Journal of Artificial Intelligence Research 30 (2007) 361-412 Submitted 07/07; published 11/07

Natural Events

John Bell JB@DCS.QMUL.AC.UK
Department of Computer Science,
Queen Mary, University of London,
London E1 4NS, UK

Abstract

This paper develops an inductive theory of predictive common sense reasoning. The theory provides the basis for an integrated solution to the three traditional problems of reasoning about change; the frame, qualification, and ramification problems. The theory is also capable of representing non-deterministic events, and it provides a means for stating defeasible preferences over the outcomes of conflicting simultaneous events.

1. Introduction

A great deal has been written on the logical representation of common sense reasoning about change since the publication of McCarthy and Hayes's (1969) seminal paper, and many theories have been proposed; see, for example, the monographs by Sandewall (1994), Shanahan (1997), and Reiter (2001).

Most theories treat events[1] deductively, along the lines of the representation of actions used in the planner STRIPS (Fikes & Nilsson, 1971). Each event type is defined by its preconditions and effects. For example, in the blocks world, the preconditions for unstacking block x from block y are that x is on y, x is clear (no block is on top of it), and the robot hand is empty. The effects are that the hand is holding x, y is clear, and each of the preconditions is false. Change is then a matter of deduction. If a particular event (a token of an event type) occurs and its particular preconditions hold, then its particular effects are deduced, and so necessarily follow, from it. Events of this kind will be called *deductive events* and the view that natural events can be represented deductively will be called *Deductionism*.

Deductive event types can be thought of as invariable regularities (or uniformities) of sequence. Viewed in this way the STRIPS representation of events can be seen to be descended from those considered by Hume and Mill in their discussions of causation. Hume (1739, Bk I, Pt III) suggests that we inductively acquire knowledge of regularities of succession of the form: A-type events are followed by B-type events. We then consider that A-type events cause B-type events, because whenever we see an A-type event we expect that it will be followed by a B-type event. Mill (1898, Bk III, Ch 5) complicates this picture by considering *assemblages* of conditions. A single assemblage might consist of an A-type event together with certain conditions which must be present (positive conditions) and certain conditions which must be absent (negative conditions). For example, an assemblage concerning the lighting of matches might include the striking of the match, the presence of oxygen, and the absence of dampness in the match head.

1. Events are assumed to include the physical actions of agents; whether intentional or unintentional.

Mill (1898) thought that it was possible, at least in principle, to define assemblages which are detailed enough to ensure their effects: "For every event there is some combination of objects or events, some given concurrence of circumstances, positive and negative, the occurrence of which is always followed by that phenomenon" (p. 214).

However, Hume (1777, pp. 36-38) had already argued against this possibility. It is always possible that a regularity, no matter how long it has continued in the past, will not continue in the future. Consequently no set of sentences which report what has been observed ever logically implies anything about what has not been observed. As Goodman (1954, p. 59) puts it, what has happened imposes no logical restrictions on what will happen. So, if Deductionism is to be plausible, it is necessary to assume that Nature is uniform; that the future will resemble the past, that past regularities will continue. Now, clearly, Uniformity cannot be justified by appealing to experience, and it is difficult to see how else it can be justified; see, Goodman's discussion (pp. 61-62). Without such a justification, Deductionism should be regarded as being suspect in theory.

Deductionism is also suspect in practice, as it is impossible in practice to define preconditions which, together with the occurrence of the event, are sufficient to ensure that its effects will follow. For example, Russell (1913, p. 7) considers the problem of conflicting events: "I put my penny in the slot, but before I can draw out my ticket there is an earthquake which upsets the machine and my calculations. In order to be sure of the expected effect, we must know that there is nothing in the environment to interfere with it. But this means that the supposed cause is not, by itself, adequate to insure the effect". Russell also observes that we cannot usefully solve the problem by complicating preconditions because: "as soon as we include the environment, the probability of repetition is diminished, until, at last, when the whole environment is included, the probability of repetition becomes almost *nil*" (pp. 7-8).

The problem of specifying preconditions which are always sufficient also arises in Mackie's account of *causal regularities* (invariable regularities of sequence). For example, "at least part of the answer [to the question of what caused a particular fire] is that there is a set of conditions (of which some are positive and some are negative), including the presence of inflammable material, the absence of a suitably placed sprinkler, and no doubt quite a number of others" (Mackie, 1975, p. 16). The list of conditions is incomplete because, even if causal regularities hold "in the objects", they are seldom, if ever, known in full: "Causal knowledge progresses gradually towards the formulation of such regularities, but it hardly ever gets there. Causal regularities *as known* are typically incomplete ... What we know are certain *elliptical* or *gappy* universal propositions" (Mackie, 1974, p. 66).

The sufficient-preconditions problem becomes more acute when we consider formal representations which are intended to be of practical use, because the preconditions have to be computationally tractable. McCarthy (1977, p. 1040) gives the example of using a boat to cross a river. Given that the boat is a rowing boat, that it is equipped with oars, and is manned by an oarsman, it can be used to convey two passengers across the river; provided that the boat does not leak, and provided that it does not hit a rock, and provided that it is not hit by another boat, and provided that it is not overturned by a hippopotamus, or sunk by a meteorite, or vapourized by a thermonuclear blast, etc. It seems that the list of qualifications which need to be added is limited only by the limits of our imagination. Accordingly, McCarthy calls the sufficient-preconditions problem the *qualification problem*.

In response, Deductionists might argue that representations are abstractions and that their approach works well for simple domains, in which it can be assumed that qualifications do not arise. In such domains, they might argue, the uniformity assumption is reasonable, and so deductive theories do provide a useful representation. This may well be true, but, theories of this kind cannot readily be extended to more complex (less uniform) domains because the additional complexity of the required preconditions quickly becomes overwhelming. Thus the Deductionist approach may be appropriate for certain applications, such as the mathematical analysis of high-level programming languages in which elementary commands are viewed as abstract operations on data. But it is inappropriate for the representation of predictive reasoning about *natural events*, the events of our everyday experience, because these are too irregular to be treated deductively. Moreover, the Deductionist abstraction is better thought of as an idealization, and a problematic one at that. If the preconditions of a deductive event hold on occurrence, then its effects are logically guaranteed to follow, and so no natural force can intervene to prevent them from doing so. Deductive events are thus not natural, but supernatural. This idealization creates technical difficulties when it comes to the representation of events with variable effects (including non-deterministic and context-dependent effects) as these effects should not always be deduced, and to the representation of conflicting events (events whose effects are individually consistent but jointly inconsistent) as their joint effects cannot consistently be deduced. Consequently Deductionist theories of these phenomena (some of which are discussed in the sequel) face unnecessary, self-imposed, difficulties. It is difficult to escape the conclusion that, in representing natural events deductively, Deductionism starts off on the wrong foot.

But this conclusion is hardly surprising when we consider that our predictive reasoning about natural events is *inductive*, rather than deductive, in nature.[2] The major purpose of predictive reasoning is to support practical reasoning; that is, reasoning about what to do. Predictive reasoning is normally based on *partial* knowledge (or incomplete belief), both of causal regularities, as Mackie observes, and of the contexts in which the events concerned occur. It also tends to be conjectural in that it seeks to produce reasonable conclusions on the basis of what is known. As a result it tends to produce conclusions which are both *supra-deductive* (which may not be deducible from the known) and *defeasible* (which may turn out to be wrong).[3] Accordingly, definitions of (practical, non-omniscient) rationality are typically couched in terms of the utility of the *expected* (rather than the actual) outcomes of actions. Russell and Norvig (2003, p. 36) illustrate this point as follows: "I am walking along the Champs Elysées one day and I see an old friend across the street. There is no traffic nearby and I am not otherwise engaged, so being rational, I start to cross the street. Meanwhile at 33,000 feet, a cargo door falls off a passing airliner, and before I make it to

2. In philosophy, the term 'inductive reasoning' is applied to any form of qualitative non-deductive reasoning. It thus includes enumerative induction, in which a general rule is inferred from a non-exhaustive set of inferences (for example, all of the emeralds which have been observed have been green, therefore all emeralds are green). This is the form of reasoning which underlies our knowledge of regularities. When it comes to inductive events the idea (as explained in the text below) is to produce reasonable conclusions about their outcomes on the basis of partial information. In AI, inductive reasoning of this kind is formalized in non-monotonic logics.
3. J. K. Galbraith once remarked that there are two kinds of forecasters. Those who don't know, and those who know that they don't know.

the other side of the street I am flattened. Was I irrational to cross the street? It is unlikely that my obituary would read 'Idiot attempts to cross street'".

This change in perspective results in a substantial simplification of the problem of specifying preconditions. We are no longer concerned with invariable regularities of sequence, with necessary connections between events and their effects, but rather with regularities of sequence which *normally* hold; with "fairly dependable regularities of sequence" (Russell, 1913, p. 8), with expected connections between events and their effects. Consequently we can define preconditions which, together with their associated events, form conditions which are *normally* sufficient for the associated effects, and which are otherwise minimal in the sense that no part of them is redundant. Preconditions of this kind tend to be both tractable (simple) and useful (to occur frequently in practice).

We can now give a Humeian account of predictions involving natural events in terms of (fairly) dependable regularities and the expectations that they engender. If an event occurs, its preconditions obtain, and we are not aware of anything which will prevent the effects from following, then we form a clear expectation (we "know") that the effects will follow, and so it is rational (reasonable) to predict them. For example, if block A is unstacked from block B, the preconditions obtain, and we are not aware of a preventer, then we clearly expect, and so predict, that A will no longer be on B. More complex cases involve conflicting events. If, in such a case, we have a clear expectation of the outcome, then it is rational to predict it. For example, we clearly expect, and so predict, that the airliner door will crush, rather than bounce off, the intrepid pedestrian. In this example we consider that two conflicting outcomes are possible, but that only one of them is probable. However, in other cases our expectations are unclear; we are torn between conflicting expectations and so do not "know" what to expect. In such cases it seems reasonable to adopt a *cautious* approach, and restrict our predictions to those effects that we clearly expect. For example, if a fair coin is tossed, then we do not have a clear expectation as to which side it will land. We "half" expect that it will land on heads and we "half" expect that it will land on tails. We consider that two conflicting outcomes are equally probable. So caution dictates that we should predict that the coin lands on one side or the other, but that we should not predict which of the two sides it will land on. Note that, as expectations are based on incomplete knowledge, the predictions which are based on them are defeasible. For example, if, unbeknown to us, block A is glued to block B when A is unstacked from B, then the event does not have the effects that we predict it will have.

We can thus begin by thinking of natural events as defeasible STRIPS events, as STRIPS-like events whose effects do not always follow them (when they occur and their preconditions are true), and inferring their effects inductively. Accordingly, events of this kind will be called *inductive events*, and the view that natural events should be represented inductively will be called *Inductionism*. So if 'logic' is understood to include both deductive and inductive inference, then the Inductionist objection to Deductionism can be stated succinctly: Deductionism is a logical mistake.

This paper can be seen as an argument for Inductionism. It begins by presenting a basic theory of inductive events and then uses this as the basis for a more comprehensive theory of natural events.

The formal language in which the theory is expressed is defined in the next section, and the basic theory of inductive events is then given in Section 3. The theory builds on the ideas of McCarthy (1986, §9), Lifschitz (1987), and Shoham (1988), and is *logico-pragmatic* in nature; that is, it consists of a set of axioms together with a formal *pragmatics*, which, given a formal theory containing the axioms, interprets it in a particular way, and in doing so, generates the predictions of the theory. The basic theory of inductive events provides the basis of a solution to the qualification problem and integrates this with the basis of a solution to the complementary *frame problem* (McCarthy & Hayes, 1969, p. 487); that is, the problem of inferring what is unchanged by the occurrence of an event (or, more generally, by the occurrence of several simultaneously occurring events).

In Section 4, the basic theory of inductive events is extended by introducing a distinction between inductive events which are *primary* and those which are *secondary*. Whereas primary events occur independently, secondary events are *invoked* by other (primary or secondary) events in appropriate contexts, and are *causally dependent* on them. In the simplest case, a primary event invokes a secondary event and is the only event to do so. In which case, the secondary event succeeds (is followed by its effects) only if the primary event which invoked it succeeds. This extension to the basic theory makes it possible for inductive events to have additional context-dependent effects, thereby providing the basis for a solution to the *ramification problem* (Ginsberg & Smith, 1988); that is, the problem of representing the indirect, context-dependent, effects of events. For example, if an agent is holding a block and the agent moves, then the move-agent event invokes a causally dependent move-block event, with the effect that the block moves only if the agent does. This extension also makes it possible to represent events with non-deterministic effects. For example, the non-deterministic event of tossing a fair coin can be represented by having the event invoke two conflicting deterministic events, one of which has the effect that the coin lands on heads, the other that it lands on tails.

The theory of natural events is completed in Section 5, which deals with the problem of representing defeasible preferences over the outcomes of conflicting simultaneous events. When two events conflict we often have a clear expectation about the outcome. For example, if two agents attempt to go through a door simultaneously and only one of them can succeed, then it is reasonable to expect that the stronger one will do so. However this expectation is defeasible. The stronger agent may fail for some independent reason (the agent may slip, say), in which case the preference is reversed and we expect that the weaker agent will succeed (although the weaker agent may also slip, etc.). In order to represent defeasible asymmetric expectations of this kind *event preferences* are introduced, and the formal pragmatics of the basic theory is refined in order to interpret them correctly.

A philosophical justification of the theory of natural events is given in Section 6, and related work is discussed in Section 7.

Although causal notions underlie much of the development of the theory of natural events, there is no explicit reference to causation in the theory. This is because it is intended to provide the basis for a definition of sufficient causation which forms part of a larger theory of causation (Bell, 2004, 2006, 2008).[4]

4. According to the theory, the occurrence of event e in context c is a sufficient cause of effect ϕ if the occurrence of e in c is sufficient to ensure ϕ; for instance, if e succeeds at time t and ϕ is a logical consequence of e's effects at $t+1$, then e is a sufficient cause of ϕ. The definition of causation is then obtained by requiring that sufficient causes also satisfy a refinement of Lewis's (1986, Ch. 21) counterfactual-dependence

2. The Event Language $\mathcal{EL}$

The theory of events is expressed in the event language $\mathcal{EL}$, which has been developed in order to represent and reason about events and their effects, on the basis of partial information, at successive points in time. This section begins with an informal introduction and then gives a formal account.

In order to represent epistemic partiality in a natural and economical way $\mathcal{EL}$ is based on Kleene's (1952, §64) strong three-valued logic.[5] Kleene introduced the truth value 'undefined' in order to accommodate undecidable mathematical sentences. However he also suggested that 'undefined' could be interpreted as 'unknown', where: "'unknown' is a category into which we can regard any proposition as falling, whose value we either do not know or choose for the moment to disregard; and it does not the exclude the other two possibilities 'true' and 'false'" (p. 335). Thus understood, 'undefined' is not a truth value on a par with 'true' and 'false', and its introduction is intended as a practical, logically conservative, way of reasoning with partial information; rather than a revolutionary attack on classical logic.[6]

In keeping with this interpretation, the truth value of a sentence should be classical (either 'true' or 'false') if enough is known to determine it. The formal semantics for the propositional case can thus be given as follows. A model, M, consists of a possibly partial evaluation function, V, which assigns at most one classical truth value to each atomic proposition. The truth ($\models$) and falsity ($=\!\!|$) of sentences in M is then defined by the following truth and falsity conditions:

$$
\begin{aligned}
M \models p \quad &\text{iff} \quad V(p) = \mathit{true}\\
M =\!\!|\, p \quad &\text{iff} \quad V(p) = \mathit{false}\\
M \models \neg\phi \quad &\text{iff} \quad M =\!\!|\, \phi\\
M =\!\!|\, \neg\phi \quad &\text{iff} \quad M \models \phi
\end{aligned}
$$

condition: the occurrence of event e is a cause of effect ϕ in context c iff (i) e is a sufficient cause of ϕ in c, and (ii) ϕ depends on e in the closest context to c in which e is the only sufficient cause of ϕ.

