Chronological Minimisation and Explanation

John Bell
Applied Logic Group
Department of Computer Science
Queen Mary and Westfield College
University of London
London E1 4NS
jb@dcs.qmw.ac.uk

Abstract

This paper discusses chronological minimisation and the problem of representing common sense explanatory causal reasoning. The background of chronological minimisation is briefly surveyed. Then prioritised chronological minimisation is introduced. Finally, it is shown that this can be used as a basis for representing explanation.

1 Introduction

In a seminal paper, McCarthy [1986] suggested a way of representing common sense causal reasoning in the framework of the Situation Calculus. He introduced an abnormality predicate in order to indicate when changes occur, and suggested using Predicate Circumscription to minimise this predicate, with the effect that changes are minimised. By considering only those models is which abnormalities (changes) were minimised, McCarthy was giving a *pragmatics* which was intended to capture the intended meaning of the theories in question. The pragmatics captured an important property of common sense causal reasoning: *common sense causal reasoning is governed by defeasible causal rules.* Common sense causal reasoning typically takes place in *partial* contexts, contexts which are incompletely specified; because obvious facts are taken for granted, or because it is impractical or impossible to state all of the facts. We overcome this by making appropriate assumptions about what is normal or typical given the context, and by using defeasible causal rules which license conclusions on the basis of such assumptions. Formal reasoning of this kind thus involves generating appropriate additional assumptions, and then restricting changes to those that follow from the rules given these assumptions. Thus "miracles", facts or events which do not follow from the intended interpretation of the rules, are ruled out.

However, Hanks and McDermott [1986] showed by means of their infamous Yale Shooting Problem that minimising changes (abnormalities) is insufficient for representing simple cases of predictive causal reasoning. Consequently they suggested that changes (abnormalities) should be minimised chronologically. This refinement on model selection captures a further property of predictive common sense causal reasoning : predictive common sense causal reasoning involves reasoning forwards in time from causes to their likely effects : *predictive common sense reasoning is temporally forwards directed.* Lifschitz [1986] showed how Circumscription can be extended to satisfy the additional condition, and Shoham outlined, and later expounded, an alternative approach in which knowledge is minimised chronologically [1986, 1988]. Shortly afterwards Haugh [1987] and Lifschitz [1987] suggested chronologically minimising causes. Various problems with the various proposals began to emerge, for example they encountered difficulty with ramifications. However the apparently fatal blow had already been struck by Kautz [1986] when he showed, by means of his Stolen Car Problem, that approaches based on chronological minimisation cannot represent simple cases of explanatory causal reasoning. A further example, the Stanford Murder Mystery, was provided by Baker [1989]. It is now commonly accepted that chronological minimisation, in whatever form it is realised, cannot be used to represent explanatory causal reasoning because of what might be called the *postponement problem*; as chronological minimisation has the effect of delaying changes it cannot, it is commonly thought, be used to represent scenarios, such as those suggested by Kautz and Baker, in which this effect is counter intuitive. In time most of the proponents of chronological minimisation abandoned it on this score, and turned to other approaches. The few exceptions, for example Sandewall [1991, 1995], went on to propose increasingly more complex forms of chronological minimisation, and the idea lost its simplicity and much of its intuitive appeal. For example, Sandewall suggests that changes should be filtered from observations, and then un-occluded changes should be minimised chronologically. While this procedure is shown to work at the formal level, the intuitions behind it are not clear. What in our everyday causal reasoning corresponds to this technical machinery?

However it seems that one of the approaches based on (simple) chronological minimisation, that proposed by Shoham, has been abandoned prematurely. Shoham's Causal Theories have been attacked for a variety of reasons, both conceptual, e.g. [Galton 1991], and technical, e.g. [Sandewall 1991, 1995]. A major problem is the lack of expressiveness of Causal Theories which arises as they are exrpessed in a propositional language, and as a result of the syntactic constraints which Shoham imposes in order to obtain the unique models property. However, Shoham's approach remains attractive as it is based on temporal logic, which, thanks to the work of generations of logicians (see e.g. [Rescher and Urquhart 1971, van Benthem 1991]) has a clear, general, ontology, and because it provides a clear and elegant model-theoretic realisation of chronological minimisation. Consequently, Prediction Theories (PTs) [Bell 1997] represent an attempt to rejuvenate and extend Shoham's approach. The syntax of PTs is much more general and there is no attempt to obtain a unique models result. The aim is thus to show that chronological minimisation can be used in a much broader setting. These generalisations are outlined in the next section. In particular, (simple) chronological minimisation is refined further to *prioritised* chronological minimisation. This captures a further property of predictive common sense causal reasoning, namely that *if an event can succeed, then, ceteris paribus, it will succeed.* If the normal preconditions of an event are true when the event occurs, then other things being equal, the event will succeed. In [Bell 1997] it is argued that prioritised chronological minimisation when used with PTs, provide a general means of representing predictive causal reasoning.

But what of explanatory common sense causal reasoning and the postponement problem? In section 3 it is argued, using Kautz and Baker's examples, that the postponement problem can be overcome using PTs and prioritised chronological minimisation. However, another more serious problem, the *generation problem*, emerges. Prioritised chronological minimisation as used with PTs cannot generate the additional information contained in explanations. But this should not be surprising as prediction and explanation are different forms of reasoning. While prediction is concerned with producing reasonable conclusions about the future on the basis of the information currently available, with reasoning forwards in time from causes to their likely effects, explanation (at least in this context) is concerned with producing reasonable conclusions about the past on the basis of the information currently available, with reasoning backwards in time from effects to their likely causes. Thus while prediction is best viewed as a form of *inductive* reasoning, explanation is best viewed as a form of *abductive* reasoning. However, as section 3 shows, prioritised chronological minimisation can be used as the basis of an abductive definition of explanatory common sense causal reasoning.

2 Prioritised Chronological Minimisation

Prioritised chronological minimisation is realised in the framework of Prediction Theories (PTs). PTs are expressed in the Temporal Calculus ($\mathcal{TC}$). $\mathcal{TC}$ is so called because it combines features of temporal logic with features analogous to those of the Situation Calculus. $\mathcal{TC}$ is a many sorted, higher order, temporal, nonmonotonic extension of Kleene's strong three-valued language [Kleene 1952]. The three-valued basis provides a natural partial setting in which prioritised chronological minimisation, the pragmatics of PTs, can be defined. In particular it provides an answer to Galton's objections to the use of the knowledge operator in Shoham's framework. The three-valued basis also facilitates implementation [Bell 1995, 1996], [Bell, Hodges and White 1997]. Finally the temporal logic basis of $\mathcal{TC}$ also makes it readily extendible [Bell 1997]. In order to make this paper self contained, a concise acount of the relevant features of $\mathcal{TC}$ is now given.

