolutely necessary in light of your knowledge and ignorance so far. The particular form of causal theories guarantees that future knowledge and ignorance will not affect past knowledge and ignorance. As a more specific example, we have the following

Theorem 4. The (unique and finite) set of basic knowledge sentences satisfied by any c.m.i. model of a finite toy causal theory can be computed in time $O(n \log n)$, where n is the size of the causal theory.

4 Related work

There is much to say about the relation of this work to previous work in computer science and philosophy, but again, given the space limitations, I will only be able to briefly mention a few points.

1. There has been considerable amount of work on nonmonotonic logics in AI. The best known systems are McCarthy's circumscription [6], Reiter's default logic [9], McDermott and Doyle NML I [7] and McDermott's NML II [8]. The reader may ask why we cannot simply adopt one of those and be done with it. The short answer is that there has been much wishful thinking in this regard; in reality, almost none of those claiming that a particluar nonmonotonic system captured the inferences they desire verified that it in fact did. This discrepancy between hopes and reality was recently made very clear when S. Hanks and D. McDermott tried to apply the first three systems to a simple problem in temporal reasoning, and *none* of them turned out to have the right properties [3] (see related paper in this volume). It is a direct corollary of the Hanks and McDermott experiment that none of the above systems can be used to achieve the effect of chronological minimality. The underlying problem is the crude criterion of what constitutes a "minimal model." Taking McCarthy's circumscription as an example, we have a "set inclusion" criterion: when you circumscribe a FO formula φ, you select models in which the extension of φ is not a superset of its extension in any other model. This turns out to be too crude a criterion of minimality for our purposes.

2. In all FO-based nonmonotonic logic one must specify explicitly what it is that is being minimized. For example, in circumscription one must supply the predicate to be circumscribed, and for the more exotic versions (e.g., parameterized) even more needs to be specified. Notice that in the logic of chronological ignorance, the object of minimization is defined once and for all: we (chronologically) minimize knowledge.

3. Recently, V. Lifschitz proposed a new form of circumscription called *pointwise circumscription* [4]. In that new formulation the minimality criterion is made much more flexible, and can be used to chronologically minimize the extension of a particular predicate (or set of predicates). It cannot, however, be used to emulate the notion of chronological ignorance, since one must still specify explicitly what it is that is being minimized. There is a way of combining pointwise circumscription with the "abnormality" predicate which bears an interesting relation to our logic, but I will refer the reader to the full paper for details.

4. The discussion in this paper has been entirely model theoretic. One of the elegant features of circumscription, in either McCarthy's original formulations or Lifschitz' recent ones, is that it comes along with

a circumscription *axiom*, a second-order axiom that when added to the theory achieves the effect of limiting the models to the "minimal" ones (in the relevant sense of minimality). The question is, though, since we understand the model theory anyway, what do the various (extremely ingenuous) circumscription axioms add to our understanding. It would seem that those would be worthwhile only if there were a way to use them to generate automatic inferences, or if they generalized any results on (say) chronological minimization to a larger class of nonmonotonic logics. I am skeptical of the first possibility: the only uses of circumscription to date have been manual and incredibly simplistic. It seems that at this point the burden of the proof that the circumscription *axiom* is of any use is on its vendors. The second possibility, however, that the circumscription axiom (and I have in mind Lifschitz' new version) would suggest results that transcend the particular criterion of chronological minimality, looks more promising.

5. The inadequacy for our purposes of the set-inclusion minimality criterion extends to recent logics of minimal knowledge, such as those discussed by Moore, Konolige, Halpern and Moses, and Vardi.

5. Causation has been the subject of much discussion in philosophy. I think it is fair to say that there has not yet been a satisfactory account of the concept, which plays such a prominent role in our everyday thinking. I am now in a position to give a precise semantic account of causation, which appears not to suffer from shortcomings of previous accounts. Since I do not have the space to give the details, I will reserve those to a fuller version of this paper. Here I will only claim that the expressiveness of causal theories on the one hand, and their simplicity on the other, explain why causal reasoning is so pervasive in everyday life.

5 Summary

The main messages of this paper have been the following.

1. One problem that arises when one tries to reason about change both efficiently and rigorously is the *initiation problem.* The logic of *chronological ignorance* offers one solution.

2. Causal theories have nice model-theoretic and complexity properties, which is one explanation why the concept of causation plays such a prominent role in everyday thinking.

3. Nonmonotonic logics are constructed semantically, by deciding on the minimality criterion for models. Here one such criterion was discussed; in [11] I discuss another.

Acknowledgements. This work was supervised by Drew McDermott. Beside Drew, I have benefitted from discussions with more people at Yale and outside it than I could possibly list, such as Joe Halpern, Steve Hanks and Vladimir Lifschitz.

Bibliography

[1] D. G. Bobrow (Ed.), *Special Volume on Qualitative Reasoning and Physical Systems,* Artificial Intelligence, 24/1-3 December (1984).

[2] J. Y. Halpern and Y. Moses, A guide to the Modal Logics of Knowledge and Belief: Preliminary Draft, *Proc. IJCAI-85,* IJCAI, 1985, pp. 480–490.

[3] S. Hanks and D. V. McDermott, *Temporal Reasoning and Default Logics,* Technical Report YALEU/C: #430, Yale University, October 1985.

[4] V. Lifschitz, Pointwise Circumscription, Manuscript 1986.

[5] J. M. McCarthy and P. J. Hayes, Some Philosophical Problems From the Standpoint of Artificial Intelligence, *Readings in Artificial Intelligence,* Tioga Publishing Co., Palo Alto, CA, 1981, pages 431–450.

[6] J. M. McCarthy, Circumscription - A Form of Non Monotonic Reasoning, *Readings in Artificial Intelligence,* Tioga Publishing Co., Palo Alto, CA, 1981, pages 466–472.

[7] D. V. McDermott and J. Doyle, *Nonmonotonic Logic I,* Artificial Intelligence, 13 (1980), pp. 41–72.

[8] D. V. McDermott, *Nonmonotonic Logic II: Nonmonotonic Modal Theories,* JACM, 29/1 (1982), pp. 33–57.

[9] R. Reiter, *A Logic for Default Reasoning,* Artificial Intelligence, 13 (1980), pp. 81–132.

[10] Y. Shoham, *Ten Requirements for a Theory of Change,* New Generation Computing, 3/4 (1985).

[11] Y. Shoham, *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence,* Ph.D. Thesis, Yale University, Computer Science Department, 1986.

[12] Y. Shoham, Reified Temporal Logics: Semantical and Ontological Considerations, *Proc. 7th ECAI,* Brighton, U.K., July 1986.