5. Kleene's logic will be familiar to readers with a background in philosophical logic (it is, for example, used by Kripke, 1975, as a basis for his theory of truth), and the choice to use it here is likely to appear a natural one to them. However, Kleene's logic may be unfamiliar to readers in the 'reasoning about actions and change' community, and they may well wonder why I have not used a more established classical language such as the Situation Calculus (McCarthy & Hayes, 1969). A full justification of my choice would involve a lengthy comparison of languages. In short, it is simpler to acknowledge that epistemic partiality is a ubiquitous feature of predictive reasoning and to deal with it directly, as in Kleene's logic, rather than indirectly in a classical logic; by means of syntactic encoding and circumscription (as in the Situation Calculus), or by using modal logic (as in TK, Shoham, 1988). The representation of partiality in Kleene's logic is also optimal; because there is no cost associated with the representation of what is unknown. By contrast, partiality in classical reasoning requires the consideration of a class of models (or possible worlds) which is large enough to ensure that unwanted "noise" (arbitrary, but compulsory, assignments of classical truth values to sentences whose truth values are not determined by the theory in question) is eliminated. This profligacy is significant when considering the contemplated model-building implementation of event theories (Bell, 1996, §1); see the remarks on implementation in Section 8. Finally, as indicated in the introduction and Footnote 4, the theory of events is intended to form part of a larger theory of causation, in which $\mathcal{EL}$ is embedded in a partial modal language.
6. Confirmed classicists can thus rest assured that they are not being threatened with anything radical, such as "the Bolshevik menace of Brouwer and Weyl" (remark on Intuitionism attributed to F.P. Ramsey by Blackburne, 1994).

$$
\begin{array}{lll}
M \models \phi \wedge \psi & \text{iff} & M \models \phi \text{ and } M \models \psi \\
M \mathrel{=\!\!|} \phi \wedge \psi & \text{iff} & M \mathrel{=\!\!|} \phi \text{ or } M \mathrel{=\!\!|} \psi
\end{array}
$$

So a sentence $\neg\phi$ is true if ϕ is false, is false if ϕ is true, and is undefined otherwise; and the sentence $\phi \wedge \psi$ is true if ϕ and ψ are both true, false if either is false, and is undefined otherwise. Note that when the evaluation function is total this semantics is equivalent to the semantics for classical propositional calculus. So the essential difference between the two semantics is the classical assumption that the evaluation function is total; the additional requirement that V assigns at least one classical truth value to each atomic proposition.

Further operators can be defined as in classical logic. In particular, inclusive disjunction is defined as: $\phi \vee \psi =_{\mathrm{Df}} \neg(\neg\phi \wedge \neg\psi)$; so $\phi \vee \psi$ is true if either disjunct is true, is false if both disjuncts are false, and is undefined otherwise. And exclusive disjunction is defined as: $\phi \oplus \psi =_{\mathrm{Df}} (\phi \wedge \neg\psi) \vee (\neg\phi \wedge \psi)$; so $\phi \oplus \psi$ is true if the truth values of ϕ and ψ are both defined and are different, false if the truth values of ϕ and ψ are both defined and are the same, and is undefined otherwise.[7]

Kleene's logic can perhaps be called "demi-classical", as it becomes classical when the truth values of all of the constituent atomic sentences are classical. Unsurprisingly then, the use of Kleene's logic does not, of itself, solve any of the problems of predictive reasoning beyond that of representing partiality. There is, for example, no reliance on a special "causal" notion of consequence, such as that of Linear Logic (Girard, 1987).

The expressiveness of Kleene's language is greatly enhanced by adding a classically-valued "definedness" operator to it. The sentence $\mathsf{D}\phi$ is true if ϕ is defined (is either true or false), and is false otherwise:

$$
\begin{array}{lll}
M \models \mathsf{D}\phi & \text{iff} & \text{either } M \models \phi \text{ or } M \mathrel{=\!\!|} \phi \\
M \mathrel{=\!\!|} \mathsf{D}\phi & \text{iff} & \text{neither } M \models \phi \text{ nor } M \mathrel{=\!\!|} \phi
\end{array}
$$

Further classically-valued operators can now be defined as follows:

$$
\begin{array}{lll}
\mathsf{T}\phi =_{\mathrm{Df}} \mathsf{D}\phi \wedge \phi & \mathsf{F}\phi =_{\mathrm{Df}} \mathsf{D}\phi \wedge \neg\phi & \mathsf{U}\phi =_{\mathrm{Df}} \neg\mathsf{D}\phi \\
\phi \rightarrow \psi =_{\mathrm{Df}} \neg\mathsf{T}\phi \vee \mathsf{T}\psi & \multicolumn{2}{l}{\phi \equiv \psi =_{\mathrm{Df}} (\mathsf{T}\phi \wedge \mathsf{T}\psi) \vee (\mathsf{F}\phi \wedge \mathsf{F}\psi) \vee (\mathsf{U}\phi \wedge \mathsf{U}\psi)}
\end{array}
$$

Thus, for sentences ϕ and ψ: $\mathsf{T}\phi$ is true if ϕ is true, and is false otherwise; $\mathsf{F}\phi$ is true if ϕ is false, and is false otherwise; $\mathsf{U}\phi$ is true if ϕ is undefined, and is false otherwise; $\phi \rightarrow \psi$ is

7. Readers who are unfamiliar with Kleene's logic may wish to check the definitions against the semantics. For example, $M \models \phi \vee \psi$ iff $M \models \neg(\neg\phi \wedge \neg\psi)$ iff $M \mathrel{=\!\!|} \neg\phi \wedge \neg\psi$ iff $[M \mathrel{=\!\!|} \neg\phi$ or $M \mathrel{=\!\!|} \neg\psi]$ iff $[M \models \phi$ or $M \models \psi]$.

true if ψ is true or ϕ is not, and is false otherwise;[8] and $\phi \equiv \psi$ is true if ϕ and ψ have the same truth value, and is false otherwise.[9]

The first-order extension, given by Kleene, is straightforward. The universal sentence $\forall x\phi$ is true if ϕ is true for all assignments to x, false if ϕ is false for some such assignment, and is undefined otherwise. The existential quantifier is defined as in classical logic: $\exists x\phi =_{\mathrm{Df}} \neg\forall x\neg\phi$; thus $\exists x\phi$ is true if ϕ is true for some assignment to x, is false if ϕ is false for all such assignments, and is undefined otherwise.

In order to represent change, events and time points are added as an additional sort. For simplicity, the order of time is assumed to be discrete and linear.

An *object atom* is an atom of the form $r(u_1, \ldots, u_n)(t)$, where r is an object relation symbol, the u_i are object terms, and t is a time-point term. For example, the object atoms $At(O, L)(1)$ and $\neg At(O, L)(2)$ state respectively that object O is at location L at time 1, and that O is not at L at time 2.

An *event atom* is an atom of the form $r(u_1, \ldots, u_n)(t)$, where r is an event relation symbol, the u_i are event terms, and t is a time-point term. For example, the event atom $Occ(Move(O, L1, L2))(3)$ states that the event consisting of object O moving from location $L1$ to location $L2$ occurs at time 3.

The intuition behind the fact-event distinction is that events are active "agents" (causes) of change, while facts are passive "patients" of change which persist through time until affected by some event (until some event causes them to change).[10]

In order to represent the persistence of facts, second-order quantification over object relations and second-order relations are added to $\mathcal{EL}$. An *object-relation atom* is an atom of the form $r(u_1, \ldots, u_n)(t)$, where r is a second-order relation symbol, the u_i are object relation symbols or object terms, and t is a time-point term. For example, the object-relation atom $Inert(At, \langle O, L\rangle)(4)$ states that the At relation is inert for objects O and L at time 4.

8. The conditional $\phi \rightarrow \psi$ captures much of the flavour of classical material implication. This can be emphasized by defining the weaker conditional of Kleene's logic: $\phi \supset \psi =_{\mathrm{Df}} \neg\phi \vee \psi$. This conditional is inadequate, at least for present purposes, because it is undefined, rather than false, when ϕ is true and ψ is undefined. This is not the case for the stronger conditional, $\rightarrow$, which could equally have been defined as: $\mathsf{T}\phi \supset \mathsf{T}\psi$. This definition makes it clear that a conditional $\phi \rightarrow \psi$ states a constraint which must be satisfied if ϕ is true, but which can otherwise be ignored. The conditional does not quite capture all of the meaning of classical material implication, as it does not satisfy the (implicitly understood) condition that if ψ is false, then so is ϕ. If desired, it is possible to define a stronger conditional, which better represents classical material implication, as follows: $\phi \longrightarrow \psi =_{\mathrm{Df}} (\phi \rightarrow \psi) \wedge (\neg\psi \rightarrow \neg\phi)$; or, equivalently, $\phi \longrightarrow \psi =_{\mathrm{Df}} (\mathsf{T}\phi \supset \mathsf{T}\psi) \wedge (\mathsf{F}\psi \supset \mathsf{F}\phi)$. The equivalence operator can then be defined as follows: $\phi \equiv \psi =_{\mathrm{Df}} (\phi \longrightarrow \psi) \wedge (\psi \longrightarrow \phi)$.
9. The addition of a truth-value designation operator is not new. Bochvar (1939) added a truth operator to a different system of connectives, and the undefined operator U is a special case of Rosser and Turquette's (1952) J_k operator. Barringer, Cheng, and Jones (1984) give a natural deduction system for Kleene's logic and the defined operator D. Their system is readily extended to include the operators defined above; the additional introduction and elimination rules for connectives such as $\rightarrow$ and $\equiv$ simply fold and unfold the definitions. Similar languages have been used as a basis for the formalization of non-monotonic reasoning (Doherty, 1996).
10. The fact-event distinction is similar to McCarthy's (1986, §9) fluent-event distinction. The case for adding events to the ontology of facts is argued by Davidson (1980). Lewis (1986, Ch. 23) goes further and treats facts as events.

A *temporally-indexed* relation is any relation whose atoms are temporally indexed; whose atoms are of the form $r(u_1, \ldots, u_n)(t)$. The primitive atemporal relations of $\mathcal{EL}$ are temporal precedence, '$<$', and identity, '$=$'. Further atemporal relations can be defined in terms of temporally-indexed relations as follows:

$$r(u_1, \ldots, u_n) =_{\mathrm{Df}} \forall t\, r(u_1, \ldots, u_n)(t) \;.$$

For example, $\forall t\, \mathit{Phys}(At)(t)$ can be abbreviated to $\mathit{Phys}(At)$, which states that At is (eternally) a physical relation.

The formal semantics of $\mathcal{EL}$ can be sketched as follows. Models contain a set of objects, a set of event types, a time frame consisting of a set of time points ordered by a (discrete and linear) precedence relation, and functions for interpreting terms and relations. The twin notions of *satisfaction* and *violation* of a formula, by a variable assignment in a model, are defined by means of a parallel recursion. The truth (falsity) of a sentence in a model is then defined in terms of satisfaction (violation) by all assignments for that model. As in classical logic, a model M is said to be a model of a sentence ϕ (a set of sentences Θ) if ϕ is true in M (if every sentence in Θ is true in M), and a set of sentences Θ semantically entails a sentence ϕ iff all models of Θ are also models of ϕ.

In the remainder of this section the formal syntax and semantics of $\mathcal{EL}$ are defined. Readers may wish to skip to the next section and return to consult the details as necessary.

Definition 1 *The four* sorts *of $\mathcal{EL}$ are identified by the following letters: O (objects), T (time points), E (events), and R (object relations). The* vocabulary *of $\mathcal{EL}$ consists of the symbols '$<$', '$=$', '$\neg$', 'D', '$\wedge$', '$\forall$', '(', and ')', and the following countable sets of symbols:*

- *C_S (constants of each sort S),*
- *V_S (variables of each sort S),*
- *F_S (function symbols of each arity $n \geq 1$ of each sort S),*
- *R_O, R_E, R_R (relation symbols of each arity $n \geq 0$ of sorts O, E,and R).*

The sets R_O and C_R are required to be the same set. Otherwise the above sets are required to be mutually disjoint. Furthermore, V_R is assumed to contain variables of each arity $n \geq 0$.

Definition 2 *The* terms *of $\mathcal{EL}$ are defined as follows.*

- *$\mathit{term}_S = C_S \cup V_S \cup \{f(u_1, \ldots, u_n) : n\text{-ary } f \in F_S, u_i \in \mathit{term}_S\}$ for $S \in \{O, T, R\}$.*
- *$\mathit{term}_E = C_E \cup V_E \cup \{f(u_1, \ldots, u_n) : n\text{-ary } f \in F_E, u_i \in \mathit{term}_O\}$.*

Definition 3 *$\mathcal{EL}$ is the minimal set which satisfies the following conditions.*

- *If $t, t' \in \mathit{term}_T$ then $t < t' \in \mathcal{EL}$.*
- *If S is any sort and $u, u' \in \mathit{term}_S$, then $u = u' \in \mathcal{EL}$.*
- *If $u_1, \ldots, u_n \in \mathit{term}_O$, r_O is an n-ary relation symbol in R_O, and $t \in \mathit{term}_T$, then $r_O(u_1, \ldots, u_n)(t) \in \mathcal{EL}$.*

- *If S is of sort E or R, $u_1, \ldots, u_n \in$ terms$_S$ and $w_1, \ldots, w_m \in$ term$_O$ (where $m = 0$ if $n = 0$), r_S is an $n+m$-ary relation symbol in R_S, and $t \in$ term$_T$, then $r_S(u_1, \ldots, u_n, w_1, \ldots, w_m)(t) \in \mathcal{EL}$.*
- *If $u_1, \ldots, u_n \in$ term$_O$, v_R is an n-ary variable in V_R, and $t \in$ term$_T$, then $v_R(u_1, \ldots, u_n)(t) \in \mathcal{EL}$.*
- *If $\phi, \psi \in \mathcal{EL}$, then $\neg\phi \in \mathcal{EL}$, $\mathsf{D}\phi \in \mathcal{EL}$, and $(\phi \wedge \psi) \in \mathcal{EL}$.*
- *If S is any sort, $v \in V_S$ and $\phi \in \mathcal{EL}$, then $\forall v\phi \in \mathcal{EL}$.*

The members of $\mathcal{EL}$ are called formulas *(of $\mathcal{EL}$). Those formulas in which no variable occurs free are called* sentences *(of $\mathcal{EL}$).*

Models for $\mathcal{EL}$ consist of a set $\mathcal{O}$ of objects, a set $\mathcal{E}$ of event types, a temporal frame $\langle \mathcal{T}, \prec_\mathcal{T} \rangle$ (where $\mathcal{T}$ is a set of time points and $\prec_\mathcal{T}$ is the before-after relation on $\mathcal{T}$), and interpretation functions for terms and relations. For simplicity, time is assumed to be isomorphic to the integers. The denotations of terms are always defined and do not vary over time. By contrast, temporally-indexed relations may be partial and may vary over time. Consequently each temporally-indexed relation is interpreted by a function from time points to partial characteristic functions. Where defined, the partial characteristic function associated with a time point maps instances of the relation to $\{$*true*, *false*$\}$.

Definition 4 *A* model *for $\mathcal{EL}$ is a structure $\langle \mathcal{O}, \mathcal{E}, \langle \mathcal{T}, \prec_\mathcal{T} \rangle, \mathcal{F}, \mathcal{R}, \mathcal{V} \rangle$, where:*

- *$\mathcal{O}$, $\mathcal{E}$ and $\mathcal{T}$ are mutually disjoint, non-empty, countable sets,*
- *$\prec_\mathcal{T}$ is a binary relation on $\mathcal{T}$ which is isomorphic to the integers,*
- *$\mathcal{R} = \langle \mathcal{R}_O, \mathcal{R}_E, \mathcal{R}_R \rangle$. $\mathcal{R}_O$ is a set of partial functions of each arity $n \geq 0$ of type $\mathcal{O}^n \to \{$true, false$\}$. For $\langle S, \mathcal{S} \rangle \in \{\langle E, \mathcal{E} \rangle, \langle R, \mathcal{T} \to \mathcal{R}_O \rangle\}$, $\mathcal{R}_S$ is a set of partial functions of each arity $n + m \geq 0$ of type $\mathcal{S}^n \times \mathcal{O}^m \to \{$true, false$\}$.*
- *$\mathcal{F} = \langle \mathcal{F}_O, \mathcal{F}_T, \mathcal{F}_E, \mathcal{F}_R \rangle$. For each $\langle S, \mathcal{S} \rangle \in \{\langle O, \mathcal{O} \rangle, \langle T, \mathcal{T} \rangle, \langle R, \mathcal{T} \to \mathcal{R}_O \rangle\}$, $\mathcal{F}_S$ is a set of functions of each arity $n \geq 1$ of type $\mathcal{S}^n \to \mathcal{S}$. $\mathcal{F}_E$ is a set of functions of each arity $n \geq 1$ of type $\mathcal{O}^n \to \mathcal{E}$.*
- *$\mathcal{V} = \langle \langle \mathcal{V}_O^C, \mathcal{V}_T^C, \mathcal{V}_E^C, \mathcal{V}_R^C \rangle, \langle \mathcal{V}_O^F, \mathcal{V}_T^F, \mathcal{V}_E^F, \mathcal{V}_R^F \rangle, \langle \mathcal{V}_O^R, \mathcal{V}_E^R, \mathcal{V}_R^R \rangle \rangle$ is an interpretation function such that: $\mathcal{V}_S^C : C_S \to \mathcal{S}$ for each $\langle S, \mathcal{S} \rangle \in \{\langle O, \mathcal{O} \rangle, \langle T, \mathcal{T} \rangle, \langle E, \mathcal{E} \rangle, \langle R, \mathcal{T} \to \mathcal{R}_O \rangle\}$, $\mathcal{V}_S^F : F_S \to \mathcal{F}_S$, $\mathcal{V}_S^R : R_S \to (\mathcal{T} \to \mathcal{R}_S)$, and $\mathcal{V}_O^R = \mathcal{V}_R^C$.*