Kleene introduced the third truth value *undefined* in order to be able to describe computations which may not, in principle, terminate. The language can also be given an epistemic, resource bounded, interpretation. Typically we have incomplete, or partial, information and our reasoning resources are limited. Consequently we cannot, as the classical ideal suggests, determine the truth or falsity of every sentence. Rather, we concentrate our efforts on attempting to establish the truth or falsity of those sentences that we consider relevant and ignore, treat as undefined, the truth values of the rest. Accordingly, a sentence can be true (established by the reasoner(s) as being true), false (established by the reasoner(s) as being false) or undefined (ignored by the reasoner(s) as irrelevant, or unestablishable within the reasoner(s) resource limitations). Appropriately, the truth conditions of the logical constants yield a Boolean truth value wherever possible. For example, $\neg\phi$ is true if ϕ is false, false of ϕ is true, and undefined otherwise. Similarly, $\phi \wedge \psi$ is true if both ϕ and ψ are both true, false if either is false, and undefined otherwise. Disjunction, $\vee$, and material implication, $\supset$, can be defined in the usual way. The language is then readily extended to the first-order case. Atomic sentences are interpreted using partial functions. Thus an atomic sentence, a *Kleene atom*, $r(u_1, \ldots, u_n)$ may be true, false or undefined in a model. A universally quantified sentence $\forall v\phi$ is then true if every instance of ϕ is true, false if one instance is false, and is undefined otherwise. The existential quantifier, $\exists$, is then defined in the usual way.

As it stands, Kleene's logic is appropriate for representing monotonic reasoning with partial information. In order to be able to represent nonmonotonic reasoning, an *undefined* operator, ?, is added. Informally $?\phi$ states that the truth value of ϕ is undefined; that is, that neither the truth nor the falsity of ϕ is established. Thus $?\phi$ should be true if ϕ is undefined, and false otherwise. This operator can be used to define *weak affirmation*, *weak negation*, and *defined* operators, making it possible to express meta-level distinctions such as "ϕ is not false" and "ϕ is not true", and consequently nonmonotonic inference rules, in the object language.

$$\begin{array}{rcll} \circ\phi & \stackrel{\mathrm{def}}{=} & ?\phi \vee \phi & \text{“}\phi\textit{ is not false”} \\ \bullet\phi & \stackrel{\mathrm{def}}{=} & ?\phi \vee \neg\phi & \text{“}\phi\textit{ is not true”} \\ !\phi & \stackrel{\mathrm{def}}{=} & \neg ?\phi & \text{“}\phi\textit{ is not undefined”} \end{array}$$

Similar extensions to Kleene's language, starting with the weak negation operator, are discussed in [Doherty 1996]. A formula is said to be *hypothetical* if the undefined operator occurs in it, otherwise the formula is said to be *categorical*.

A drawback with Kleene's semantics is that the defined conditional operator, $\supset$, does not provide a satisfactory account of conditionality. In particular, if ϕ is true and ψ is undefined, then $\phi \supset \psi$ is undefined, rather than false.

Moreover, in keeping with the resource-bounded interpretation, it is desirable to have a conditional which is true if the antecedent is not true, regardless of the truth value of the consequent. A conditional of this kind can be thought of as expressing a constraint which must be met if the antecedent is true, but which can otherwise be ignored. It can be defined as follows:

$$\phi \rightarrow \psi \quad \stackrel{\text{def}}{=} \quad \neg\bullet\phi \supset \neg\bullet\psi$$

where the sentence $\neg\bullet\phi$ makes the classical claim that ϕ is true; that is, it is true if ϕ is, and is false otherwise. The new conditional is itself classically valued. It is false if the antecedent is true and the consequent is not, and true otherwise.

Kleene added his own classical operator. The sentence $\phi \equiv \psi$ states that ψ and ψ are equivalent; that is, that they always have the same truth value. The sentence is true if this is the case and is false otherwise. This operator can be defined as follows:

$$\phi \equiv \psi \quad \stackrel{\text{def}}{=} \quad (\neg\bullet\phi \wedge \neg\bullet\psi) \vee (\neg\circ\phi \wedge \neg\circ\psi) \vee (?\phi \wedge ?\psi)$$

where the sentence $\neg\circ\phi$ makes the classical claim that ϕ is false.

In order to represent time, an explicit temporal index is added to each atomic formula of the underlying language, to each Kleene atom; thus time is not reified, rather it is introduced as a separate sort and a predicative approach is adopted. For example the ordinary atom $r(u_1, \ldots, u_n)(t)$ states that the Kleene literal $r(u_1, \ldots, u_n)$ is true at the time point t. The temporal precedence operator, $<$, allows comparison of time points, thus $t < t'$ states that the time point t temporally preceeds the time point t'. The semantics for time are as in standard temporal logic. Models now include a temporal frame $(\mathcal{T}, \mathcal{R}_\mathcal{T})$ where $\mathcal{R}_\mathcal{T} \subseteq \mathcal{T} \times \mathcal{T}$. Intuitively $\mathcal{T}$ is the set of time points and $(\mathcal{T}, \mathcal{R}_\mathcal{T})$ represents the flow of time. For convenience in this paper we will assume that time is discrete and linear. The truth value of Kleene literals may now vary from time point to time point.

As recommended by Galton and as practised in the Situation Calculus, the difference in aspect between facts and events is signified. Accordingly, (primitive) events are introduced as a third sort.

In order to gain comparable expressivity with the Situation Calculus, *Kleene literals*, Kleene atoms and their negations, are added as a fourth sort. This higher-order quantification over literals is convenient, and can easily be replaced by treating Kleene literals as particulars and using a $Holds$ predicate; as is done with fluents in the Situation Calculus. The final extension is a set of higher-order predicates of the form $hr(e_1, \ldots, e_n, \lambda_1, \ldots, \lambda_m)(t)$; where $n + m > 0$, the e_i are event terms, and the λ_j are Kleene literals. For example, a sentence of the form $Occ(e)(t)$ states that a token of an event of type e occurs at t. Two further higher-order predicates, *Aff* and *Qual*, will be introduced in the sequel.

The relevant formal semantics of $\mathcal{TC}$ are now outlined. A *model* of $\mathcal{TC}$ is a structure $M = (\mathcal{D}, \mathcal{E}, \mathcal{T}, \mathcal{R}_\mathcal{T}, \mathcal{F}, \mathcal{R}, \mathcal{HR}, \mathcal{V})$, where: $\mathcal{D}$ is the set of domain objects, $\mathcal{E}$ is the set of primitive events, $(\mathcal{T}, \mathcal{R}_\mathcal{T})$ is a (discrete, linear) temporal frame, $\mathcal{F}$ interprets the functions for each sort, $\mathcal{R}$ is a partial interpretation function for relations, $\mathcal{HR}$ is a partial interpretation function for higher-order relations, and $\mathcal{V}$ interprets the terms of each sort.