Definition 5 *A* variable assignment *for an $\mathcal{EL}$ model is a function $g = \langle g_O, g_T, g_E, g_R \rangle$, where for $\langle S, \mathcal{S} \rangle \in \{\langle O, \mathcal{O} \rangle, \langle T, \mathcal{T} \rangle, \langle E, \mathcal{E} \rangle\}$, $g_S : V_S \to \mathcal{S}$, and $g_R : V_R \to (\mathcal{T} \to \mathcal{R}_O)$. For $\mathcal{EL}$-model M, with interpretation function $\mathcal{V}$ and variable assignment g for M, the* term evaluation function *$\mathcal{V}_g$ is defined, on the terms and relation symbols of $\mathcal{EL}$, as follows:*

$$\mathcal{V}_g(u) = \begin{cases} g_S(u) & \text{if } u \in V_S, \\ \mathcal{V}_S^C(u) & \text{if } u \in C_S, \\ \mathcal{V}_S^F(f)(\mathcal{V}_g(u_1), \ldots, \mathcal{V}_g(u_n)) & \text{if } u = f(u_1, \ldots, u_n) \in \textit{terms}_S, \\ \mathcal{V}_S^R(u) & \text{if } u \in R_S. \end{cases}$$

Table 1: Satisfaction and violation conditions for $\mathcal{EL}$ (see Definition 6)

$$
\begin{array}{rcl}
M, g \models t < t' & \text{iff} & \langle \mathcal{V}_g(t), \mathcal{V}_g(t') \rangle \in \prec_{\mathcal{T}} \\
M, g \mathrel{=\!\!|} t < t' & \text{iff} & \langle \mathcal{V}_g(t), \mathcal{V}_g(t') \rangle \notin \prec_{\mathcal{T}} \\
M, g \models u = u' & \text{iff} & \mathcal{V}_g(u) \text{ is } \mathcal{V}_g(u') \\
M, g \mathrel{=\!\!|} u = u' & \text{iff} & \mathcal{V}_g(u) \text{ is not } \mathcal{V}_g(u') \\
M, g \models u(u_1, \ldots, u_n)(t) & \text{iff} & (\mathcal{V}_g(u)(\mathcal{V}_g(t)))(\mathcal{V}_g(u_1), \ldots, \mathcal{V}_g(u_n)) = \mathit{true} \\
M, g \mathrel{=\!\!|} u(u_1, \ldots, u_n)(t) & \text{iff} & (\mathcal{V}_g(u)(\mathcal{V}_g(t)))(\mathcal{V}_g(u_1), \ldots, \mathcal{V}_g(u_n)) = \mathit{false} \\
M, g \models \neg\psi & \text{iff} & M, g \mathrel{=\!\!|} \psi \\
M, g \mathrel{=\!\!|} \neg\psi & \text{iff} & M, g \models \psi \\
M, g \models \mathsf{D}\psi & \text{iff} & \text{either } M, g \models \psi \text{ or } M, g \mathrel{=\!\!|} \psi \\
M, g \mathrel{=\!\!|} \mathsf{D}\psi & \text{iff} & \text{neither } M, g \models \psi \text{ nor } M, g \mathrel{=\!\!|} \psi \\
M, g \models \psi \wedge \chi & \text{iff} & M, g \models \psi \text{ and } M, g \models \chi \\
M, g \mathrel{=\!\!|} \psi \wedge \chi & \text{iff} & M, g \mathrel{=\!\!|} \psi \text{ or } M, g \mathrel{=\!\!|} \chi \\
M, g \models \forall v\psi & \text{iff} & M, g' \models \psi \text{ for every } g' \text{ such that } g \approx_v g' \\
M, g \mathrel{=\!\!|} \forall v\psi & \text{iff} & M, g' \mathrel{=\!\!|} \psi \text{ for some } g' \text{ such that } g \approx_v g'
\end{array}
$$

Definition 6 *Let M be an $\mathcal{EL}$ model, g be a variable assignment for M, and let $g \approx_v g'$ indicate that variable assignment g' differs from g at most on the assignment to variable v. Then g* satisfies *an $\mathcal{EL}$-formula ϕ in M (written $M, g \models \phi$) or* violates *ϕ in M (written $M, g \mathrel{=\!\!|} \phi$) according to the clauses given in Table 1.*

Let M be an $\mathcal{EL}$ model. Then a formula ϕ is true *in M (written $M \models \phi$) if $M, g \models \phi$ for all variable assignments g; a formula ϕ is* false *in M (written $M \mathrel{=\!\!|} \phi$) if $M, g \mathrel{=\!\!|} \phi$ for all variable assignments g; M is a* model of a sentence *ϕ iff ϕ is true in M; and M is a* model of a set of sentences *Θ iff M is a model of every sentence in Θ.*

A set of sentences Θ (semantically) entails *a sentence ϕ (written $\Theta \models \phi$) iff every model of Θ is also a model of ϕ.*

3. Inductive Events

The formal theory of natural events is introduced in stages, beginning, in this section, with the basic theory of inductive events.

Definition 7 *The* theory of inductive events, *Θ_{Ind}, consists of the axioms given in Table 2; thus $\Theta_{Ind} = \{(1), (2), (3)\}$. An* event theory *is any set of $\mathcal{EL}$ sentences which contains Θ_{Ind}.*

Axiom (1) defines the notion of *success*, and states that event e succeeds at time t iff it is true that e occurs at t, the preconditions of e are true at t, and the effects of e are true at $t{+}1$.[11] The presence of the truth operator in this axiom ensures that the relation *Succ*

11. Thus defined, the success or failure of an event is a simple, objective, matter of whether its occurrence is accompanied by its preconditions and is followed by its effects. So when speaking of the success or failure

Table 2: The theory of inductive events, Θ_{Ind}

$$\forall e, t(Succ(e)(t) \equiv \mathsf{T}(Occ(e)(t) \wedge Pre(e)(t) \wedge Eff(e)(t{+}1))) \tag{1}$$
$$\forall R\, \mathsf{T}(Phys(R) \oplus Theo(R)) \tag{2}$$
$$\forall R, \overline{x}, t(Inert(R, \langle\overline{x}\rangle)(t) \equiv (Phys(R) \wedge (R(\overline{x})(t) \equiv R(\overline{x})(t{+}1)))) \tag{3}$$

is classical in the sense that every instance of it is either true or false. It is thus possible to reason classically about success and failure on the basis of partial information.

In view of the fact-event distinction, it is also necessary to represent *inertia*; that is, the temporal persistence of facts which are not changed by events. The definition of inertia begins with a distinction between *physical facts* (represented by *physical relations*) and *theoretical facts* (represented by *theoretical relations*). Intuitively, physical facts are facts about the world as we directly observe it, whereas theoretical facts are the product of our more complex reflection on (theorizing about) the physical facts.[12] For example, in a representation of the blocks world the locations of blocks might be represented by the object relation *At*. This can then be used to define the relation *Clear*, which is true of a location at a point in time if there are no blocks at that location at that point in time. In this theory, the relation *At* is naturally classified as a physical relation (as it represents the physical locations of blocks) while the relation *Clear* is naturally classified as a theoretical relation (as it represents a, comparatively complex, property of locations which is defined in terms of the locations of blocks). Note that, as theoretical facts are (ultimately) defined in terms of physical facts, the theoretical facts *supervene* on the physical facts; that is, fixing the physical facts at any point in time also fixes the theoretical facts at that point in time.

In event theories, physical and theoretical relations are identified by means of the second-order predicates *Phys* and *Theo* respectively. Thus Axiom (2) states that every object relation is either a physical relation or a theoretical relation. As a matter of notational convenience the convention is adopted that any object relation which is not declared to be a theoretical relation is a physical one. This convention is enforced by the formal pragmatics discussed below.

The inertia of physical facts is defined by Axiom schema (3); which, for simplicity, will henceforth be called an axiom. In the axiom, R is an n-ary object relation symbol, $\overline{x}$ is a vector $x_1, \ldots, x_n$ of object variables, and $\langle\overline{x}\rangle$ is the list $\langle x_1, \ldots, x_n\rangle$. The axiom states that R is inert for the objects referred to by $\overline{x}$ at time t iff R is a physical relation and

of an event no end or purpose (no teleology) is implied; we could equally talk of an event occurrence being complete or incomplete. It is natural to talk informally of intentional actions succeeding or failing, for example of an agent succeeding or failing in their intention to move to a particular location. But no attempt is made to represent this intentionality formally.

12. Physical facts can be thought of as Quine's (1995, Chs. 2-3) *observational predications*. These are compounds of more primitive *observation sentences*, which are "the human equivalent of bird-calls and apes' cries" (p. 22). For example, the observational sentences 'Black' (or 'That's black') and 'Dog' (or 'That's a dog') might be combined in the observational predication 'Black dog' (or 'The dog is black'). Theoretical facts are the result of more complex compounding, involving logical connectives and, especially, reification.

the truth values of the object atoms $R(\overline{x})(t)$ and $R(\overline{x})(t{+}1)$ are equivalent. Note that the relation *Inert* is classical in event theories; because *Phys* is classical (by Axiom (2)) and the right-hand equivalence is classically valued. Note also that, for event theories which do not contain occurrences of the *Theo* relation, Axiom (2) is unnecessary and Axiom (3) can be simplified accordingly.

We turn now to the intended interpretation of event theories.

On the intended interpretation, Axiom (1) is used to generate the expected outcomes of events. If $\mathit{Pre}(e)(t)$ and $\mathit{Occ}(e)(t)$ are both true, and it is consistent to assume that the *success atom* $\mathit{Succ}(e)(t)$ is true, then this *success assumption* is made, and the axiom is used to conclude that the expected effects, $\mathit{Eff}(e)(t{+}1)$, are true. Thus interpreted the axiom states that, when accompanied by their preconditions, occurring events are normally sufficient for (are normally followed by) their effects. When interpreted in this way, the axiom amounts to a common sense law of change.

Similarly, on the intended interpretation, Axiom (3) is used to generate the expected persistence of physical facts. If R is a physical relation, the object atom $R(\overline{x})(t)$ is defined, and it is consistent to assume that the *inertia atom* $\mathit{Inert}(R, \langle\overline{x}\rangle)(t)$ is true, then this *inertia assumption* is made, and the axiom is used to conclude that the truth value of $R(\overline{x})$ persists from t to $t{+}1$.[13] Thus interpreted, the axiom states that physical facts normally persist, and so it amounts to a common sense law of inertia.

The intended interpretations of axioms (1) and (3) often conflict. For example, if an unstack-A-from-B event occurs, its preconditions are true, and no other relevant facts or events are involved, then it is consistent to assume that the unstack event succeeds, and it is consistent to assume that the fact that A is on B is inert, but the assumptions cannot both be made because the success of the unstack event implies that A is no longer on B. In such cases I suggest that change should always have priority over inertia, that success assumptions should always have priority over inertia assumptions. This conflict resolution principle can be defended by appealing to our regularity-based (Humeian) expectation of change. Thus, in the case at hand, experience has taught us that unstack events of the sort described normally succeed, and so we form a clear expectation that the effects will follow. By contrast, giving priority to inertia would, contrary to expectation, result in nothing changing, and adopting a neutral stance would, contrary to expectation, produce an unclear outcome.

The intended interpretation of event theories is enforced by their formal *pragmatics*, which defines the class of *preferred models* of any given event theory. In this section the notion of preference$_1$ is defined. This is later refined to preference$_2$ in Section 5.

In order to enforce the convention that object-relations are physical unless stated otherwise, the preferred$_1$ models of an event theory should all be models of the theory in which the (positive) domain of the *Phys* relation is as large as it can be. Let us say that a model M is a *Phys-maximal* model of an event theory Θ if M is a model of Θ and, for any model M' of Θ, if $\{R : M \models \mathit{Phys}(R)\} \subseteq \{R : M' \models \mathit{Phys}(R)\}$ then $M = M'$. Then the preferred$_1$ models of Θ should all be *Phys*-maximal models of Θ.

13. If $R(\overline{x})(t)$ is not defined, then, as will become clear, this atom can safely be ignored. For if $\mathit{Inert}(R, \langle\overline{x}\rangle)(t)$ can consistently be assumed, then the formal pragmatics (in particular, the minimization of evidential atoms at $t{+}1$) ensures that $R(\overline{x})(t{+}1)$ is undefined, thereby satisfying the axiom.

Beyond this requirement, we can get a clearer idea of what a preferred$_1$ model of an event theory should look like by considering an inductive version of the "canonical" example, known as the "Yale Shooting Problem" (Hanks & McDermott, 1987). At time 1 a gun is loaded and pointed at Fred. Nothing relevant happens at time 2. At time 3 the gun is fired. If the gun is still loaded at time 3, then, in the absence of further information, we expect that the shot will prove fatal and that Fred will no longer be alive at time 4. This example can be represented formally by the theory $\Theta_1 = \Theta_{Ind} \cup \{(4), (5), (6)\}$, where:

$$\forall t(Pre(Shoot)(t) \equiv (Alive(t) \wedge Loaded(t))) \tag{4}$$

$$\forall t(Eff(Shoot)(t) \equiv \neg Alive(t)) \tag{5}$$

$$Alive(1) \wedge Loaded(1) \wedge Occ(Shoot)(3) \tag{6}$$

Thus the preconditions for a *Shoot* event are that the gun is loaded and the victim is alive (Axiom (4)), and its effect (if successful) is that the victim is not alive (Axiom (5)).

For model M and time point t, let M/t denote the set of all object or occurs literals with temporal index $t' \leq t$ which are true in M, and let $M(t) = M/t \setminus M/t{-}1$.[14] So M/t can be thought of as the history that is represented by M up to (and including) t. And, at t, $M(t)$ can be thought of as representing what is known at the present moment, as the *evidential context* on which predictions about $t{+}1$ are based. We can also think of there being a dynamic *conjectural context* at t, which consists of the set of success and inertia assumptions which correspond to our expectations about $t{+}1$. Assumptions are added to the conjectural context if they are consistent with the *current context* (the union of the evidential context and the current conjectural context) and the *background theory* (the laws of the given event theory).

The generation of a preferred$_1$ model, M, of Θ_1 should proceed as follows:

$$\begin{aligned}
M/1 &= \{Alive(1), Loaded(1)\}\ , \\
M/2 &= M/1 \cup \{Alive(2), Loaded(2)\}\ , \\
M/3 &= M/2 \cup \{Alive(3), Loaded(3), Occ(Shoot)(3)\}\ , \\
M/4 &= M/3 \cup \{\neg Alive(4), Loaded(4)\}\ , \\
M/5 &= M/4 \cup \{\neg Alive(5), Loaded(5)\}\ ,\ \ldots\ .
\end{aligned}$$

Thus the only atoms which should be in $M/1$, and so in the evidential context $M(1)$, are those required by the *boundary conditions* of Θ_1, which are stated by Axiom (6). This restriction of the evidential context ("Ockham's razor") is appropriate because prediction should be based on all *and only* the available evidence. Now, *Alive* is not declared to be a theoretical relation in Θ_1, so by the notational convention, $Phys(Alive)$ should be true in M. And it is consistent in the current context $M(1) \cup \emptyset$ (given the background theory $\Theta_1 \setminus \{(6)\}$) to assume $Inert(Alive)(1)$, so this assumption should be added to the conjectural context. Consequently $Alive(2)$ should be in $M(2)$ by the inertia axiom (Axiom (3)). Similarly $Phys(Loaded)$ should be true in M, and it is consistent in the current context $M(1) \cup \{Inert(Alive)(1)\}$ to assume $Inert(Loaded)(1)$, so this assumption should be added to the conjectural context. By inertia, $Loaded(2)$ should be in $M(2)$. And, in accordance with

14. As usual, a *literal* is either an atom, α, or its negation, $\neg\alpha$.

Ockham's razor, no other atoms should be in $M(2)$. By analogous reasoning beginning with the current context $M(2) \cup \emptyset$, *Alive*(3) and *Loaded*(3) should be in $M(3)$, as should the remaining boundary condition *Occ*(*Shoot*)(3). And, by Ockham's razor, no other atoms should be in $M(3)$. Now, in the current context $M(3) \cup \emptyset$ it is consistent to assume either *Succ*(*Shoot*)(3) or *Inert*(*Alive*)(3). However both cannot be assumed; for if they were, then it would follow by the axioms for change ((1) and (5)) and inertia that both *Alive*(4) and $\neg$*Alive*(4) would be in $M(4)$. In keeping with the principle that change is preferred to inertia, *Succ*(*Shoot*)(3) should be assumed and added to the conjectural context, and so $\neg$*Alive*(4) should be in $M(4)$. It is consistent in the current context $M(3) \cup \{Succ(Shoot)(3)\}$ to assume *Inert*(*Loaded*)(4), so by inertia, *Loaded*(4) should be in $M(4)$. And, by Ockham's razor, no further atoms should be in $M(4)$. The remainder of M/∞ should then be generated by repeated applications of the inertia axiom and Ockham's razor.

The example suggests that event theories should be interpreted *chronologically*. This fits naturally with our experience of "time's arrow"; of the asymmetry between the past (which is fixed) and the future (which is open, which is yet to exist). In particular, our understanding of events in terms of dependable regularities of sequence is founded on this asymmetry. The example also suggests that at each successive time point (at each new present moment) we should first fix the evidential context and then generate the appropriate conjectural context. The evolving context and the background theory then produce the expected changes and persistences. Fixing the evidential context consists of *minimizing* it; that is, restricting it to those object and event literals which are required by the boundary conditions or the earlier interpretation of the theory. Generating the conjectural context consists of *maximizing* success and inertia assumptions (that is, assuming them when they are suggested by the evidential context and they are consistent with the current context and the background theory) giving *priority* to the former in case of conflict. The definition of a preferred$_1$ model of an event theory should thus reflect the *prioritized chronological minimaximization* involved in its intended interpretation.

We begin by defining the preference relation $\prec_1$. In this definition (and in the subsequent definition of $\prec_2$) 'fewer' and 'more' should be understood in terms of set inclusion rather than cardinality.[15] In keeping with the above discussion, let an *evidential atom* be either an object atom or an event atom other than a success atom, and let a *conjectural atom* be either a success atom or an inertia atom.

Definition 8 (Preference$_1$) *Let M and M' be models which differ at most on the interpretation of temporally-indexed relations. Then M* is preferred$_1$ to *M' (written $M \prec_1 M'$)*[16] *iff there is a time point t such that M and M' agree before t, and at t:*

1. *fewer evidential atoms are defined in M, and M and M' agree on the truth values of all evidential atoms which are defined in M; or*

2. *M and M' differ only on conjectural atoms, and more success atoms are true in M; or*

15. Thus if $At(T, v, t, M) = \{\alpha : \alpha$ is an atom of type T with truth value v at time t in model $M\}$, then fewer atoms of type T have truth value v at time t in model M than do so in model M' if $At(T, v, t, M) \subset At(T, v, t, M')$. Similarly, but with $\supset$ replacing $\subset$, in the case for 'more'.

16. This way of writing preferences is based on the comparison of evidential contexts.

3. M and M' differ only on inertia atoms, and more inertia atoms are true in M.

For example, suppose that M and M' are models which differ at most on temporally indexed relations. (1) If M and M' agree before time 2 and disagree then only in that $Occ(Shoot)(2)$ is undefined in M and true in M', then M is preferred$_1$ to M' by clause 1 of the definition. (2) If M and M' agree before time 3 and disagree then only in that $Succ(Shoot)(3)$ is true in M and false in M', then M is preferred$_1$ to M' by clause 2 of the definition. (3) If M and M' agree before time 3 and disagree then only in that $Succ(Shoot)(3)$ and $\neg Inert(Alive)(3)$ are true in M whereas $\neg Succ(Shoot)(3)$ and $Inert(Alive)(3)$ are true in M', then M is preferred$_1$ to M' by clause 2 of the definition. (4) If M and M' agree before time 2 and disagree then only in that $Inert(Alive)(2)$ is true in M and false in M', then M is preferred$_1$ to M' by clause 3 of the definition.

The preferred$_1$ models of an event theory Θ should thus be obtained by focussing on the class of all *Phys*-maximal models of Θ and then selecting the $\prec_1$-minimal models of Θ from it. Accordingly, the definition of the preferred$_1$ models of an event theory and the definition of the predictions based on them are instances of the following generic definitions.

Definition 9 (Preferred Models, Prediction) *A model M is said to be a* preferred$_i$ model *of an event theory Θ if M is a Phys-maximal model of Θ and there is no other model M' of Θ which is preferred$_i$ to M (which is such that $M' \prec_i M$). If Θ is an event theory and ϕ is a sentence, then Θ* predicts$_i$ *ϕ (written $\Theta \mathrel{|\!\approx_i} \phi$) iff all preferred$_i$ models of Θ are also models of ϕ.*

In keeping with the discussion in the introduction, the definition of prediction is *cautious*. A clearer picture of this emerges if we consider the preferred$_i$ models of a given event theory at a more abstract level.

Definition 10 (Equivalence, Determinism, Representative Preferred Model) *Let Θ be an event theory, and let M and M' be preferred$_i$ models of Θ. Then M and M' are said to be* preference$_i$ equivalent *(written $M \sim_i M'$) if M and M' agree on the interpretation of all evidential and conjectural atoms.*[17] *An event theory Θ is* deterministic$_i$ *if it has a single $\sim_i$-equivalence class, and is* non-deterministic$_i$ *otherwise. The representative member of each $\sim_i$-equivalence class is called a* representative preferred$_i$ model *of Θ.*

Each $\sim_i$-equivalence class of preferred$_i$ models of an event theory represents a possible history which is defined by the theory. So deterministic$_i$ theories define a single possible history, and predictions can safely be based on it. However, non-deterministic$_i$ theories define more than one possible history, and so caution dictates that their predictions should be restricted to those sentences which are true in all of the possible histories that they define.[18] The representative preferred$_i$ models of an event theory provide a concrete way of thinking about possible histories.

We can now return to the (inductive version of the) Yale Shooting Problem.

17. Any two such models may differ on the interpretation of terms, or on the truth values of atoms not considered in the definition of preference$_i$.

18. It is possible to define a *risky* notion of prediction based on a single $\sim_i$-equivalence class c for event theory Θ. Thus $\Theta \mathrel{|\!\approx^c_i} \phi$ iff ϕ is true in all models in c. This relation can be used to obtain information about a particular possible history, but it does not serve as a basis for reliable prediction in non-deterministic$_i$ theories because it does not take other, equally possible, histories into account.

Example 1 *As before, let $\Theta_1 = \Theta_{Ind} \cup \{(4), (5), (6)\}$. Then Θ_1 is deterministic$_1$. The evidential literals which are true in its representative preferred$_1$ model agree with those in the set M/∞ discussed earlier. Thus Θ_1 predicts$_1$ that the Shoot event succeeds at time 3, with the effect that Fred is not alive at time 4.*

Proposition $\Theta_1 \approx_1 \mathit{Succ}(\mathit{Shoot})(3) \wedge \neg \mathit{Alive}(4)$.

Proof *By Definition 9, it is sufficient to prove that the conclusion follows in all preferred$_1$ models of Θ_1. So, let M be a preferred$_1$ model of Θ_1. Then, by Definition 9, $\mathit{Phys}(\mathit{Alive})$ and $\mathit{Phys}(\mathit{Loaded})$ are true in M.*[19] *By Axiom (6), $\mathit{Alive}(1)$ and $\mathit{Loaded}(1)$ are both true in M. By Definition 8.3, $\mathit{Inert}(\mathit{Alive})(1)$ is true in M.*[20] *So it follows by Axiom (3) that $\mathit{Alive}(2)$ is true in M. By similar reasoning, $\mathit{Inert}(\mathit{Loaded})(1)$ and $\mathit{Loaded}(2)$ are also true in M (Axiom (3), Definition 8.3). As $\mathit{Alive}(2)$ is true in M, it follows by inertia (Axiom (3), Definition 8.3) that $\mathit{Alive}(3)$ is true in M. Similarly, as $\mathit{Loaded}(2)$ is true in M, it follows by inertia that $\mathit{Loaded}(3)$ is true in M. So, by Axiom (4), $\mathit{Pre}(\mathit{Shoot})(3)$ is true in M. By Axiom (6), the occurs atom $\mathit{Occ}(\mathit{Shoot})(3)$ is true in M. By Definition 8.2, $\mathit{Succ}(\mathit{Shoot})(3)$ is true in M. It follows, by Axiom (1), that $\mathit{Eff}(\mathit{Shoot})(4)$ is true in M, and so, by Axiom (5), $\neg\mathit{Alive}(4)$ is true in M.*

The Yale Shooting Problem is of interest because, as Hanks and McDermott (1987) show, it poses problems for theories which do not take account of time's arrow. The example suggests that reasoning about inertia should be chronological. A related example involving reasoning about change was suggested by Lifschitz (1987, p. 37). The point of his example can be illustrated by adding a second shot to the Yale Shooting Problem. Let $\Theta_1' = \Theta_1 \cup \{\mathit{Occ}(\mathit{Shoot})(4)\}$. Then we expect that, as before, the first shot will succeed and that the second shot will fail (because Fred is no longer alive when the second shot occurs). And, indeed, this is what transpires in all preferred$_1$ models of Θ_1'. However, if Θ_1' were not interpreted chronologically, then there would be preferred models of it in which the second shot succeeds and the first shot fails; as the success of the second shot requires that Fred is alive at time 4.

The inductive version of the Yale Shooting Problem considered here (in which the *Shoot* event is treated inductively rather than deductively) also illustrates the need to give priority to change (*Succ* assumptions) over inertia (*Inert* assumptions) in case of conflict. Without

19. As M is a preferred$_1$ model of Θ_1 it follows by Definition 9 that M is a *Phys*-maximal model of Θ_1. So, if $\mathit{Phys}(\mathit{Alive})$ were not true in M, then there would be a model M' of Θ_1 in which $\mathit{Phys}(\mathit{Alive})$ and all of the *Phys* atoms which are true in M were true. But then M would not be a *Phys*-maximal model of of Θ_1, contradicting the assumption that it is. An analogous argument justifies all subsequent appeals to Definition 9 regarding the relation *Phys*.

20. If $\mathit{Inert}(\mathit{Alive})(1)$ were not true in M, then there would be a model M' of Θ_1 in which $\mathit{Inert}(\mathit{Alive})(1)$ was true and which was therefore preferred$_1$ to M on this basis by clause 3 of Definition 8 at time 1. (M and M' would disagree at most on the interpretation of temporally-indexed relations, M and M' would agree on the interpretation of all temporally-indexed relations at all time points before time 1, M and M' would agree on the interpretation all evidential and success atoms at time 1, and in M' more inertia atoms with temporal index 1 (all of those which are true in M together with $\mathit{Inert}(\mathit{Alive})(1)$) would be true.) But then it would follow by Definition 9 that M would not be a preferred$_1$ model of Θ_1, contradicting the assumption that it is. An analogous argument justifies all subsequent appeals to some clause n of the definition of preference$_i$ regarding the truth value of some atom with temporal index t.

this requirement there would be preferred models of Θ_1 in which $\mathit{Inert}(\mathit{Alive})(3)$ is true and $\mathit{Succ}(\mathit{Shoot})(3)$ is false, and so, contrary to expectation, Fred remains alive at time 4.

In subsequent examples it will often be assumed that different names (whether constants or functional expressions) denote different individuals. In order to enforce this convention, *uniqueness of names* axioms (Lifschitz, 1987, p. 50) are used. Let $f_1, \ldots, f_n$ be functions returning values of the same sort, and let $x_1, \ldots, y_1, \ldots$ be variables of the appropriate sorts. Then $U[f_1, \ldots, f_n]$ is the conjunction of the axioms in the set:

$$\{\forall x_1, \ldots, x_k, y_1, \ldots, y_l \neg f_i(x_1, \ldots, x_k) = f_j(y_1, \ldots, y_l) : 1 \leq i < j \leq n\} \cup$$
$$\{\forall x_1, \ldots, x_k, y_1, \ldots, y_k (f_i(x_1, \ldots, x_k) = f_i(y_1, \ldots, y_k) \rightarrow (x_1 = y_1 \wedge \ldots \wedge x_k = y_k)) : 1 \leq i \leq n\} \ .$$

These axioms express the fact that the functions $f_1, \ldots, f_n$ are injections with different ranges. This notation is extended to constants by treating them as 0-ary functions. Thus, for example, given $U[A, B, L1, L2]$ and $U[\mathit{Move}]$, it follows that the constants A, B, etc., denote different objects, and that the functional expressions $\mathit{Move}(A, L1, L2)$ and $\mathit{Move}(B, L1, L2)$ denote different events.

The next example illustrates the need to restrict the inertia axiom to physical relations.

Example 2 *Block B is moved from location $L1$ to location $L2$. We expect that $L1$ will be clear as a result. This example can be represented as follows:*

$$\forall x, l, l', t(\mathit{Pre}(\mathit{Move}(x, l, l'))(t) \equiv \mathit{At}(x, l)(t)) \tag{7}$$
$$\forall x, l, l', t(\mathit{Eff}(\mathit{Move}(x, l, l'))(t) \equiv (\mathit{At}(x, l')(t) \wedge \neg \mathit{At}(x, l)(t))) \tag{8}$$
$$\forall l, t(\mathit{Clear}(l)(t) \equiv \neg \mathsf{T} \exists x \mathit{At}(x, l)(t)) \tag{9}$$
$$\mathit{Theo}(\mathit{Clear}) \tag{10}$$
$$U[B, L1, L2] \wedge \mathit{At}(B, L1)(1) \wedge \mathit{Occ}(\mathit{Move}(B, L1, L2))(1) \tag{11}$$

Axioms (7) and (8) define the preconditions and effects of move events. Axiom (9) defines a location to be clear if it is not true that there exists a block which is at that location. The use of the truth operator in this definition allows for the fact that the At relation may be partial; a location is considered to be clear if none of the blocks whose locations are defined are at the location. Axiom (10) declares Clear to be a theoretical relation. Finally, Axiom (11) states the boundary conditions.

Let $\Theta_2 = \Theta_{\mathit{Ind}} \cup \{(7), \ldots, (11)\}$. Then Θ_2 has a single representative preferred$_1$ model in which $\mathit{Clear}(L1)(2)$ is true; because, in the model, the move event succeeds, thereby vacating $L1$, and no other object replaces B at $L1$. However, non-replacement at $L1$ depends on the fact that Clear is a theoretical relation, and so is exempt from the law of inertia. If $\Theta_2' = \Theta_2 \setminus \{(10)\}$, then in every preferred$_1$ model of Θ_2' an object mysteriously replaces B at $L1$.

Proposition $\Theta_2 \approx\!\!\!\!|_1 \mathit{Clear}(L1)(2)$*, but* $\Theta_2' \approx\!\!\!\!|_1 \neg \mathit{Clear}(L1)(2)$.

Proof *For the first part, let M be a preferred$_1$ model of Θ_2. Then, by axioms (7) and (11), $\mathit{At}(B, L1)(1)$, $\mathit{Pre}(\mathit{Move}(B, L1, L2))(1)$, and $\mathit{Occ}(\mathit{Move}(B, L1, L2))(1)$ are all true in M.*

By Definition 8.2, $Succ(Move(B, L1, L2))(1)$ *is true in* M*. So it follows, by axioms (1) and (8) that* $At(B, L2)(2)$ *and* $\neg At(B, L1)(2)$ *are both true in* M*. As* $At(B, L1)(1)$ *is true in* M*, it follows by Axiom (9) that* $\neg Clear(L1)(1)$ *is true in* M*. By Axiom (10)* $Theo(Clear)$ *is true in* M *and so, by Axiom (2),* $\neg Phys(Clear)$ *is true in* M*. It follows by Axiom (3) that* $\neg Inert(Clear, \langle L1 \rangle)(1)$ *is true in* M *(consequently* $\neg Clear(L1)(2)$ *is no longer true in* M *by inertia). By Definition 8.1 it follows that, for any* x *other than* B*,* $At(x, L1)(2)$ *is undefined in* M*. So, as* $\neg At(B, L1)(2)$ *is true in* M*, it follows that for any* x*,* $\neg \mathsf{T} \exists x At(x, L1)(2)$ *is true in* M*. And so it follows by Axiom (9) that* $Clear(L1)(2)$ *is true in* M*.*

For the second part, let M *be a preferred*$_1$ *model of* Θ_2'*. Then, as before, the atoms* $At(B, L1)(1)$*,* $\neg Clear(L1)(1)$*,* $Succ(Move(B, L1, L2))(1)$*,* $At(B, L2)(2)$ *and* $\neg At(B, L1)(2)$ *are all true in* M*. However (in the absence of Axiom (10)) it follows by Definition 9 that* $Phys(Clear)$ *is true in* M*. By Definition 8.3,* $Inert(Clear, \langle L1 \rangle)(1)$ *is true in* M*. So it follows by Axiom (3) that* $\neg Clear(L1)(2)$ *is true in* M*.*

The restriction of the inertia axiom can be justified in terms of the physical-theoretical distinction as follows. The law of inertia is a law of *physical* inertia; its task is to represent the persistence of those physical facts which are not changed by events. Applying it to theoretical relations (as in Θ_2') results in mysterious consequences; which arise because maintaining the inertia of a theoretical fact ($\neg Clear(L1)$) introduces an additional real change in a physical fact (a change in the At relation). Moreover, as theoretical facts supervene on physical facts, changes (persistences) in theoretical facts supervene on changes (persistences) in physical facts. So it is sufficient to represent changes (persistences) in physical facts, and to let changes (persistences) in theoretical facts "take care of themselves". This is done in the case of Θ_2, where a change in the At relation results in a change in the *Clear* relation. In more complex examples, several blocks may be moved to or from a location simultaneously, and each move event may or may not succeed. In such cases, the axioms for change and inertia represent changes and persistences in the At relation, and changes (persistences) in the *Clear* relation take care of themselves once the new At facts are fixed; once the real changes have occurred and the dust has settled.[21]

21. Another, more artificial, example involves the interaction between Goodman's (1954, p. III.4) predicate *Grue* and common sense inertia. Call an object "grue" if it is green at time t and t is before time 2 or it is blue thereafter. Now suppose that object O is green at time 1 and that we don't know of any events which occur at time 1 which affect O. Then it seems natural to conclude by inertia that O is green at time 2. However, as O is green at time 1, O is also grue at time 1, so it is equally reasonable to predict that O will be grue (that is, blue) at time 2. The example can be represented formally by the theory Θ_G, which consists of Θ_{Ind} together with the following axioms:

$$\forall x, t(Grue(x)(t) \equiv ((t < 2 \wedge Green(x)(t)) \vee (t \geq 2 \wedge Blue(x)(t)))) \ ,$$
$$\forall x, t(Green(x)(t) \equiv \neg Blue(x)(t)) \ ,$$
$$Green(O)(1) \ .$$

Then there are intended preferred$_1$ models of Θ_G in which $Green(O)(2)$ is true, and unintended preferred$_1$ models of Θ_G in which $Blue(O)(2)$ is true. This problem can be solved by declaring that *Grue* is a theoretical predicate (in keeping with Goodman's doctrine of entrenchment and Quine's advocation of similarity), and so should not be projected by the law of inertia.