The interpretation of terms is straightforward. If is M model with term interpretation function $\mathcal{V}$, and g is a variable assignment for M, then the term evaluation function $\mathcal{V}_g$ interprets ordinary, event, and temporal terms in the usual way, while literal variables are assigned literals, and literals are "interpreted" as themselves.

The truth and falsity conditions of sentences are then defined by means of the intermediary notions of satisfaction and violation.

Definition 2.1 *Let $M = (\mathcal{D}, \mathcal{E}, \mathcal{T}, \mathcal{R}_\mathcal{T}, \mathcal{F}, \mathcal{R}, \mathcal{HR}, \mathcal{V})$ be a $\mathcal{TC}$-model, g be a variable assignment for M, and ϕ be a $\mathcal{TC}$-formula. Then g* satisfies *ϕ in M (written $M, g \models \phi$) or* violates *ϕ in M (written $M, g \models\!\!\!\mid \phi$) as follows:*

$M, g \models t < t'$	*iff*	$(\mathcal{V}_g(t), \mathcal{V}_g(t')) \in \mathcal{R}_\mathcal{T}$
$M, g \dashv t < t'$	*iff*	$(\mathcal{V}_g(t), \mathcal{V}_g(t')) \notin \mathcal{R}_\mathcal{T}$
$M, g \models r(u_1, \ldots, u_n)(t)$	*iff*	$\mathcal{R}(r, \mathcal{V}_g(t))(\mathcal{V}_g(u_1), \ldots, \mathcal{V}_g(u_n)) = \mathit{true}$
$M, g \dashv r(u_1, \ldots, u_n)(t)$	*iff*	$\mathcal{R}(r, \mathcal{V}_g(t))(\mathcal{V}_g(u_1), \ldots, \mathcal{V}_g(u_n)) = \mathit{false}$
$M, g \models$ $hr(e_1, \ldots, e_n, \lambda_1, \ldots, \lambda_m)(t)$	*iff*	$\mathcal{HR}(hr, \mathcal{V}_g(t))(\mathcal{V}_g(e_1), \ldots, \mathcal{V}_g(e_n),$ $\mathcal{V}_g(\lambda_1), \ldots, \mathcal{V}_g(\lambda_m)) = \mathit{true}$
$M, g \dashv$ $hr(e_1, \ldots, e_n, \lambda_1, \ldots, \lambda_m)(t)$	*iff*	$\mathcal{HR}(hr, \mathcal{V}_g(t))(\mathcal{V}_g(e_1), \ldots, \mathcal{V}_g(e_n),$ $\mathcal{V}_g(\lambda_1), \ldots, \mathcal{V}_g(\lambda_m)) = false$
$M, g \models \neg\psi$	*iff*	$M, g \dashv \psi$
$M, g \dashv \neg\psi$	*iff*	$M, g \models \psi$
$M, g \models ?\psi$	*iff*	*neither* $M, g \models \psi$ *nor* $M, g \dashv \psi$
$M, g \dashv ?\psi$	*iff*	*either* $M, g \models \psi$ *or* $M, g \dashv \psi$
$M, g \models \psi \wedge \chi$	*iff*	$M, g \models \psi$ *and* $M, g \models \chi$
$M, g \dashv \psi \wedge \chi$	*iff*	$M, g \dashv \psi$ *or* $M, g \dashv \chi$
$M, g \models \forall v \psi$	*iff*	$M, g' \models \psi$ *for all* g' *such that* $g \overset{v}{\approx} g'$
$M, g \dashv \forall v \psi$	*iff*	$M, g' \dashv \psi$ *for some* g' *such that* $g \overset{v}{\approx} g'$

A formula ϕ is true *in a model M (written $M \models \phi$) if $M, g \models \phi$ for all variable assignments g. A formula ϕ is* false *in M (written $M \dashv \phi$) if $M, g \dashv \phi$ for all variable assignments g. We write $M \not\models \phi$ if it is not the case that $M \models \phi$, and $M \not\dashv \phi$ if it is not the case that $M \dashv \phi$. M is said to be a* model *of a set of sentences Θ (written $M \models \Theta$), if $M \models \phi$ for all $\phi \in \Theta$.*

Prediction Theories (PTs) are expressed in $\mathcal{CA}$. However, their definition is part syntactic and part semantic. For present purposes, the formal details are not essential, so an informal outline is given. PTs may contain *boundary conditions*, *prediction rules*, and *domain rules*.

Boundary conditions are categorical unconditional sentences which state facts and the occurrence of events; for example (4) and (8) below.

Domain rules are categorical temporally undirected conditionals or biconditionals. For any instance of a domain rule, the evaluation of the antecedent and the evaluation of the consequent starts at the same time point and is completed at the same time point. Domain rules are used for a variety of purposes, including the definition of domain constraints and ramifications; for example (9) and (19) below.

Prediction rules are temporally forwards-directed conditionals whose antecedents contain hypothetical formulas. For any instance of a prediction rule, the evaluation of the antecedent can be completed at a time point which temporally precedes the earliest time point at which the evaluation of consequent can be completed. Prediction rules are used to represent inertia and change (the direct effects of events).

In order to represent inertia, the *Aff* predicate is introduced. Intuitively, for Kleene literal λ and time point t, $\mathit{Aff}(\lambda)(t)$ states that the truth value of λ is affected at t; that is, that there is reason to doubt that the truth value of λ persists beyond t. A general *inertia rule* can then be stated:

$$\forall \lambda, t(\lambda(t) \wedge \bullet \mathit{Aff}(\lambda)(t) \rightarrow \lambda(t+1)) \tag{1}$$

Thus, if λ is true at t and it is not true that λ is affected at t, then λ remains true at $t+1$. In the examples that follow we adopt the usual convention that free variables are implicitly universally quantified. Thus (1) is abbreviated to:

$$\lambda(t) \wedge \bullet \mathit{Aff}(\lambda)(t) \rightarrow \lambda(t+1) \tag{2}$$

The intention is to use this rule from left to right only in order to reason forwards in time from instances of its antecedent to instances of its consequent. Typically also we want to be able to infer the second conjunct of each instance nonmonotonically; whenever it is consistent to do so. Thus, on the intended interpretation of the inertia rule, Kleene literals, and consequently the formulas composed of them, persist by default.

In order to represent change, the *Qual* predicate is introduced. Intuitively, $\mathit{Qual}(e)(t)$ states that event e is qualified at time t; that is, that there is reason to doubt that e will succeed if e occurs at t. Then a general *change rule* can be stated as follows (again omitting the quantification):

$$\mathit{Pre}(e)(t) \wedge \mathit{Occ}(e)(t) \wedge \bullet \mathit{Qual}(e)(t) \rightarrow \mathit{Post}(e)(t+1) \tag{3}$$

Thus if event e occurs at t and the normal preconditions of e, $Pre(e)(t)$, are true at t and e is not qualified at t, then infer that the normal postconditions of e, $Post(e)(t)$, are true at $t+1$; that is, that e succeeds at t. Once again, the intention is to use this rule from left to right only in order to reason forwards in time from instances of its antecedent to instances of its consequent. Typically also we want to be able to infer the second conjunct of each instance nonmonotonically; whenever it is consistent to do so. Thus, on the intended interpretation of the change rule, events succeed by default if their normal preconditions are true.