4. Primary and Secondary Events

The theory of inductive events provides the basis of an integrated solution to the qualification problem and the frame problem. On the intended interpretation of the success axiom, events are, given their preconditions, normally sufficient for their effects. On the intended interpretation of the inertia axiom, physical facts which are not affected by events persist. However, whereas the effects of successful inductive events are certain and invariable, the effects of natural events may be uncertain and they, or some of them at least, may vary according to the context in which the events occur.

It may seem that context-dependent effects, or *ramifications*, can be represented by domain axioms. However, the following example, which is based on Lifschitz's (1990) lamp-circuit example and Baker's (1991) ice-cream example, shows that this approach is too simplistic.

Example 3 *Ollie is at location $L1$, he is holding block B, and he moves to location $L2$. We expect that, as a result, Ollie will reach $L2$. Moreover, as Ollie is holding the block when he moves, we expect that it will move with him to $L2$.*

It may seem that this example can be represented by the event theory $\Theta_3 = \Theta_{Ind} \cup \{(12), \ldots, (16)\}$; where:

$$\forall x, l, l', t(Pre(Move(x, l, l'))(t) \equiv At(x, l)(t)) \tag{12}$$

$$\forall x, l, l', t(Eff(Move(x, l, l'))(t) \equiv At(x, l')(t)) \tag{13}$$

$$\forall x, l, l', t((At(x, l)(t) \wedge \neg l = l') \rightarrow \neg At(x, l')(t)) \tag{14}$$

$$\forall x, y, l, t((At(x, l)(t) \wedge Holding(x, y)(t)) \rightarrow At(y, l)(t)) \tag{15}$$

$$U[O, B, L1, L2] \wedge At(O, L1)(1) \wedge Holding(O, B)(1) \wedge Occ(Move(O, L1, L2))(1) \tag{16}$$

Thus the effects of Move have been simplified. The fact that the moved object is no longer where it was is now to be inferred from Axiom (14), which states that no object can be at two different locations simultaneously, together with the appropriate inequality. The intention is to use Axiom (15) to infer that Ollie's movement results in the movement of block B. For, given that Ollie is holding B when he gets to $L2$, it follows from the axiom that B is at $L2$ also; and so, by axioms (14) and (16), B is not at $L1$.

However, there are two representative preferred$_1$ models of Θ_3, which can be partially described as follows:

$$M_1 \supseteq \{At(O, L2)(2), Holding(O, B)(2), At(B, L2)(2)\},$$
$$M_2 \supseteq \{At(O, L2)(2), At(B, L1)(2)\}.$$

As change is preferred to inertia, Ollie succeeds in moving to $L2$ in both models. In M_1, the fact that Ollie is holding B is inert, and so it follows that the block moves with him to $L2$ as expected. In M_2, the fact that B is at $L1$ is inert, and so, contrary to expectation, B remains at $L1$.

One reaction to this failure is to seek to strengthen axioms such as Axiom (15) by making them "causally" directed, so that they can be interpreted "causally" (positively, as in M_1), rather than declaratively (positively as in M_1, or contrapositively as in M_2). However it seems to me that this response (which is discussed further in Section 7) is mistaken because it misdiagnoses the problem; taking it to be a logical problem rather than a representational one.

Let us consider the problem posed by the example afresh. The theory Θ_3 has two representative preferred models only one of which corresponds to our expectation. What accounts for the asymmetry in our expectation, and how is the formal symmetry to be broken?

The intended positive interpretation of Axiom (15) depends on appropriate reasoning about inertia, and in particular on the appropriate use of the inertia axiom; it is necessary to conclude that Ollie keeps hold of the block, rather than concluding that it remains at $L1$. But it seems odd to be using the inertia axiom when reasoning about change; to be using the inertia axiom (together with Axiom (15)) to get a block to move without there being an event which causes it to move. In doing so we are violating the fundamental intuition which underlies the fact-event distinction; which has it that events are the only causes of physical change.

This consideration provides the key to the correct understanding of the problem. We expect that the block will move when Ollie does because we are told that he is holding the block when he moves and are not told that he releases it. Consequently we discount the possibility of the block remaining at $L1$ because an additional event would be required in order to account for this. However the movement of the block is itself an additional event which has the additional effect that the block is at $L2$. The missing, symmetry-breaking, causal element in this example is thus an event, and the choice between a move event and a release event seems clear. But note that the block's moving differs from Ollie's moving. The block moves only because Ollie moves and only because he is holding it when he does so.

In order to reflect this difference, a distinction is drawn between *primary* and *secondary* events. Primary and secondary events are inductive events of the kind that we have been considering so far. However while primary events occur independently, secondary events are *invoked* (in the non-mystical Computer Science sense in which one program (procedure, process, ...) is said to invoke another) by other events, and are *causally dependent* on them in the sense that a secondary event can succeed only if it is invoked invoked by a successful event.[22] Given this distinction, ramifications can be represented by invoking appropriate secondary events in appropriate contexts. Thus, in Example 3, Ollie's moving can be represented as a primary event which, because he is holding block B when he moves, invokes the secondary move-B event. The move-B event occurs because it is invoked, and it succeeds only if the move-Ollie event does. Note that secondary events may, in turn, invoke other events which are causally dependent on them. For instance, if in the current example block B' is placed on top of block B, then the invoked move-B event should in turn invoke a move-B' event. There can thus be tertiary events, and events of ever higher order.

22. This condition should perhaps be called "success dependence" or "effect dependence" in order to distinguish it from counterfactual dependence (Lewis, 1986, Ch. 21).

Table 3: The theory of invocation, Θ_{Inv}

$$\forall e, e', t(Inv_1(e, e')(t) \rightarrow (Occ(e)(t) \wedge Occ(e')(t))) \tag{17}$$
$$\forall e, t((Succ(e)(t) \wedge \exists e' Inv_1(e', e)(t)) \rightarrow \exists e'(Inv_1(e', e)(t) \wedge Succ(e')(t))) \tag{18}$$
$$\forall e, e', t(Inv(e, e')(t) \equiv (Inv_1(e, e')(t) \vee \exists e''(Inv_1(e, e'')(t) \wedge Inv(e'', e')(t)))) \tag{19}$$
$$\forall e, t\, \neg \mathsf{T} Inv(e, e)(t) \tag{20}$$

However, for the sake of convenience, all invoked events will be referred to as 'secondary events'.[23]

Invocations are represented in $\mathcal{EL}$ by *invocation atoms*. An invocation atom is an event atom of the form $Inv_1(e, e')(t)$, which states that event e directly invokes event e' at time t. Secondary events are typically invoked by *invocation axioms* of the form:

$$\forall e, e', t((Occ(e)(t) \wedge \phi) \rightarrow Inv_1(e, e')(t)) \ ;$$

where ϕ is a formula which distinguishes those contexts in which e invokes e'. The properties of secondary events are stated by axioms (17)-(20) given in Table 3. Axiom (17) requires that both the invoking and the invoked events occur. Axiom (18) represents (causal) dependence, and states that a secondary event succeeds only if one of the events which invoked it succeeds. The axiom is stated in this way in order to allow for cases in which a secondary event is invoked by more than one event. Axioms (19) and (20) ensure that invocation is acyclic. This is achieved by defining the auxiliary (indirect invocation) relation Inv to be the transitive closure of the (direct invocation) relation Inv_1 (Axiom (19)) and requiring that Inv is irreflexive (Axiom (20)).[24]

Events which invoke others can be thought of in two ways: as elementary inductive events, or as complex events which have a causal structure. The *invocation graph* for an

23. Primary and secondary events are so called by loose analogy with the philosophical distinction between primary and secondary qualities. Primary qualities (such as size, shape and motion) are the fundamental qualities used in science. By contrast, secondary qualities are sensory qualities (such as colour, taste, smell, felt warmth or texture, and sound) which exist only in certain contexts (to individual observers under specific conditions) and which are causally dependent on primary qualities. Similarly, tertiary qualities are qualities which an object has in virtue of its secondary qualities; for example a flower may be attractive to a butterfly because of its colour, or a wine may be expensive because of its taste (Blackburne, 1994).

24. The axiomatization of secondary events is intended to be minimal. For example, there is no prohibition on an event occurring as both a primary event and as a secondary event. If this were to happen, then it would follow from Axiom (18) that the event would have secondary status. In certain circumstances it may be desirable to define the order of an event, and to require that each event has exactly one order. This can be done by adding axioms such as the following:

$$\begin{aligned} \forall e, t(Ord(e, n)(t) \equiv ((n = 1 \wedge Occ(e)(t) \wedge \neg \mathsf{T} \exists e' Inv_1(e', e)(t)) \\ \vee\ (n > 1 \wedge \exists e'(Ord(e', n{-}1)(t) \wedge Inv_1(e', e)(t))))) \ , \\ \forall e, n, m, t((Ord(e, n)(t) \wedge \neg n = m) \rightarrow \neg \mathsf{T} Ord(e, m)(t)) \ . \end{aligned}$$

event e (at some time point t) is a directed acyclic graph whose initial vertex is e, whose remaining vertices are the events invoked by e (either directly or indirectly), and whose edges represent the direct invocation relation. Event e's *success graph* (at t) is the subgraph of e's invocation graph which consists of all those chains in the invocation graph which begin with e and which consist entirely of successful events. The *direct* effects of e are its defined (invariant) effects. The *indirect* effects of e are the effects of all other events in its success subgraph. So, in particular, successful inductive events can be thought of as having all of the effects of the successful secondary events that they invoke (either directly or indirectly). Thus, in the working example, the move-Ollie event has the direct effect that he moves, and can be thought of as having the indirect context-dependent effects that B and B' move with him. Note that the effects of an event may now be context-dependent in two ways. An event may invoke different events in different contexts (recall that invocation axioms may be context-dependent), and so its invocation graph may vary according to the context in which it occurs. Moreover, the same invocation graph may result in different success graphs if the context varies in other ways. For instance, in the working example, move-B' may be invoked in two different contexts (in both of which B' is on B), and succeed in one (the one we have been contemplating) but fail in another (say because B is also being held by another agent).

Definition 11 *The* theory of invocation*,* Θ_{Inv}*, consists of the axioms given in Table 3; thus* $\Theta_{\mathit{Inv}} = \{(17), \ldots, (20)\}$.

As invocation atoms are event atoms, the pragmatics given in the last section can be used without change.

We can now give a formal version of an extension of the block-carrying example.

Example 4 *Ollie is at location* $L1$*. He is holding block* $B1$*, block* $B2$ *is stacked on* $B1$*, and block* $B3$ *is stacked on* $B2$*. Ollie moves to location* $L2$*. We expect that the move event will succeed and that the stack of blocks will move with him. However if, for some independent reason,* $B2$ *does not move, then we expect that* $B3$ *will also remain at* $L1$*.*

The first part of this example can be represented by the event theory $\Theta_4 = \Theta_{\mathit{Ind}} \cup \Theta_{\mathit{Inv}} \cup \{(12), \ldots, (15)\} \cup \{(21), \ldots, (24)\}$*, where:*

$$\forall x, y, l, t((\mathit{On}(x, y)(t) \wedge \mathit{At}(y, l)(t)) \rightarrow \mathit{At}(x, l)(t)) \tag{21}$$

$$\forall x, y, l, l', t((\mathit{Occ}(\mathit{Move}(x, l, l'))(t) \wedge \mathit{Holding}(x, y)(t)) \rightarrow \\ \mathit{Inv}_1(\mathit{Move}(x, l, l'), \mathit{Move}(y, l, l'))(t)) \tag{22}$$

$$\forall x, y, l, l', t((\mathit{Occ}(\mathit{Move}(x, l, l'))(t) \wedge \mathit{On}(y, x)(t)) \rightarrow \\ \mathit{Inv}_1(\mathit{Move}(x, l, l'), \mathit{Move}(y, l, l'))(t)) \tag{23}$$

$$U[O, B1, B2, L1, L2] \wedge U[\mathit{Move}] \wedge \mathit{At}(O, L1)(1) \wedge \mathit{Holding}(O, B1)(1) \\ \wedge\ \mathit{On}(B2, B1)(1) \wedge \mathit{On}(B3, B2)(1) \wedge \mathit{Occ}(\mathit{Move}(O, L1, L2))(1) \tag{24}$$

Axiom (21) states that if object x *is on object* y *then* x *is at the same location as* y *is, while axioms (22) and (23) are invocation axioms representing ramifications. Axiom (22) states that a move-*x *event invokes a move-*y *event in contexts in which* x *is holding* y*. Similarly,* *Axiom (23) states that a move-x event invokes a move-y event in contexts in which y is on x.*

There is a single representative preferred$_1$ *model of* Θ_4*, in which Ollie's movement successfully invokes the movement of* $B1$ *(because Ollie is holding* $B1$ *when he moves), this in turn successfully invokes the movement of* $B2$ *(because* $B2$ *is on* $B1$ *when* $B1$ *moves), and this in turn successfully invokes the movement of* $B3$ *(because* $B3$ *is on* $B2$ *when* $B2$ *moves). Thus, in accordance with expectation,* Θ_4 *predicts that Ollie succeeds in moving to* $L2$ *with the entire stack of blocks. Note that the success of each invoked event depends on the success of the event which invoked it. Thus, for example, if* Θ_4 *is extended such that* $B2$ *remains at* $L1$*, then the extended theory predicts that* $B3$ *also remains at* $L1$*.*

Proposition $\Theta_4 \mathrel{|\!\!\approx}_1 At(B3, L2)(2)$*, and* $\Theta_4 \cup \{At(B2, L1)(2)\} \mathrel{|\!\!\approx}_1 At(B3, L1)(2)$*.*

Proof *For the first part, let* M *be a preferred*$_1$ *model of* Θ_4*. Then* $Occ(Move(O, L1, L2))(1)$ *and* $Holding(O, B1)(1)$ *are true in* M *(Axiom (24)). So it follows (axioms (17), (22)) that* $Inv_1(Move(O, L1, L2), Move(B1, L1, L2))(1)$ *and* $Occ(Move(B1, L1, L2))(1)$ *are true in* M*. So, as* $On(B2, B1)(1)$ *is true in* M *(Axiom (24)), it follows (axioms (17), (23)), that* $Inv_1(Move(B1, L1, L2), Move(B2, L1, L2))(1)$ *and* $Occ(Move(B2, L1, L2))(1)$ *are true in* M*. And so, as* $On(B3, B2)(1)$ *is true in* M *(Axiom (24)), it follows (axioms (17), (23)) that* $Inv_1(Move(B2, L1, L2), Move(B3, L1, L2))(1)$ *and* $Occ(Move(B3, L1, L2)(1)$ *are true in* M*. No further invocation atoms with temporal index* 1 *are defined in* M *(Definition 8.1), so the four events* $Move(O, L1, L2)$*,* $Move(B1, L1, L2)$*,* $Move(B2, L1, L2)$*,* $Move(B3, L1, L2)$*, which occur at time* 1 *in* M*, are linked by a chain of invocations at time* 1 *in* M*.*[25] *Moreover, the preconditions of each of the four events are true in* M *(axioms (12), (15), (21), (24)). As* $Move(O, L1, L2)$ *is the only primary event at time* 1 *in* M*, it follows (Definition 8.2) that* $Succ(Move(O, L1, L2))(1)$ *is true in* M*. Moreover, as* $Move(B1, L1, L2)$ *is directly invoked by a successful event at time* 1 *in* M*, it follows (Definition 8.2) that* $Succ(Move(B1, L1, L2))(1)$ *is true in* M*. Similar reasoning shows that success is propagated down the rest of the invocation chain; that, in turn,* $Succ(Move(B2, L1, L2))(1)$ *and hence* $Succ(Move(B3, L1, L2))(1)$ *are true in* M*. So, as* $Succ(Move(B3, L1, L2))(1)$ *is true in* M*, it follows (axioms (1), (13)) that* $At(B3, L2)(2)$ *is true in* M*.*

For the second part, let M *be a preferred*$_1$ *model of* $\Theta'_4 = \Theta_4 \cup \{At(B2, L1)(2)\}$*. Then, as above, the events* $Move(O, L1, L2)$*,* $Move(B1, L1, L2)$*,* $Move(B2, L1, L2)$*, and* $Move(B3, L1, L2)$ *all occur at time* 1 *in* M*, their preconditions are all true at time* 1 *in* M*, and they are all linked by a unique invocation chain at time* 1 *in* M*. As above, the first two events in the chain succeed. However, as* M *is a model of* Θ'_4*,* $At(B2, L1)(2)$ *is true in* M*. So it follows (axioms (14), (24)) that* $\neg At(B2, L2)(2)$ *is true in* M*. It therefore follows (axioms (1) and (13)) that* $\neg Succ(Move(B2, L1, L2))(1)$ *is true in* M*. So, as* $Move(B2, L1, L2)$ *is the only event which directly invokes* $Move(B3, L1, L2)$ *at time* 1 *in* M*, it follows that* $\neg\mathsf{T}\exists e(Inv_1(e, Move(B3, L1, L2))(1) \wedge Succ(e)(1))$ *is true in* M*. And so* $\neg\mathsf{T}(Succ(Move(B3, L1, L2))(1) \wedge \exists e\, Inv_1(e, Move(B3, L1, L2))(1))$ *is true in* M *(Axiom (18)). As* $Inv_1(Move(B2, L1, L2), Move(B3, L1, L2))(1)$ *is true in* M *and the* $Succ$

25. The unique-names axiom for *Move* in Axiom (24) ensures that the events are distinct. In view of axioms (19) and (20) this appeal to the unique-names axiom is not strictly necessary. However, in this and in subsequent examples involving multiple events it is simpler to add unique-names axioms for events, and then assume in the proofs that distinct event terms denote distinct events.

relation is classical (Axiom (1)), it follows that $\neg Succ(Move(B3, L1, L2))(1)$ *is true in* M*. However,* $Phys(At)$ *and* $At(B3, L1)(1)$ *are true in* M *(axioms (15), (21), (24), Definition 9), so it follows by inertia (Axiom (3), Definition 8.3) that* $At(B3, L1)(2)$ *is true in* M*.*

Note that with the introduction of secondary events as the missing causal elements in this example, the inertia axiom and the domain axioms ((14), (15) and (21)) are confined to their proper tasks; namely, representing inertia, and defining or constraining the At relation respectively.

In philosophical terms, invocations provide a means for representing contemporaneous sufficient causation between events. Thus if event e directly invokes event e' at time t, then e is a sufficient cause of e' at t. The direct invocation relation, Inv_1, thus represents causal directedness (causal priority) between contemporaneous events. Axioms (17) and (18) respectively ensure that contemporaneous causation occurs between actual events, and that an invoked event is efficacious only if it is invoked by an efficacious event. This account of contemporaneous sufficient causation can be thought of in terms of regularities, however the reasoning is now more complex; for example, we might discover that event e directly invokes event e' by noting that whenever e occurs and condition ϕ is true e' also occurs, and that whenever this is the case e' succeeds only if e does.

Contemporaneous (sufficient) causation between events is naturally required to be asymmetric (axioms (19) and (20)). However, it has been suggested that there are also cases of symmetric contemporaneous causation between events. Taylor (1975) gives the example of a locomotive and a caboose which are coupled together in such a way that the locomotive moves iff the caboose does. An analogous example, suggested by Denecker, Dupré, and Belleghem (1998), involves a pair of gears which are interlocked, so that each gear rotates iff the other does.

As will become clear, it is better to view these as cases involving *symmetric constraints* on contemporaneous events, rather than symmetric causation between them. Now, clearly, symmetric constraints cannot be represented as invocations.[26] However, they can be represented on an individual basis by adding particular axioms. For example, the symmetric constraint on the rotation of the gears can be represented by the following axiom:

$$\begin{aligned}\forall g, g', t(Intl(g, g')(t) \rightarrow ((Occ(Rot(g))(t) \equiv Occ(Rot(g'))(t))\\ \wedge (Succ(Rot(g))(t) \equiv Succ(Rot(g'))(t)))) \ .\end{aligned}$$

The first conjunct of the consequent of this axiom is required in order to ensure that the rotations of pairs of interlocked gears co-occur; without it, it would be possible for one of the rotate events to occur and fail without the other occurring. Note also that, when introducing the problem, Denecker et al. (1998, p. 34) require that the representation of the behaviour of the interlocked gears should not be such that they can rotate spontaneously. This can present a problem for theories based on classical logic; because of the semantics of

26. An attempt to do so in the case of the gears would be to use the following axiom:

$$\forall g, g', t((Occ(Rot(g))(t) \wedge Intl(g, g')(t)) \rightarrow Inv_1(Rot(g), Rot(g'))(t)) \ .$$

But clearly if $Intl(g, g')(t)$, $Intl(g', g)(t)$ and $Occ(e)(t)$ hold, then a contradiction results by the above axiom and axioms (17), (19) and (20).

Table 4: The theory of symmetrically constrained events, Θ_{SCon}

$$\forall e, t\, \neg \mathsf{T} SCon(e, e)(t) \tag{25}$$

$$\forall e, e', t(SCon(e, e')(t) \rightarrow SCon(e', e)(t)) \tag{26}$$

$$\forall e, e', t(SCon(e, e')(t) \rightarrow ((Occ(e)(t) \equiv Occ(e')(t)) \wedge (Succ(e)(t) \equiv Succ(e')(t)))) \tag{27}$$

$$\begin{aligned} &\forall e, e', t(SCSet(e, e')(t) \equiv \\ &\qquad (SCon(e, e')(t) \vee \exists e''(SCon(e, e'')(t) \wedge SCSet(e'', e')(t)))) \end{aligned} \tag{28}$$

$$\begin{aligned} &\forall e, e', t(EInv(e, e')(t) \equiv \\ &\qquad (Inv_1(e, e')(t) \wedge \exists e''(SCSet(e', e'')(t) \wedge \neg \mathsf{T} SCSet(e, e'')(t)))) \end{aligned} \tag{29}$$

$$\begin{aligned} &\forall e, e', n, t(IPath(e, e', n)(t) \equiv \\ &\qquad ((n = 0 \wedge e = e' \wedge \exists e'' EInv(e'', e)(t)) \\ &\qquad\quad \vee (n > 0 \wedge \exists e''(IPath(e, e'', n-1)(t) \wedge SCon(e'', e')(t))))) \end{aligned} \tag{30}$$

$$\begin{aligned} &\forall e, e', t((SCon(e, e')(t) \\ &\quad \wedge \exists e'', n(IPath(e'', e, n)(t) \\ &\qquad \wedge \forall e''', m(IPath(e''', e', m)(t) \rightarrow m > n)) \rightarrow \; Inv_1(e, e')(t))) \end{aligned} \tag{31}$$

classical biconditionals. However this potential pitfall is effortlessly avoided in event theories because of the accurate representation of partiality; for if only $Intl(g, g')(t)$ is given, then $Occ(Rot(g))(t)$ and $Occ(Rot(g'))(t)$ are both undefined.

A more general treatment of symmetric constraints on events can be given by introducing the symmetric constraint relation *SCon*. This relation represents the co-occurrence of and co-dependence between pairs of events. Accordingly we expect *SCon* to be irreflexive, symmetric, and to hold between pairs of co-occurring events which are also co-dependent. These properties are stated by axioms (25)-(27) in Table 4.

When considered in isolation, the fact that two events are symmetrically constrained provides no compelling evidence that either is the cause of the other. For example, in the gears case, if all we know is that g and g' are interlocked and that g is rotating, then it is reasonable to conclude that g' is also rotating, but it is not reasonable to conclude that the rotation of g is the cause of the rotation of g' or vice versa. Indeed, the two events may not even be causally connected; each of the gears might be rotating because the shaft that it is attached to is being driven.

However, if we have additional, external, information about the direction of causation, then we can use it to infer the direction of causation between pairs of symmetrically constrained events.[27] For example, if we know that gear g is being driven (say by rotating the shaft that it is attached to), and we don't know that gear g' is being driven, then it is reasonable to conclude that the driving of g is causally prior to the rotation of g, and that the rotation of g is in turn causally prior to the rotation of g'. More formally, if we have just $Inv_1(Drive(g), Rot(g))(t)$ and $SCon(Rot(g), Rot(g'))(t)$, then it seems reasonable

27. Taylor makes a similar appeal to an external cause in his discussion of the locomotive-caboose example.

to conclude that $Inv_1(Rot(g), Rot(g'))(t)$. The external invocation chain is thus extended across the symmetric constraint link. The construction of invocation chains of this kind is defined by axioms (28)-(31) given in Table 4.

Axiom (28) states that, at time t, events e and e' are in the same *symmetrically constrained set* iff they are symmetrically constrained or they both occur in a chain of symmetric constrained events.

Axiom (29) defines the conditions for an *external invocation* of a symmetrically constrained event. Thus, at time t, event e externally invokes event e' iff e invokes e' and e' is in a symmetric constraint set which does not include e.

An *invocation path* in a symmetrically constrained set begins with an externally invoked event and consists of a chain of events each of which symmetrically constrains its neighbour. The length of an invocation path is determined by the number of links that it contains; so that an invocation path consisting only of an externally invoked event has length 0, one consisting of such an event and its neighbour has length 1, etc. Accordingly, Axiom (30) states the conditions under which there is an invocation path of length n between events e and e'.

Finally, Axiom (31) defines invocation paths in symmetrically constrained sets. It does so by requiring that an invocation link exist between symmetrically constrained events e and e' whenever the shortest invocation path to e in their symmetrically constrained set is shorter than the shortest invocation path to e' in that set.

Definition 12 *The* theory of symmetric constraints, Θ_{SCon}, *consists of the axioms given in Table 4; thus* $\Theta_{SCon} = \{(25), \ldots, (31)\}$.

These ideas are illustrated by the following elaboration of the gears example.

Example 5 *Five interlocking gears, $G_1, \ldots, G_5$, are arranged in a row. If only G_1 is driven, then this invokes the rotation of G_1, and, for each $\langle G_i, G_{i+1}\rangle$ pair, the rotation of G_i invokes the rotation of G_{i+1}. However if both G_1 and G_5 are driven, then the rotation of each invokes the rotation of its neighbour, and, in turn, each of these rotations invokes the rotation of G_3.*

The first part of this example can be represented by the event theory $\Theta_5 = \Theta_{Ind} \cup \Theta_{Inv} \cup \Theta_{SCon} \cup \{(32), \ldots, (39)\}$, where:

$$\forall g, t(Pre(Drive(g))(t) \equiv Free(g)(t)) \tag{32}$$

$$\forall g, t(Eff(Drive(g))(t) \equiv Rotd(g)(t)) \tag{33}$$

$$\forall g, t(Pre(Rot(g))(t) \equiv Free(g)(t)) \tag{34}$$

$$\forall g, t(Eff(Rot(g))(t) \equiv Rotd(g)(t)) \tag{35}$$

$$\forall g, g', t(Intl(g, g')(t) \equiv Intl(g', g)(t)) \tag{36}$$

$$\forall g, t(Occ(Drive(g))(t) \rightarrow Inv_1(Drive(g), Rot(g))(t)) \tag{37}$$

$$\forall g, g', t(Intl(g, g')(t) \rightarrow SCon(Rot(g), Rot(g'))(t)) \tag{38}$$

$$U[G_1, \ldots, G_5] \wedge U[Drive, Rot] \wedge \bigwedge_{i=1}^{5} Free(G_i)(1) \wedge \bigwedge_{i=1}^{4} Intl(G_i, G_{i+1})(1) \wedge Occ(Drive(G_1))(1) \tag{39}$$

For the sake of simplicity the preconditions and effects of drive events are the same as those of rotate events and all objects are assumed to be gears. Thus a gear can be driven (can rotate) if it is free to do so, and the effect of its being driven (rotating) is that it has rotated (in a direction and by a degree which are, again for simplicity, not represented). Axiom (36) states that the Intl relation (which represents pairs of interlocked gears) is symmetric. Axiom (37) states that a drive-g event invokes a rotate-g event, and Axiom (38) states that if gears g and g' are interlocked, then their rotation is symmetrically constrained.

The theory Θ_5 has a single representative preferred$_1$ model in which all five gears rotate and in which there is a single invocation chain $\langle Drive(G_1), Rot(G_1), Rot(G_2), Rot(G_3), Rot(G_4), Rot(G_5)\rangle$.

Moreover, the extended theory $\Theta_5 \cup \{Occ(Drive(G_5))(1)\}$, has a single preferred$_1$ model in which all five gears rotate, and in which there are two invocation chains; $\langle Drive(G_1), Rot(G_1), Rot(G_2), Rot(G_3)\rangle$ and $\langle Drive(G_5), Rot(G_5), Rot(G_4), Rot(G_3)\rangle$.

Proposition $\Theta_5 \approx\!\!\!\models_1 Succ(Rot(G_i))(1)$ *for* $1 \leq i \leq 5$, *and* $\Theta_5 \approx\!\!\!\models_1 Inv(Rot(G_1), Rot(G_5))(1)$. *Moreover, if* $\Theta'_5 = \Theta_5 \cup \{Occ(Drive(G_5))(1)\}$, *then* $\Theta'_5 \approx\!\!\!\models_1 Succ(Rot(G_i))(1)$ *where* $1 \leq i \leq 5$, *but now both* $\Theta'_5 \approx\!\!\!\models_1 Inv(Rot(G_1), Rot(G_3))(1)$ *and* $\Theta'_5 \approx\!\!\!\models_1 Inv(Rot(G_5), Rot(G_3))(1)$.

Proof *Let M be a preferred$_1$ model of Θ_5. Then it follows, by axioms (17), (37) and (39), that $Occ(Drive(G_1))(1)$, $Inv_1(Drive(G_1), Rot(G_1))(1)$, and $Occ(Rot(G_1))(1)$ are true in M. By Axiom (39), $Intl(G_1, G_2)(1)$ is true in M, so it follows by Axiom (38) that $SCon(Rot(G_1), Rot(G_2))(1)$ is true in M. And so it follows, by axioms (26) and (27), that $SCon(Rot(G_2), Rot(G_1))(1)$ and $Occ(Rot(G_2))(1)$ are true in M. Similar reasoning shows that for each i such that $1 \leq i \leq 4$, both $SCon(Rot(G_i), Rot(G_{i+1}))(1)$ and $SCon(Rot(G_{i+1}), Rot(G_i))(1)$ are true in M, and that, for each i such that $1 \leq i \leq 5$, $Occ(G_i)(1)$ is true in M. By axioms (32), (34) and (39), $Pre(Drive(G_1))(1)$ is true in M, as is each $Pre(Rot(G_i))(1)$ where $1 \leq i \leq 5$. So by Definition 8.2, $Succ(Drive(G_1))(1)$ is true in M, as is each $Succ(Rot(G_i))(1)$ where $1 \leq i \leq 5$.*

By Axiom (28), the set $SI = \{Rot(G_i) : 1 \leq i \leq 5\}$ is a symmetric constraint set at time 1 in M, and by Definition 8.1 it is the only such set. So it follows by Axiom (29) that $EInv(Drive(G_1), Rot(G_1))(1)$ is true in M. Moreover it follows, by Definition 8.1 and Axiom (29), that $Rot(G_1)$ is the only initially invoked event in SI. So it follows by Axiom (30) that $\langle Rot(G_1), Rot(G_1)\rangle$ is an influence path of length 0 in SI, and that the shortest influence path to G_2 in SI is $\langle Rot(G_1), Rot(G_2)\rangle$ which is of length 1. So it follows by Axiom (31) that $Inv_1(Rot(G_1), Rot(G_2))(1)$ is true in M. By analogous reasoning, the shortest influence path to G_3 in SI is the path $\langle Rot(G_1), Rot(G_2), Rot(G_3)\rangle$ which is of length 2, and so $Inv_1(Rot(G_2), Rot(G_3))(1)$ is true in M. Similar reasoning shows that $Inv_1(Rot(G_3), Rot(G_4))(1)$ and $Inv_1(Rot(G_4), Rot(G_5))(1)$ are true in M. So it follows by Axiom (19) that $Inv(Rot(G_1), Rot(G_5))(1)$ is true in M.

Now let M be a preferred$_1$ model of Θ'_5. By reasoning analogous to that given above, $Succ(Drive(G_1))(1)$ is true in M, as is each $Succ(Rot(G_i))(1)$ where $1 \leq i \leq 5$. We also have that $SI = \{Rot(G_i) : 1 \leq i \leq 5\}$ is the only symmetric constraint set at time 1 in M. However this time $Rot(G_1)$ and $Rot(G_5)$ are both initially invoked events in SI. The shortest influence path to G_2 in SI is the path $\langle Rot(G_1), Rot(G_2)\rangle$ of length 1, and so, as before, $Inv_1(Rot(G_1), Rot(G_2))(1)$ is true in M. Analogous reasoning shows that

$\langle Rot(G_5), Rot(G_5)\rangle$ is an influence path in SI of length 0 and that the shortest influence path to G_4 in SI is the path $\langle Rot(G_5), Rot(G_4)\rangle$ of length 1, and so $Inv_1(Rot(G_5), Rot(G_4))(1)$ is true in M. Moreover, the shortest influence paths to $Rot(G_3)$ in SI are the paths $\langle Rot(G_1), Rot(G_2), Rot(G_3)\rangle$ and $\langle Rot(G_5), Rot(G_4), Rot(G_3)\rangle$, each of length 2. So it follows by Axiom (31) that $Inv_1(Rot(G_2), Rot(G_3))(1)$ and $Inv_1(Rot(G_4), Rot(G_3))(1)$ are true in M. Thus by Axiom (19), $Inv(Rot(G_1), Rot(G_3))(1)$ and $Inv(Rot(G_5), Rot(G_3))(1)$ are both true in M.