In order to ensure that PTs are interpreted as intended, weneed to define their pragmatics. The definition is based on the partiality of $\mathcal{TC}$. The relevant models are among those in which *information* (what is defined/established) is minimised chronologically. The effect is that prediction rules are interpreted from left to right only, and the appropriate nonmonotonic inferences are generated. However, simple chronological minimisation is insufficient as it does not resolve possible conflicts between (instances of) the inertia rule and the change rule. When there is such a conflict, it seems that the (instance of the) prediction rule should take precedence. This corresponds to the intuition that, other things being equal, if an event occurs and it can succeed (that is, if its necessary preconditions are true), then it will succeed. In order to capture this principle, simple chronological minimisation is refined to *prioritised chronological minimisation*. The minimisation is still chronological. However in order to give preference to potential change, qualifications (reasons why an event may not succeed as expected) are minimised in preference to affectations (reasons why a fact should not persist) at each time point. The following abbreviations are used in the definition. Let M be a $\mathcal{TC}$ model with term interpretation function $\mathcal{V}$. Then for any time point t:

$$\begin{aligned}
M_{\mathcal{R}}/t &= \{\lambda(t') \mid \mathcal{V}(t') \le t \text{ and } M \models !\lambda(t')\},\\
M_{\mathcal{A}}/t &= \{Aff(\lambda)(t') \mid \mathcal{V}(t') \le t \textit{ and } M \models !Aff(\lambda)(t')\},\\
M_{\mathcal{O}}/t &= \{Occ(e)(t') \mid \mathcal{V}(t') \le t \textit{ and } , M \models !Occ(e)(t')\},\\
M_{\mathcal{Q}}/t &= \{Qual(e)(t') \mid \mathcal{V}(t') \le t \textit{ and } M \models !Qual(e)(t')\}.
\end{aligned}$$

Thus $M_{\mathcal{R}}/t$ is the set of all Kleene literals which are defined (which are either true or false) in M up to t, and similarly $M_{\mathcal{A}}/t$, $M_{\mathcal{O}}/t$, and $M_{\mathcal{Q}}/t$ are respectively the sets of all *Aff*, *Occ* and *Qual* literals which are defined in M up to t.

Definition 2.2 *Let M and M' be $\mathcal{TC}$-models which differ at most on the interpretation of Kleene literals and the Occ, Aff and Qual predicates. Then M is* prioritised chronologically less defined *than M' (written $M \prec_{pcld} M'$) if there exists a time point t such that:*

- $M_{\mathcal{R}}/t \subset M'_{\mathcal{R}}/t$ *and* $M_{\mathcal{O}}/t \subseteq M'_{\mathcal{O}}/t$, *or*
- $M_{\mathcal{R}}/t \subseteq M'_{\mathcal{R}}/t$ *and* $M_{\mathcal{O}}/t \subset M'_{\mathcal{O}}/t$, *or*

- $M_{\mathcal{R}}/t \subseteq M'_{\mathcal{R}}/t$, $M_{\mathcal{O}}/t \subseteq M'_{\mathcal{O}}/t$ *and* $M_{\mathcal{Q}}/t \subset M'_{\mathcal{Q}}/t$, *or*
- $M_{\mathcal{R}}/t \subseteq M'_{\mathcal{R}}/t$, $M_{\mathcal{O}}/t \subseteq M'_{\mathcal{O}}/t$, $M_{\mathcal{Q}}/t \subseteq M'_{\mathcal{Q}}/t$ *and* $M_{\mathcal{A}}/t \subset M'_{\mathcal{A}}/t$.

Thus $M \prec_{pcld} M'$ if M and M' are comparable and there is a time point t such that M and M' agree for any earlier time point, and:

- at least one more Kleene or *Occ* literal is established in M' at t, or
- M and M' agree on Kleene and *Occ* literals, and at least one more *Qual* literal is established in M' at t, or
- M and M' agree on all Kleene, *Occ*, and *Qual* literals, and at least one more *Aff* literal is established in M' at t.

Definition 2.3 *A model M is a* prioritised chronologically least defined *model (a* p.c.l.d. model*) of ϕ if $M \models \phi$ and there is no other model M' such that $M' \models \phi$ and $M' \prec_{pcld} M$. Likewise, M is a p.c.l.d. model of a set of sentences Θ if $M \models \Theta$ and there is no other model M' such that $M' \models \Theta$ and $M' \prec_{pcld} M$.*

Definition 2.4 *A PT Θ* predicts *a sentence ϕ, written $\Theta \approx_{pcld} \phi$, if ϕ is true in all p.c.l.d. models of Θ.*

The following example motivates the refinement of simple chronological minimisation.

Example 2.5 (The Dodgy Pickup.) A block on the table is picked up. We want to be able to conclude that it is not on the table as a result. Let Θ be the PT $\{(2),(4),(5)\}$; where:

$$OnTable(B)(1) \land Occ(Pickup(B))(1) \tag{4}$$

$$OnTable(b)(t) \land Occ(Pickup(b))(t) \land \bullet Qual(Pickup(b))(t) \rightarrow \neg OnTable(b)(t+1) \tag{5}$$

$\Theta \approx_{pcld} \neg OnTable(B)(2)$ as required. By virtue of the definition, $?Qual(Pickup(B))(1)$ is true in all p.c.l.d. models of Θ. The intended models correspond to those in which the defeasible change rule is applied first, blocking the application of the inertia rule. By contast, simple chronological minimisation would also yield unintended models corresponding to those in which the inertia rule is applied first, blocking the application of the defeasible change rule. $\square$

[Bell 1997] shows that, in addition to representing qualifications and inertia, this realisation of prioritised chronological minimisation in PTs can be used to represent ramifications, nondeterminism, and simultaneous interacting events, and, with the full definition of the $\mathcal{TC}$ language, events with durations, continuous change, and complex events. Moreover the language is readily extended to incorporate partial modal operators ([Thijse 1992, Bell 1995]) allowing for the representation of mental attitudes. Despite all of this however, there remains the question as to the representation of explanation.

3 Explanation

It is commonly believed that a major drawback of approaches based on chronological minimisation is that they cannot represent explanatory causal reasoning because of the postponement problem. However, it seems clear from the standard examples that the problem lies elsewhere.

Example 3.1 (The Stolen Car Problem [Kautz 1986].) You park your car in the lot at time 1. You discover that it is gone at time 5. You conclude that it could have been stolen at any time in between. It is widely believed that that approaches based on chronological minimisation cannot represent this problem as they lead to the conclusion that the car was stolen as late as possible; that is, at time 4.

In order to see the problem, let Θ be the PT $\{(2),(6),(7),(8),(10),(11)\}$; where:

$$InLot(1) \tag{6}$$

$$\neg InLot(5) \tag{7}$$

$$\exists t(2 \leq t \leq 4 \wedge Occ(Steal)(t)) \tag{8}$$

$$2 \leq t \leq 4 \rightarrow\ !Occ(Steal)(t) \tag{9}$$

$$InLot(t) \wedge Occ(Steal)(t) \rightarrow \neg InLot(t+1) \tag{10}$$

$$Occ(Steal)(t) \rightarrow InLot(t) \tag{11}$$

Then, indeed, $\Theta \mathrel{|\!\!\approx}_{pcld} Occ(Steal)(4)$; as the p.c.l.d. models of Θ are those models in which the *Steal* event occurs as late as possible.

However, given the present formulation, the problem is not one of explanation but of the misinterpretation of the existential quantifier in the boundary condition (8). The pragmatic interpretation of (8), the meaning of (8) in all p.c.l.d. models of Θ, is $Occ(Steal)(4)$. Clearly however, the semantic interpretation of (8), the meaning of (8) in all models, is required. The semantic interpretation is ensured by adding rule (9), which states that the occurrence or non-occurrence of the *Steal* event is defined (is relevant) at each of the time points in question. Thus there are three classes of p.c.l.d. models of $\Theta' = \Theta \cup \{(9)\}$, those in which $Occ(Steal)(2)$ is true, those in which $Occ(Steal)(3)$ is true, and those in which $Occ(Steal)(4)$ is true. Consequently $\Theta' \mathrel{|\!\!\approx}_{pcld} Occ(Steal)(2) \vee Occ(Steal)(3) \vee Occ(Steal)(4)$, while $\Theta' \not\mathrel{|\!\!\approx}_{pcld} Occ(Steal)(2)$, $\Theta' \not\mathrel{|\!\!\approx}_{pcld} Occ(Steal)(3)$ and $\Theta' \not\mathrel{|\!\!\approx}_{pcld} Occ(Steal)(4)$ as required.

The real problem with this formulation is that it provides only a partial solution, as (8) explicitly states that a *Steal* event occurs. □

Example 3.2 (The Stanford Murder Mystery [Baker 1991].) In this variant of the Yale Shooting Problem, Fred is alive at time 1. As before, if the gun is loaded when it is fired, then Fred is dead as a result. We are also told that a *Shoot* event occurs at time 1, that a *Wait* event occurs at time 2 (that is, that nothing changes at time 2), and that Fred is dead at time 3. When and why did Fred die? Intuitively we think that the gun must have been loaded at time 1 and that, as a result of the shooting, Fred is dead at time 2.

Baker claims that chronological minimisation gives the wrong answer. To see this, let Θ be the PT $\{(2), (12), (13), (14), (15)\}$; where:

$$Alive(1) \wedge Occ(Shoot)(1) \tag{12}$$

$$Occ(Wait)(2) \tag{13}$$

$$\neg Alive(3) \tag{14}$$

$$Loaded(t) \wedge Occ(Shoot)(t) \rightarrow \neg Alive(t+1) \tag{15}$$

$$\lambda(t) \wedge Occ(Wait)(t) \rightarrow \lambda(t+1) \tag{16}$$

Then $\Theta \mid\!\approx_{pcld} Alive(2)$; (prioritised) chronological minimisation results in the unintuitive conclusion that Fred is still alive at time 2, and so dies mysteriously as a result of the $Wait$ event. Even worse, $\Theta \mid\!\approx_{pcld} \bullet Loaded(1)$; (prioritised) chronological minimisation leads to the conclusion that the gun was not loaded at time 1.

In defence of chronological minimisation, it seems fair to argue that the $Wait$ event is not correctly represented in Θ. In Baker's Situation Calculus framework a $Wait$ event leaves the values of all fluents unchanged. In order to repesent this correctly in the present framework, Rule (16) is added. Consequently, $\Theta \cup \{(16)\} \mid\!\approx_{pcld} \neg Alive(2)$; Fred dies as a result of the $Shoot$ event at time 1.

However, this is only a partial solution as $\Theta \cup \{(16)\} \not\mid\!\approx_{pcld} Loaded(1)$; it is not possible to generate the additional fact that the gun was loaded at time 1. $\square$

As these examples show, the real problem with the present approach is that it cannot be used to *generate* the additional information commonly contained in explanations. This is hardly surprising as PTs were not designed to do so. Prediction is concerned with drawing reasonable conclusions about the future on the basis of the information currently available, with reasoning from causes to their likely effects. By contrast, explanation (at least in this context) is concerned with drawing reasonable conclusions about the past on the basis of the information currently available, with reasoning from effects to their likely causes. Thus while prediction is a form of *inductive* reasoning, abduction is best viewed as a form of *abductive* reasoning.

Standardly, e.g. [Brewka, Dix and Konolige 1997], given a theory Θ such that $\Theta \not\models \phi$, abduction is presented as the process of finding appropriate minimal sets Θ_i such that $\Theta \cup \Theta_i \models \phi$. The Θ_i represent the various alternative explanations for ϕ given Θ. In order that the explanations are relevant, the sentences in each Θ_i are usually required to be drawn from a restricted set of *abducible* sentences. Explanations are also required to be *non-trivial*, in that $\Theta_i \not\models \phi$. This can be called the *credulous* definition of abduction. It is typically used as the basis of proof-theoretic approaches; for example, those based on Logic

Programming [Kakas et.al. 1992]. It is possible to give a *sceptical* definition of abduction, according to which abduction is the process of finding a single appropriate minimal set Θ' such that $\Theta \cup \Theta' \models \phi$. Intuitively, Θ' represents the "disjunction" of all of the explanations for ϕ given Θ. For example, where there are a finite number of finite Θ_i, Θ' often contains a single sentence in disjunctive normal form, where each disjunct is a conjunction of the sentences in one of the Θ_i. The sceptical definition is used in model-theoretic approaches; for example, those based on Circumscription [Shanahan 1993, 1997].

Fortunately the framework that has been developed to represent prediction as a form of induction can be readily extended to represent explanation as a form of abduction. When Θ is a PT, the abductive task can be sharpened to that of finding appropriate minimal Θ_i such that $\Theta \cup \Theta_i \approx_{pcld} \phi$ or, in the sceptical case, an appropriate minimal Θ' such that $\Theta \cup \Theta' \approx_{pcld} \phi$. The abducibles are thus drawn from the instantiations of the antecedents of the rules in Θ. Using $\approx_{pcld}$ rather than $\models$ as the consequence relation should result in fewer, more relevant, explanations; as the obvious pragmatic inferences given $\Theta \cup \Theta'$ are already included in the $\approx_{pcld}$ relation. The present approach also suggests two ways in which the credulous explanations, the Θ_i, can be *generated*, one proof theoretic in character, the other purely model-theoretic.