Secondary events can also be used to represent non-deterministic effects.

Non-deterministic effects may arise because of uncertainty in the preconditions. Suppose, for example, that the initial positions of blocks B and B' are uncertain, either B is on B' or conversely, and that B is moved. Intuitively the resulting location of B' should be uncertain. If B' was on B, then B' should have moved to the same location as B, otherwise B' should have remained where it was. It is easy to see how Example 4 can be adapted to represent this example faithfully. If B' is on B initially, then the move-B event invokes a secondary move-B' event (Axiom (23)) which results in B' moving with B, otherwise the location of B' remains unchanged. The resulting theory thus has two representative preferred$_1$ models; one in which B' moves, and one in which B' does not move.

non-deterministic effects may also arise because the outcome of the events in question is uncertain. If successful, a non-deterministic event is not regularly followed by a definite effect, but rather by any one of a set of mutually exclusive possible effects; as in the classic example of tossing a fair coin. In the terms used in the introduction, our expectations regarding the outcome of such events are unclear. In order to see why secondary events are needed to represent these, consider the following attempt to represent coin-tossing.

Example 6 *Suppose that a fair coin is showing heads initially and that the coin is tossed. As the coin is fair, the result should be uncertain; the coin may show heads or it may show tails. Let $\Theta_6 = \Theta_{Ind} \cup \{(40), (41), (42)\}$ where:*

$$\forall t(Pre(Toss)(t) \equiv (Heads(t) \oplus Tails(t))) \tag{40}$$

$$\forall t(Eff(Toss)(t) \equiv (Heads(t) \oplus Tails(t))) \tag{41}$$

$$Heads(1) \land \neg Tails(1) \land Occ(Toss)(1) \tag{42}$$

Then, contrary to intention, Θ_6 is deterministic$_1$. Θ_6 has a single representative preferred$_1$ model in which the Toss event succeeds and its effect $Heads(2) \oplus Tails(2)$ is true. However, as $Heads(1)$ is true in the model, the normal application of the inertia axiom removes the uncertainty by determining that $Heads(2)$ is true.

If, as in this example, one of the alternative effects of an event preserves the status quo, then inertia will favour that outcome and so determine the outcome of the event. The fact that inertia intervenes in this way in the example suggests that something is missing from the representation of the toss event. As defined, the event does not do what it is intended to do. In succeeding it does not ensure that there are two distinct outcomes. Metaphorically speaking, it does not introduce a fork at this point in history, resulting in two alternative futures. There seems to be a hidden causal element which accounts for our intuitive understanding of the example but which is missing from the formalization of it.

I suggest that the missing component is the causal structure of the toss event, and that this can be faithfully represented by two conflicting secondary events; that tossing the coin invokes two conflicting deterministic events, one resulting in the coin showing heads, the other resulting in the coin showing tails.

When formalizing this and subsequent examples involving nondeterministic events, the following abbreviation is useful:

$$Inv_1(e, \{e_1, \ldots, e_n\})(t) \quad =_{\mathrm{Df}} \quad Inv_1(e, e_1)(t) \wedge \ldots \wedge Inv_1(e, e_n)(t)\ .$$

Example 7 *Let $\Theta_7 = \Theta_{Ind} \cup \Theta_{Inv} \cup \{(40), \ldots, (46)\}$ where:*

$$\forall t(Pre(TossH)(t) \equiv Pre(Toss)(t)) \wedge \forall t(Eff(TossH)(t) \equiv Heads(t)) \tag{43}$$
$$\forall t(Pre(TossT)(t) \equiv Pre(Toss)(t)) \wedge \forall t(Eff(TossT)(t) \equiv Tails(t)) \tag{44}$$
$$\forall t(Occ(Toss)(t) \rightarrow Inv_1(Toss, \{TossH, TossT\})(t)) \tag{45}$$
$$U[Toss, TossH, TossT] \tag{46}$$

Then there are two representative preferred$_1$ models of Θ_7. In both models, the primary event Toss succeeds and invokes the conflicting secondary events TossH and TossT. In one of the models, TossH succeeds and TossT fails. In the other model, TossT succeeds and TossH fails.

Proposition $\Theta_7 \approx_1 Heads(2) \oplus Tails(2)$, $\Theta_7 \not\approx_1 Heads(2)$, *and* $\Theta_7 \not\approx_1 Tails(2)$.

Proof *For the first part, let M be a preferred$_1$ model of Θ_7. Then the Toss event occurs at time 1 in M (Axiom (42)) and its occurrence invokes the secondary TossH and TossT events (Axiom (45)), which also occur at time 1 in M (Axiom (17)). The preconditions of all three events are true at time 1 in M (axioms (40), (42)-(44)). However, in view of their effects (axioms (41), (43), (44)), all three events cannot succeed at time 1 in M. As Toss invokes the other two events, and is the only event which does so (Definition 8.1), TossH or TossT can only succeed if Toss does (Axiom (18)). Moreover, it is consistent to assume that Toss does succeed at time 1 in M, so it follows (Definition 8.2) that Toss does succeed at time 1 in M, with the effect that $Heads(2) \oplus Tails(2)$ is true in M (axioms (1), (41)).*

For the second part it is sufficient to show that there is a preferred$_1$ model of Θ_7 in which $Heads(2)$ is false. So, let M be an $\mathcal{EL}$ model in which $\forall R\, Phys(R)$ is true, which satisfies Axiom (46), and which satisfies the following conditions. The only evidential atoms which are defined in M are those which satisfy the following set of literals:

$$\{Heads(1), \neg Tails(1), Occ(Toss)(1), Inv_1(Toss, TossH)(1), Occ(TossH)(1),$$
$$Inv_1(Toss, TossT)(1), Occ(TossT)(1)\} \cup \{Tails(t) : t \geq 2\} \cup \{\neg Heads(t) : t \geq 2\}\ ;$$

thus, for example, $Heads(1)$ is true in M and $Tails(1)$ is false in M. The success atoms in the set:

$$\{Succ(Toss)(1), Succ(TossT)(1)\}$$

are both true in M, and every other success atom is false in M. Finally the inertia atoms in the set:

$$\{Inert(Heads)(0), Inert(Tails)(0), Inert(Heads)(1), Inert(Tails)(1)\}$$

are all false in M, and every other inertia atom is true in M. Clearly M exists. Moreover, inspection shows that M is a preferred$_1$ model of Θ_7. In particular, the success of Toss at time 1 in M is required (on the grounds given in the proof of the first part) as is the success of either TossH or TossT at time 1 in M (axioms (1), (41), (43), (44), Definition 8.2). In the given model, TossT succeeds at time 1, and the combined effect of the two successful events is that Heads(2) is false in M (axioms (1), (41), (44)).

The proof of the third part is similar; but with TossH (rather than TossT) succeeding at time 1 in the given model.

A similar problem is posed by the Russian Shooting Problem (Sandewall, 1994). A revolver is loaded with a single bullet and the cylinder is spun. The result should be uncertain, as the cylinder could come to rest with the bullet in any one of six possible positions. However in the naive representation, inertia will, once again, favour the outcome in which the bullet returns to, and so effectively remains in, its original position. A faithful representation of the expected outcome of the spin event can be obtained by having it invoke six conflicting secondary spin events, with the result that the bullet is equally likely to come to rest in any of the six possible positions.

More generally, a non-deterministic event can be seen as invoking a set of conflicting deterministic events, each resulting in one of its possible outcomes.

When there are more than two possible outcomes, the following general form of exclusive alternation is useful:

$$\phi_1 \,|\, \phi_2 \,|\, \ldots \,|\, \phi_n =_{\mathrm{Df}} \bigvee_{m=1}^{n} (\bigwedge_{i=1}^{m-1} \neg\phi_i \wedge \phi_m \wedge \bigwedge_{j=m+1}^{n} \neg\phi_j) \;.$$

Thus $\phi_1 \,|\, \phi_2 \,|\, \ldots \,|\, \phi_n$ is true if exactly one of the alternatives is true and the remainder are all false, is false if all of the alternatives are false or if more than one of them is true, and is undefined otherwise.[28]

The techniques for representing ramifications and non-deterministic events can readily be combined to represent events with conditional effects, such as crossing the points on a railway line. Suppose that point P has entry point PE, left exit point PL, and right exit point PR. Suppose further that if P is set to 'left', then a train crossing P should emerge at PL, otherwise if P is set to 'right', then the train should emerge at PR. Then the preconditions for crossing P can simply be that the train is at PE and that P is set to either 'left' or 'right', and the effect can be that the train is at either PL or PR. If P is set to 'left' when the train crosses, then this event should invoke a cross-P-left event, which has the precondition that the train is at PE and that P is set to 'left', and the effect that the train is at PL. Similarly, if P is set to 'right' when the train crosses, then this event should invoke a cross-P-right event with the effect that the train emerges at PR.

5. Event Preferences

Conflicts are the reductio ad absurdum of Deductionism. If two events conflict, then their effects cannot, on pain of inconsistency, be deduced. Consequently, Deductionists wishing

28. So, if $n = 2$, then exclusive alternation is just exclusive disjunction. But if $n > 2$, then the two notions differ; for example, the exclusive disjunction $\phi_1 \oplus (\phi_2 \oplus \phi_3)$ is true if $\phi_1 \wedge \phi_2 \wedge \phi_3$ is.

to represent simultaneous events are forced to police their events and to regulate their effects. Gelfond, Lifschitz, and Rabinov (1991) suggest that this can be done by means of *cancellations*. Suppose, for example, that we have a bowl full of soup. If only one side of the bowl is lifted, then the soup is spilt. However, if both sides of the bowl are lifted simultaneously, then the soup is not spilt. These interactions are represented by means of two elementary lift events (lift-left-side, lift-right-side) and their composition (lift-both-sides). If either of the elementary events occurs in isolation, then it has the effect that the soup is spilt. But if the complex event occurs, then the spill-effects of its component actions should be cancelled, with the result that the bowl is lifted and the soup is not spilt.

However cancellations are only appropriate if we ignore the possibility of failure. Suppose that in reality one of the elementary lift events fails because a hand slips. Then the effects of the other component event should not be cancelled and the soup should be spilt. But how can a cancellation be cancelled?

Conflicts do not pose the same problem for inductive events. When faced with conflicting events we do not expect that they will both succeed. Some, and possibly all, of our expectations regarding these events are uncertain. This is reflected in the formal theory: conflicting effects give rise to conflicting success atoms, resulting in failure rather than inconsistency. However the possibility of failure raises the problem of "over-weak" predictions (which amusingly complements the "over-strong" problem of Deductionism). If two inductive events conflict (and there are no other interactions involving them), then each succeeds in a preferred model in which the other fails, consequently nothing more definite than the disjunction of their effects is predicted. This is appropriate when the events are of equal status; indeed, it provides the basis for the representation of non-deterministic events given in the previous section. But it is not appropriate when we expect that one of the events will succeed. Suppose, for example, that Stan and Ollie attempt to move to the same location simultaneously but that only one of them can succeed. Suppose further that Ollie's success is more likely than Stan's, say because he is bigger. Then we normally expect that Ollie will succeed and that Stan will fail. However, if there are abnormal independent circumstances which lead us to expect that Ollie will fail, then these do not lead us to expect that Stan will fail also. Indeed, under the circumstances, we expect that Stan will succeed. For example, if Ollie trips, then we expect that Stan will succeed; although, of course, he may also trip, etc.

Asymmetric expectations of this kind can be thought of as *event preferences*, as preferences over the outcomes of events. Thus if events e and e' conflict, and we expect that e will succeed, then the success of e is preferred to that of e'. However, as the examples show, preferences of this kind should be defeasible in order not to prejudice the success of the non-preferred event should the preferred event fail for some independent reason.

Event preferences can be represented in $\mathcal{EL}$ by *preference atoms*, event atoms of the form $\mathit{Pref}(e, e')(t)$. In keeping with the above discussion, these should be interpreted as stating that the success of event e is *normally* preferred to that of event e' should they conflict at time t. The temporal index accommodates the possibility that event preferences may vary over time; as in Example 10 below.

The only logical restriction on event preferences is that they are required to be asymmetric (Axiom (47) in Table 2). Further conditions, such as transitivity, can, of course, be added where appropriate.

Table 5: The theory of event preferences, Θ_{Pref}

$$\forall e, e', t(Pref(e, e')(t) \rightarrow \neg \top Pref(e', e)(t)) \tag{47}$$

Definition 13 *The* theory of event preferences *is given by the axioms in Table 5; thus* $\Theta_{Pref} = \{(47)\}$*. The* theory of natural events*,* Θ_{NE}*, consists of the axioms given in tables 2, 3, 4, and 5; thus* $\Theta_{NE} = \Theta_{Ind} \cup \Theta_{Inv} \cup \Theta_{SCon} \cup \Theta_{Pref}$*.*

The intended interpretation of event preferences cannot be enforced by adding the axiom:

$$\forall e, e', t((Pref(e, e')(t) \wedge Pre(e)(t) \wedge Occ(e)(t) \wedge Succ(e')(t)) \rightarrow Succ(e)(t)) \ .$$

Adding this axiom would ensure that if e and e' were to conflict, then e would succeed (and so e' would fail). However, if e were to fail for some independent reason, then the axiom would have the undesirable effect of forcing the failure of e'.

Clearly, if event preferences are to be interpreted correctly, then a more flexible approach is needed, and so it is necessary to extend the pragmatics of event theories which contain them. In doing so the aim is to produce a consistent interpretation of the applicable event preferences where possible, and to ignore them otherwise.

To begin with, a distinction is drawn between preferential events, the events to which the preferences can consistently be applied, and non-preferential events, all of the remaining events under consideration. An event e is said to be *supported* in model M at time t if e occurs at t and e's preconditions are true in M at t (if $Pre(e)(t)$ and $Occ(e)(t)$ are both true in M). If e is supported at some time point in a model, and the time point and model are clear from the context, then e will simply be said to be supported. Now, if the applicable event preferences form an acyclic chain $Pref(e, e')(t)$, $Pref(e', e'')(t)$, $\ldots$, then the supported events occurring in them can be ordered lexicographically; thus given that e, e' and e'' are all supported, e has order 1, e' order 2, and e'' order 3. More generally, for model M and time point t, a supported event e may (or may not) be assigned a *preference rank* as follows:

- e has preference rank 1 if there is some event e' such that $Pref(e, e')(t)$ is true and there is no supported event e' such that $Pref(e', e)(t)$ is true, and
- e has preference rank n if e does not have a preference rank $m < n$ and there is a supported event e' with preference rank $n-1$ which is directly preferred to e (which is such that $Pref(e', e)(t)$ is true and there is no supported event e'' such that the preferences $Pref(e, e'')(t)$ and $Pref(e'', e')(t)$ are both true).

Let e be a supported event (at time t in model M). Then e is a *preferential event* (at t in M) if e has a preference rank (at t in M), otherwise e is a *non-preferential event* (at time t in model M).

Now suppose that at time t models M and M' agree on event preferences, preferential events, and non-preferential events. Then, at t:

- *M is better than M' on preferential events* if there is some preference rank n such that M and M' agree on the success of all events with preference rank $m < n$, and more events with preference rank n succeed in M,

- *M is as good as M' on preferential events* if M is better than M' on preferential events, or M and M' agree on the success of preferential events,

- M is *better than M' on non-preferential events* if more non-preferential events succeed in M,

- M is *as good as M' on non-preferential events* if M is better than M' on nonpreferential events, or M and M' agree on the success of non-preferential events.

The definition of a preferred model can now be refined.

Definition 14 (Preference$_2$) *Let M and M' be models which differ at most on the interpretation of temporally-indexed relations. Then M is* preferred$_2$ *to M' (written $M \prec_2 M'$) iff there is a time point t such that M and M' agree before t, and at t:*

1. *fewer evidential atoms are defined in M, and M and M' agree on the truth values of all evidential atoms which are defined in M; or*

2. *M and M' differ only on conjectural atoms, and either*

   (a) *M is better than M' on preferential events, and M is as good as M' on non-preferential events; or*

   (b) *M is as good as M' on preferential events, and M is better than M' on non-preferential events; or*

3. *M and M' differ only on inertia atoms, and more inertia atoms are true in M.*

Note that preference$_2$ reduces to preference$_1$ when event theories do not include event preferences.

Four examples are now given. The first illustrates the interpretation of event preferences.