From a "proof theoretic" perspective, the Θ_i can be regarded as sets of additional assumptions which are required in order to predict ϕ given Θ. These could be generated chronologically by starting at the initial time point and then reasoning forwards time point by time point, assuming only what is necessary in order to predict the conclusion. The additional assumptions used in each such "proof" then constitute one of the Θ_i. The assumptions could also be generated antichronologically by starting from the latest time point in the conclusion and then reasoning backwards time point by time point again assuming only what is necessary in order to predict the conclusion.

The purely model-theoretic perspective uses the $\prec_{pcld}$ ordering on models. If $\Theta \approx_{pcld} \phi$, the relevant models are simply the p.c.l.d. models of Θ; no further explanation is given. If not, the relevant models of $\Theta \cup \{\phi\}$ occur in the ordering above the p.c.l.d. models of Θ. As one proceeds up any chain of models above a p.c.l.d. model, M, of Θ one encounters models in which additional literals are introduced antichronologically. These models thus represent the various ways in which the course of events represented in M could have differed as a result of changes introduced at progressively earlier moments. The relevant models will thus be among those in which the changes introduced in order that ϕ is true are minimal. However, several restrictions are required in order to eliminate models which do not represent genuine alternative histories, models which do not represent natural alternative courses of events.

Clearly, the relevant models must be models in which the principle of inertia is respected; changes should not occur without reason. In practice this means that a literal should not become affected unless this is necessary on grounds of consistency. Similarly, the relevant models must be models in which the principle of change is respected; events with the potential to succeed should not fail without reason. In practice this means that an event should not be qualified unless this is necessary on grounds of consistency.

Definition 3.3 *Let M be a model of a PT Θ, and Λ be a literal of the form $\mathit{Aff}(\lambda)(t)$ or $\mathit{Qual}(e)(t)$. Then Λ is* grounded by Θ in M *if $M \models \Lambda$ and there is no model M' of Θ such that M/t and M/t' differ only in that $M' \models \bullet\Lambda$. M is a* gounded model of Θ *if every literal Λ of the form $\mathit{Aff}(\lambda)(t)$ or $\mathit{Qual}(e)(t)$ which is true in M is grounded by Θ in M.*

For example, any p.c.l.d. model of a PT is a grounded model of it. Note that affectations are not grounded in preference to qualifications. If, given Θ, ϕ can be explained by an event succeeding or by it failing, then both possibilities are equally plausible as credulous explanations. Examples arise when there is a possible over-determination of causes. Suppose that in circuit diagnosis the task is to explain why the output of an OR gate is on. Given the initial inputs to the circuit, both inputs to the gate should be on. However, the circuit has not been tested, so it is possible that only one of the inputs to the gate is on. It seems that three explanations are plausible: either or both of the inputs to the gate are on. Consequently we should allow the possibility that one of the inputs is not on because a failure has occurred in a gate leading to it.

If M is a grounded model of a PT Θ, then Θ also grounds any $Kleene$ or Occ literals which have to be true in M given Θ.

Definition 3.4 *Let M be a grounded model of a PT Θ. A literal Λ of the form $\lambda(t)$ or $\mathit{Occ}(e)(t)$ is* grounded by Θ in M *if $M \models \Lambda$ and there is no grounded model M' of Θ such that $M'/t-1 = M/t-1$ and $M' \models \bullet\Lambda$.*

Clearly if M is a p.c.l.d. model of a PT Θ, then Θ grounds all of the literals which are true in M.

Given Θ and the task of explaining ϕ the relevant models should be minimal grounded extensions of p.c.l.d. models of Θ to models of ϕ. The relevant models of $\Theta \cup \{\phi\}$ should thus be among the grounded models of Θ which are also models of ϕ. The explanation for ϕ in such a model, M, is the set of literals which are not grounded by Θ in M.

Definition 3.5 *Let $\Theta \cup \{\phi\}$ be a PT. Then M is a grounded model of $\langle\Theta, \phi\rangle$ if M is a grounded model of Θ and a model of ϕ. Let M be a grounded model of $\langle\Theta, \phi\rangle$. The* grounding set for M *given $\langle\Theta, \phi\rangle$ is the set of all literals which are true in M and which are not grounded by Θ.*

However explanations should also, wherever possible, be non-miraculous; wherever possible, explanations should not contain constituents which are themselves explicable. Wittgenstein once remarked that explanations have to stop somewhere, however in the present case it is reasonable to require that they continue for as long as possible. Thus the relevant models of $\Theta \cup \{\phi\}$ should also be *strongly grounded*:

Definition 3.6 *Let $\Theta \cup \{\phi\}$ be a PT and M_i be a grounded model of $\langle \Theta, \phi \rangle$ with grounding set Θ_i. Then a literal $\Lambda \in \Theta_i$ is* explainable *if there is a grounded model M_j of $\langle \Theta, \phi \rangle$ with grounding set Θ_j and which is such that $M_i \prec_{pcld} M_j$, $M_j \models \Lambda$, and $\Lambda \notin \Theta_j$. M is said to be a* strongly grounded *model of $\langle \Theta, \phi \rangle$ if no literal $\Lambda \in \Theta_i$ is explainable.*

Moreover, explanations should be non-trivial:

Definition 3.7 *Let $\Theta \cup \{\phi\}$ be a PT and M_i be a strongly grounded model of $\langle \Theta, \phi \rangle$ with grounding set Θ_i. Then M_i is a* non-trivial *model of $\langle \Theta, \phi \rangle$ if $\Theta_i \not\models \phi$.*

Finally, explanations should be parsimonious, or non-redundant. The relevant models should thus be the minimal extensions of the p.c.l.d. models of Θ to strongly grounded, non-trivial, models of $\Theta \cup \{\phi\}$.

Definition 3.8 *Let $\Theta \cup \{\phi\}$ be a PT and let M be a p.c.l.d. model of Θ. Then M_i is an* extended *p.c.l.d. model of Θ given ϕ (an e.p.c.l.d. model of $\Theta \cup \{\phi\}$) if $M \preceq_{pcld} M_i$, M_i is a strongly grounded non-trivial model of $\langle \Theta, \phi \rangle$ with grounding set Θ_i, and there is no strongly grounded non-trivial model M_j of $\langle \Theta, \phi \rangle$ with grounding set Θ_j such that $M \prec_{pcld} M_j \prec_{pcld} M_i$, and $\Theta_j \subseteq \Theta_i$.*

The credulous explanations for ϕ given Θ, the Θ_i, are thus the grounding sets corresponding to the e.p.c.l.d. models of $\Theta \cup \{\phi\}$. The sceptical form of abduction can now be defined as follows:

Definition 3.9 *Let Θ be a PT and ϕ be a sentence such that $\Theta \not\approx_{pcld} \phi$. Then the set of sentences Θ'* explain ϕ given Θ *if $\Theta \cup \{\phi\} \approx_{epcld} \Theta'$, $\Theta' \not\models \phi$, $\Theta \cup \Theta' \approx_{pcld} \phi$, and Θ' is minimal in that there is no $\psi \in \Theta'$ such that $\Theta \cup (\Theta' - \{\psi\}) \approx_{pcld} \psi$.*

The first example motivates the groundedness condition.

Example 3.10 (The Stanford Murder Mystery continued.) In Example 3.2 the missing information that the gun was loaded at time 1 can be generated abductively.

Let Θ be the PT $\{(2), (12), (13), (15), (16)\}$. Then $\{Loaded(1)\}$ explains $\neg Alive(3)$ given Θ. Let M be a p.c.l.d. model of Θ. Then M can be represented as the following set of literals:

$$M = \{Alive(1), Occ(Shoot)(1), Alive(2), Occ(Wait)(2), Alive(3), \ldots\}.$$

Then any e.p.c.l.d. model M_i of $\Theta \cup \{\neg Alive(3)\}$ can be represented as follows:

$$M_i = \{Alive(1), Occ(Shoot)(1), Loaded(1), Aff(Alive)(1), \neg Alive(2), \ldots\}.$$

There is a model M_j of $\Theta \cup \{\neg Alive(3)\}$ such that $M \prec_{pcld} M_j \prec_{pcld} M_i$:

$$M_j = \{Alive(1), Occ(Shoot)(1), Aff(Alive)(1), \neg Alive(2), \ldots\}.$$

However, M_j is not a grounded model of Θ; in M_j there is no reason why Fred dies. Consequently $\Theta \cup \{\neg Alive(3)\} \not\approx_{epcld} \{Loaded(1)\}$. Clearly, the remaining conditions in the definition of explanation are satisfied. □

The following example involves both explanation and a nondeterministic event, and illustrates the groundedness condition.

Example 3.11 (The Magic Bus [Kartha 1994].) A commuter travels to work by bus. He can get there by taking a red bus or a yellow bus. Before boarding he needs to buy a ticket, which he does at time 1. At time 2 he then boards the first bus that comes along. We are told that as a result he is on the red bus at time 3. The problem is to show that buying the ticket does not magically transport him onto the red bus, as happens in an unintended model on Baker's approach.

Let Θ be the PT $\{(2), (17), (18), (19), (20), (21)\}$; where:

$$\neg Ticket(1) \wedge Occ(Buy)(1) \wedge Occ(Board)(2) \tag{17}$$

$$\neg(On(Red)(t) \wedge On(Yellow)(t)) \tag{18}$$

$$\neg Ticket(t) \rightarrow \neg On(Red)(t) \wedge \neg On(Yellow)(t) \tag{19}$$

$$\neg Ticket(t) \wedge Occ(Buy)(t) \rightarrow Ticket(t+1) \tag{20}$$

$$\begin{aligned} &Ticket(t) \wedge \neg On(Red)(t) \wedge \neg On(Yellow)(t) \wedge Occ(Board)(t) \rightarrow \\ &On(Red)(t+1) \vee On(Yellow)(t+1) \end{aligned} \tag{21}$$

Then $\Theta \cup \{On(Red)(3)\} \approx_{pcld} \neg On(Red)(2) \wedge \neg On(Yellow)(2)$ as required. Let M be a p.c.l.d. model of $\Theta \cup \{On(Red)(3)\}$. By (17) and (19), $M \models \neg On(Red)(1) \wedge \neg On(Yellow)(1)$. So, by inertia, $M \models \neg On(Red)(2) \wedge \neg On(Yellow)(2)$.

The problem can be extended by supposing that we are not told which bus the commuter is on at time 3, and then asking for an explanation as to why he is on the red bus at time 4. Naturally we should infer that he must have boarded the red bus at time 2 and was thus on the red bus at time 3.

Clearly $\Theta \cup \{On(Red)(4)\} \approx_{epcld} On(Red)(3)$ as required. There are two classes of p.c.l.d. models of Θ, those in which $On(Red)(3)$ is true and those in which $On(Yellow)(3)$ is true. The models in the former class are all e.p.c.l.d. models, while none of the models in the latter class can be extended to e.p.c.l.d. models of $On(Red)(4)$; in particular, those models in which $On(Red)(2)$ is added are not grounded by Θ. □

The next example motivates the strong groundedness condition.

Example 3.12 (Variant of the Stanford Murder Mystery.) Suppose, in a variant of Example 3.2, that we are not told anything about time 2, and instead are told that the gun can become loaded as a result of a *Load* event. It seems that there are two possibilities, either the gun was already loaded at time 1 and so the *Shoot* event at time 1 was lethal, alternatively, a *Load* event could have occurred at time 1 followed by a lethal *Shoot* event at time 2. Let Θ be the PT $\{(2), (12), (15), (22)\}$; where:

$$Occ(Load)(t) \rightarrow Loaded(t+1) \tag{22}$$

Then $\{Loaded(1) \vee (Occ(Load)(1) \wedge Occ(Shoot)(2))\}$ explains $\neg Alive(3)$ given Θ. Let M be a p.c.l.d. model of Θ. Then M can be represented as the following set of literals:

$$M = \{Alive(1), Occ(Shoot)(1), Alive(2), Alive(3), \ldots\}.$$

Progressing up the p.c.l.d. ordering one finds the following grounded models of $\langle \Theta, \neg Alive(3) \rangle$; where $M \prec_{pcld} M'$ if $M \subset M'$.

$$\begin{aligned}
M_1 &= \{Alive(1), Occ(Shoot)(1), Alive(2), Loaded(2), Occ(Shoot)(2),\\
&\quad \neg Alive(3) \ldots\}\\
M_2 &= \{Alive(1), Occ(Shoot)(1), Loaded(1), \neg Alive(2), \ldots\}\\
M_3 &= \{Alive(1), Occ(Shoot)(1), Occ(Load)(1), Alive(2), Loaded(2),\\
&\quad Occ(Shoot)(2), \neg Alive(3), \ldots\}
\end{aligned}$$

Then M_2 and M_3 are e.p.c.l.d. models with respective grounding sets $\Theta_2 = \{Loaded(1)\}$ and $\Theta_3 = \{Occ(Load)(1), Occ(Shoot)(2)\}$. While M_1 is not strongly grounded as $Loaded(2)$ can be explained by Θ_3. Thus the "explanation" that the *Shoot* event at time 2 was lethal because, miraculously, $Loaded(2)$ is true while $Loaded(1)$ is not, is excluded. □

The next example shows that the postponement problem has been overcome.

Example 3.13 (The Stolen Car Problem reconsidered.) In Example 3.1 the explanation that the car must have been stolen between time 2 and time 4 can be generated abductively.

Let $\Theta = \{(2), (6), (10), (11)\}$. Then, $\Theta \cup \{\neg InLot(5)\} \approx_{epcld} (8)$. Let M be a p.c.l.d. model of Θ. Then M can be represented as follows:

$$M = \{InLot(1), InLot(2), \ldots, InLot(5), \ldots\}$$

Progressing up a chain in the p.c.l.d. ordering one successively finds the following e.p.c.l.d. models of $\Theta \cup \{\neg InLot(5)\}$:

$$\begin{aligned}
M_1 &= \{InLot(1), \ldots, InLot(4), Occ(Steal)(4), Aff(InLot)(4), \neg InLot(5), \ldots\}\\
M_2 &= \{InLot(1), \ldots, InLot(3), Occ(Steal)(3), Aff(InLot)(3), \neg InLot(4), \ldots\}\\
M_3 &= \{InLot(1), InLot(2), Occ(Steal)(2), Aff(InLot)(2), \neg InLot(3), \ldots\}
\end{aligned}$$

Consequently $\Theta \cup \{\neg InLot(5)\} \approx_{epcld} (8)$. □

The final example illustrates generation on the basis of partiality, strong groundedness, non-postponement, and the minimality of explanations.

Example 3.14 (The Loaded Gun Mystery (Suggested by Graham White).) We know that a gun is loaded at time 3. We also know that if a *Load* event occurs at time t then the gun is loaded at $t+1$. We are asked to explain how the gun became loaded.

Let $\Theta = \{(2), (22), (23)\}$; where:

$$Loaded(t) \vee Occ(Load)(t) \rightarrow t \geq 1 \tag{23}$$

Then $\Theta \cup \{\neg Loaded(3)\} \approx_{epcld} Loaded(1) \vee Occ(Load)(1) \vee Occ(Load)(2)$. Let M be a p.c.l.d. model of Θ. Then $M = \emptyset$. Progressing up the ordering from M one finds the following grounded models of $\langle \Theta, Loaded(3) \rangle$; where $M \prec_{pcld} M'$ if $M \subset M'$.

$$\begin{aligned}
&M_1 = \{Loaded(3)\}\\
&M_2 = \{Occ(Load)(2), Loaded(3)\}\\
&M_3 = \{Loaded(2), Loaded(3)\}\\
&M_4 = \{Occ(Load)(1), Loaded(2), Loaded(3)\}\\
&M_5 = \{Occ(Load)(1), Loaded(2), Occ(Load)(2), Loaded(3)\}\\
&M_6 = \{Loaded(1), Loaded(2), Loaded(3)\}\\
&\ldots
\end{aligned}$$

Then M_1 and M_3 are eliminated as they not strongly grounded, and M_5 is eliminated as its grouding set is contained in the grounding set for M_4. □

Acknowledgements

Thanks to Rob Miller, Murray Shanahan and Graham White for helpful discussions and comments. This work forms part of the Dynamo project and is supported by the United Kingdom Engineering and Physical Sciences Research Council under grant number GR/K19266.

References

[1] A. Baker (1989) A Simple Solution to the Yale Shooting Problem, Proceedings KR'89, pp. 11-20.

[2] A. Baker (1991) Nonmonotonic Reasoning in the Framework of the Situation Calculus. *Artificial Intelligence* 49 pp. 5-23.

[3] J. Bell (1995) Pragmatic Reasoning; a Model-Based Theory. In M. Masuch and L. Polos (eds.), *Applied Logic, How, What and Why?* A selection of papers from the Applied Logic Conference, Amsterdam, December 1992. Kluwer Academic Publishers, Amsterdam, pp. 1-28.

[4] J. Bell (1996) A Model-Based Approach to Predictive Causal Reasoning. In [Doherty 1996], pp. 169-195.

[5] J.Bell (1997) Prediction Theories. Working paper.

[6] J. Bell, W. Hodges and G. White (1997) The Model-Building Implementation of Prediction Theories. Working paper.

[7] G. Brewka, J. Dix, and K. Konolige (1997) *Nonmonotonic Reasoning, An Overview*, CSLI Lecture Notes No. 73, CSLI Publications, Stanford.

[8] P. Doherty (ed.) (1996) *Partiality, Modality, and Non-Monotonicity.* FoLLI Studies in Logic Language and Information series, CSLI Publications, Stanford.

[9] T. Eiter, G. Gottlob and Y. Gurevich (1993) Curb Your Theory! A Circumscriptive Approach for Inclusive Interpretation of Disjunctive Information. Proceedings of IJCAI'93, pp. 634-639.

[10] A. Galton (1991) A Critique of Yoav Shoham's Theory of Causal Reasoning. AAAI'91, pp. 355-359.

[11] S. Hanks and D. McDermott (1986) Default Reasoning, Nonmonotonic Logics, and the Frame Problem. Proceedings of AAAI'86, pp. 328-333.

[12] B. Haugh (1987) Simple Causal Minimisations for Temporal Persistence and Projection. Proceedings of AAAI'87, pp. 218-223.

[13] A.C. Kakas, R.A. Kowalski and F. Toni (1992) Abductive Logic Programming. *Journal of Logic and Computation* Vol 2, No. 6, pp. 719-770.

[14] H.A. Kautz (1986) The Logic of Persistence. AAAI'86, pp. 401-405.

[15] G.N. Kartha (1994) Two counterexamples related to Baker's approach to the frame problem. *Artificial Intelligence* 69, pp. 379-391.

[16] S.C. Kleene (1952) *Introduction to Metamathematics.* North-Holland, Amsterdam.

[17] V. Lifschitz (1994) Circumscription. In: *Handbook of Artificial Intelligence and Logic Programming.* D.M. Gabbay, C.J. Hogger and J.A. Robinson (eds.). Oxford University Press.

[18] J. McCarthy (1986) Applications of Circumscription to Formalising Commonsense Knowledge. *Artificial Intelligence* 28, pp. 98-116.

[19] N. Rescher and A. Urquhart (1971) *Temporal Logic*. Springer, Berlin.

[20] E. Sandewall (1991) *Features and Fluents*. Research Report LiTH-IDA-R-91-29, Dept. of Computer Science, Univ. of Linköping, Linköping, Sweden.

[21] E. Sandewall (1995) *Features and Fluents, The Representation of Knowledge About Dynamical Systems, Volume 1.* Oxford University Press, Oxford.

[22] M. Shanahan (1993) Explanation in the Situation Calculus. Proceedings of IJCAI-93, pp. 160- 165.

[23] M. Shanahan (1997) *Solving the Frame Problem*, M.I.T. Press, Cambridge, Mass.

[24] Y. Shoham (1988) *Reasoning About Change* M.I.T. Press, Cambridge, Mass.

[25] E. Thijse (1992) *Partial Logic and Knowledge Representation*. PhD Thesis, University of Tilburg. Eburon Publishers, Delft.

[26] J. van Benthem (1991) *The Logic of Time*, 2nd edition. D. Reidel, Dordrecht.