Example 8 *Suppose that three agents, Stan, Ollie, and Charlie, are at locations $L1$, $L2$ and $L3$ respectively. It is assumed that at most one of the agents can be at a location at a point in time, so if any two of them attempt to move to the same location simultaneously, then at most one of them can succeed. If Ollie and Stan both attempt to move to location $L4$ then, as Ollie is bigger, Ollie's success is expected. Similarly, if Stan and Charlie both attempt to move to $L4$ simultaneously, then, as Stan is bigger, Stan's success is expected. However, Charlie is much faster than Ollie, so if they both attempt to move to $L4$ simultaneously, then Charlie's success is expected. Now, suppose that in fact Stan and Ollie both attempt to move to $L4$ simultaneously, then, as indicated, Ollie's success is expected. However, if Charlie also attempts to move to $L4$, then the result is uncertain because the preferences among the move events can no longer be interpreted consistently.*

Let $\Theta_8 = \Theta_{Ind} \cup \Theta_{Pref} \cup \{(12), (13), (14)\} \cup \{(48), \ldots, (52)\}$*; where:*

$$\forall x, y, l, t((At(x, l)(t) \wedge \neg x = y) \rightarrow \neg At(y, l)(t)) \tag{48}$$
$$Pref(Move(O, L1, L4), Move(S, L2, L4))(1) \tag{49}$$
$$Pref(Move(S, L2, L4), Move(C, L3, L4))(1) \tag{50}$$
$$Pref(Move(C, L3, L4), Move(O, L1, L4))(1) \tag{51}$$
$$\begin{aligned} &U[O, S, C, L1, L2, L3, L4] \wedge U[Move] \\ &\quad \wedge At(O, L1)(1) \wedge At(S, L2)(1) \wedge At(C, L3)(1) \\ &\quad \wedge Occ(Move(O, L1, L4))(1) \wedge Occ(Move(S, L2, L4))(1) \end{aligned} \tag{52}$$

Then Θ_8 *has a single representative preferred*$_2$ *model in which Ollie succeeds in moving to L4. However the extended theory* $\Theta'_8 = \Theta_8 \cup \{Occ(Move(C, L3, L4))(1)\}$ *has three representative preferred*$_2$ *models, and a different member of the trio succeeds in each of them.*

Proposition $\Theta_8 \approx_2 At(O, L4)(2)$*. However* $\Theta'_8 \not\approx_2 At(O, L4)(2)$*,* $\Theta'_8 \not\approx_2 At(S, L4)(2)$*, and* $\Theta'_8 \not\approx_2 At(C, L4)(2)$*.*

Proof *For the first part, let* M *be a preferred*$_2$ *model of* Θ_8*. Then (Definition 14.1) the only events which occur at time* 1 *in* M *are* $Move(O, L1, L4)$ *and* $Move(S, L2, L4)$*. If follows from axioms (12) and (52) that both of these events are both supported at time* 1 *in* M*. In view of their effects, only one of these events can succeed (axioms (1), (13), (48), (52)). By axioms (49)-(51) and Definition 14.1, the only preference atoms which are true in* M *are* $Pref(Move(O, L1, L4), Move(S, L2, L4))(1)$*,* $Pref(Move(S, L2, L4), Move(C, L3, L4))(1)$*, and* $Pref(Move(C, L3, L4), Move(O, L1, L4))(1)$*. In view of the first of these, there is an event* e *such that* $Pref(Move(O, L1, L4), e)(1)$ *is true in* M*. And, as* $Move(C, L3, L4)$ *is not supported at time* 1 *in* M*, there is no supported event* e *such that* $Pref(e, Move(O, L1, L4))(1)$ *is true in* M*. So* $Move(O, L1, L4)$ *has preference rank* 1 *at time* 1 *in* M*. Moreover,* $Move(S, L2, L4)$ *does not have preference rank* 1 *at time* 1 *in* M*, because* $Move(O, L1, L4)$ *is a supported event at time* 1 *in* M *and* $Pref(Move(O, L1, L4), Move(S, L2, L4))(1)$ *is true in* M*. So, as* $Move(O, L1, L4)$ *is directly preferred to* $Move(S, L2, L4)$ *at time* 1 *in* M*, it follows* $Move(S, L2, L4)$ *has preference rank* 2 *at time* 1 *in* M*. It therefore follows by Definition 14.2(a) that* $Succ(Move(O, L1, L4))(1)$ *is true in* M*. Consequently the effect* $At(O, L4)(2)$ *is true in* M *(axioms (1), (13)).*

For the second part, let M *be an* $\mathcal{EL}$ *model in which* $\forall R\, Phys(R)$ *is true, and which satisfies* $U[O, S, C, L1, L2, L3, L4]$ *and* $U[Move]$*. Suppose further that* M *satisfies the following conditions; where* $Loc = \{O, S, C, L1, L2, L3, L4\}$*. The only object or event atoms which are defined in* M *are those which satisfy the following sets of literals:*

$$\begin{aligned} &\{At(O, L1)(t) : t \geq 1\}, \{\neg At(O, l)(t) : t \geq 1, l \in Loc, l \neq L1\}, \\ &\{At(S, L2)(1)\}, \{At(S, L4)(t) : t \geq 2)\}, \\ &\quad \{\neg At(S, l)(1) : l \in Loc, l \neq L2\}, \{\neg At(S, l)(t) : t \geq 2, l \in Loc, l \neq L4\}, \\ &\{At(C, L3)(t) : t \geq 1\} \cup \{\neg At(C, l)(t) : t \geq 1, l \in Loc, l \neq L3\}, \\ &\{Occ(Move(O, L1, L4))(1), Occ(Move(S, L2, L4))(1), Occ(Move(C, L3, L4))(1)\}, \end{aligned}$$

$$\{Pref(Move(O, L1, L4), Move(S, L2, L4))(1),$$
$$Pref(Move(S, L2, L4), Move(C, L3, L4))(1),$$
$$Pref(Move(C, L3, L4), Move(O, L1, L4))(1)\}.$$

The success atom $Succ(Move(S, L2, L4))(1)$ *is true in* M *and all other success atoms are false in* M*. Finally, the inertia atoms in the following sets are all false in* M*:*

$$\{Inert(At, \langle O, l\rangle)(0) : l \in Loc\}, \{Inert(At, \langle C, l\rangle)(0) : l \in Loc\},$$
$$\{Inert(At, \langle S, l\rangle)(0) : l \in Loc\}, \{Inert(At, \langle S, l\rangle)(1) : l \in \{L2, L4\}\},$$

and all other inertia atoms are true in M*. Clearly* M *exists. Moreover, inspection shows that* M *is a preferred*$_2$ *model of* Θ'_8*. In particular, the success of* $Move(S, L2, L4)$ *at time* 1 *can be justified as follows. The only three events which are supported at time* 1 *in* M *are* $Move(O, L1, L4)$*,* $Move(S, L2, L4)$*, and* $Move(C, L3, L4)$ *(axioms (12), (52), definition of* Θ'_8*, Definition 14.1). And the only preference atoms which are true in* M *are those in the set given above (axioms (49)-(51), Definition 14.1). So none of the three supported events has preference rank* 1 *at time* 1 *in* M*;* $Move(O, L1, L4)$ *does not because* $Move(C, L3, L4)$ *is supported and is preferred to it,* $Move(S, L2, L4)$ *does not because* $Move(O, L1, L4)$ *is supported and is preferred to it, and* $Move(C, L3, L4)$ *does not because* $Move(S, L2, L4)$ *is supported and is preferred to it. It follows that all three events are non-preferential at time* 1 *in* M*. At most one of these events can succeed at time* 1 *in* M *(axioms (1), (13), (48), (52)). And it follows by Definition 14.2(b) that one of them does succeed. In* M*,* $Move(S, L2, L4)$ *succeeds at time* 1*, with the effect that* $At(S, L4)(2)$ *and consequently* $\neg At(O, L4)(2)$ *are both true in* M *((1), (13), (48), (52)).*

The given model also establishes the fourth part of the proposition. For the third part a preferred$_2$ *model of* Θ'_8 *can be given in which either* $Move(O, L1, L4)$ *or* $Move(C, L3, L4)$ *succeeds at time* 1*.*

The next example shows how event preferences and secondary events can be combined in order to represent implicit cancellation and the implicit cancellation thereof.

Example 9 *The soup-bowl example can be represented as follows:*

$$\forall t(Pre(LiftL)(t) \equiv \neg HoldingL(t)) \wedge \forall t(Eff(LiftL)(t) \equiv HoldingL(t)) \tag{53}$$
$$\forall t(Pre(LiftR)(t) \equiv \neg HoldingR(t)) \wedge \forall t(Eff(LiftR)(t) \equiv HoldingR(t)) \tag{54}$$
$$\forall t(Pre(Spill)(t) \equiv \neg Spilt(t)) \tag{55}$$
$$\forall t(Eff(Spill)(t) \equiv (Spilt(t) \wedge (HoldingL(t) \oplus HoldingR(t)))) \tag{56}$$
$$\forall e, t((Occ(e)(t) \wedge (e = LiftL \vee e = LiftR) \wedge \neg Spilt(t)) \rightarrow Inv_1(e, Spill)(t)) \tag{57}$$
$$\forall t(Pre(Lift)(t) \equiv (Pre(LiftL)(t) \wedge Pre(LiftR)(t))) \tag{58}$$
$$\forall t(Eff(Lift)(t) \equiv (Eff(LiftL)(t) \wedge Eff(LiftR)(t))) \tag{59}$$
$$\forall t(Occ(Lift)(t) \equiv (Occ(LiftL)(t) \wedge Occ(LiftR)(t))) \tag{60}$$
$$\forall t Pref(Lift, Spill)(t) \tag{61}$$
$$U[Lift, LiftL, LiftR, Spill] \wedge \neg HoldingL(1) \wedge \neg HoldingR(1) \wedge \neg Spilt(1) \wedge Occ(Lift)(1) \tag{62}$$

Thus axioms (53)-(56) define elementary lift-left (LiftL), lift-right (LiftR) and spill events. Axiom (57) states that if the soup is not spilt, then the occurrence of either elementary lift event invokes a spill event. Axioms (58)-(60) define the complex lift-both event (Lift), and Axiom (61) states that the success of a lift-both event is normally preferred to that of a spill event.

Let $\Theta_9 = \Theta_{Ind} \cup \Theta_{Inv} \cup \Theta_{Pref} \cup \{(53), \ldots, (62)\}$*. Then* Θ_9 *has a single representative preferred*$_2$ *model in which the Lift event succeeds and the Spill event fails; with the result that the soup is not spilt. Thus the success of the Lift event implicitly cancels the (secondary) Spilt effect of its component LiftL and LiftR events. Moreover, the extended theory* $\Theta_9 \cup \{\neg Succ(Lift)(1)\}$ *has two representative preferred*$_2$ *models. In one LiftL succeeds (and LiftR fails), in the other LiftR succeeds (and LiftL fails). So in both Spill succeeds with the effect that the soup is spilt. The implicit cancellation of the Spilt effect is thus itself implicitly cancelled.*

Proposition $\Theta_9 \approx\!\!\!\mid_2 \neg Spilt(2)$*, and* $\Theta_9 \cup \{\neg Succ(Lift)(1)\} \approx\!\!\!\mid_2 Spilt(2)$*.*

Proof *For the first part, let* M *be a preferred*$_2$ *model of* Θ_9*. Then the events Lift, LiftL, and LiftR are all supported at time* 1 *in* M *(axioms (53), (54), (58), (60), (62)). Moreover the LiftL and LiftR events both invoke the Spill event at time* 1 *(axioms (57), (62)), and this is also supported at time* 1 *in* M *(axioms (17), (55), (62)). Now the Lift and Spill events conflict at time* 1 *in* M*; if Lift succeeds, then Spill fails, and vice-versa (axioms (1), (53), (54), (56), (59)). And the preference atom* $Pref(Lift, Spill)(1)$ *is the only preference atom with temporal index* 1 *which is true in* M *(Axiom (61), Definition 14.1). Consequently, at time* 1 *in* M*, Lift and Spill have preference ranks* 1 *and* 2 *respectively, and LiftL and LiftR are non-preferential. The success of Lift implies the success of its component events (axioms (1), (53), (54), (59)). So (Definition 14.2(a)) Lift succeeds (and Spill fails) at time* 1 *in* M*. And so, as* $Phys(Spilt)$ *and* $\neg Spilt(1)$ *are true in* M *(Axiom (62), Definition 9), it follows by inertia (Axiom (3), Definition 14.3) that* $\neg Spilt(2)$ *is true in* M*.*

For the second part, let M *be a preferred*$_2$ *model of* $\Theta_9 \cup \{\neg Succ(Lift)(1)\}$*. Then, as above, the events Lift, LiftL, LiftR, and Spill are all supported at time* 1 *in* M*. As Lift fails (when supported) at time* 1 *in* M*, one of its component lift-events must also fail at time* 1 *in* M *(axioms (1), (53), (54), (59)). However the other component event succeeds at time* 1 *in* M *(Definition 14.2(b)), and consequently so does Spill (axioms (1), (53) (54), (56), Definition 14.2(a)), with the effect that* $Spilt(2)$ *is true in* M *(axioms (1), (56)).*

The next example shows how event preferences and secondary events can be used to represent changing expectations regarding the outcome of non-deterministic events.

Example 10 *A race between a fast horse and a strong horse will be run the day after tomorrow. The course is currently dry. Given that this remains the case, the fast horse is expected to win. However if it were to rain tomorrow, then the strong horse would be expected to win.*

This example can be represented by the theory $\Theta_{10} = \Theta_{Ind} \cup \Theta_{Inv} \cup \Theta_{Pref} \cup \{(63), \ldots, (71)\}$*; where:*

$$\forall t(Pre(RaceFS)(t) \equiv Occ(RaceFS)(t)) \tag{63}$$

$$\forall t(\mathit{Eff}(\mathit{RaceFS})(t) \equiv (\mathit{WinnerF}(t) \oplus \mathit{WinnerS}(t))) \tag{64}$$
$$\forall t(\mathit{Pre}(\mathit{WinF})(t) \equiv \mathit{Pre}(\mathit{RaceFS})(t)) \wedge \forall t(\mathit{Eff}(\mathit{WinF})(t) \equiv \mathit{WinnerF}(t)) \tag{65}$$
$$\forall t(\mathit{Pre}(\mathit{WinS})(t) \equiv \mathit{Pre}(\mathit{RaceFS})(t)) \wedge \forall t(\mathit{Eff}(\mathit{WinS})(t) \equiv \mathit{WinnerS}(t)) \tag{66}$$
$$\forall t(\mathit{Occ}(\mathit{RaceFS})(t) \rightarrow \mathit{Inv}_1(\mathit{RaceFS}, \{\mathit{WinF}, \mathit{WinS}\})(t)) \tag{67}$$
$$\forall t(\mathit{Pre}(\mathit{Rain})(t) \equiv \mathit{Occ}(\mathit{Rain})(t)) \wedge \forall t(\mathit{Eff}(\mathit{Rain})(t) \equiv \neg \mathit{Dry}(t))) \tag{68}$$
$$\forall t(\mathit{Dry}(t) \rightarrow \mathit{Pref}(\mathit{WinF}, \mathit{WinS})(t)) \tag{69}$$
$$\forall t(\neg \mathit{Dry}(t) \rightarrow \mathit{Pref}(\mathit{WinS}, \mathit{WinF})(t)) \tag{70}$$
$$U[\mathit{RaceFS}, \mathit{WinF}, \mathit{WinS}, \mathit{Rain}] \wedge \mathit{Dry}(1) \wedge \mathit{Occ}(\mathit{RaceFS})(3) \tag{71}$$

Thus Axiom (64) states that the outcome of a successful race between the two horses (RaceFS) results in either the fast horse winning (WinnerF) or the strong horse winning (WinnerS). The causal structure of the RaceFS event (the fact that it involves the competition between two horses) is represented by its invocation of the conflicting WinF and WinS events (axioms (64)-(67)). Rain results in the course being wet (not dry) (Axiom (68)). If the course is dry, then the fast horse is expected to win (Axiom (69)), otherwise the slow horse is expected to win (Axiom (70)).

Proposition $\Theta_{10} \approx_2 \mathit{WinnerF}(4)$*, and* $\Theta_{10} \cup \{\mathit{Occ}(\mathit{Rain})(2)\} \approx_2 \mathit{WinnerS}(4)$*.*

Proof *For the first part, let M be a preferred$_2$ model of Θ_{10}. Then $\mathit{Phys}(\mathit{Dry})$ and $\mathit{Dry}(1)$ are true in M (Axiom (71), Definition 9), and so it follows by inertia (Axiom (3), Definition 14.3) that $\mathit{Dry}(3)$ is true in M. Moreover, the event RaceFS occurs at time 3 in M and invokes the WinF and WinS events (axioms (67), (71)). All three events are supported at time 3 in M (axioms (17), (63), (65), (66), (71)). As RaceFS invokes WinF and WinS, they can succeed only if it does (Axiom (18), Definition 14.1). So it follows (Definition 14.2(b)) that RaceFS succeeds at time 3 in M. The success of exactly one of the two invoked events at time 3 is consistent with the success of RaceFS (axioms (64)-(66)). As $\mathit{Dry}(3)$ is true in M it follows (Axiom (69)) that $\mathit{Pref}(\mathit{WinF}, \mathit{WinS})(3)$ is true in M. Moreover this is the only preference atom with temporal index 3 which is true in M (Definition 14.1). So, as WinF and WinS are both supported at time 3 in M, it follows that WinF has preference rank 1 and WinS has preference rank 2 at time 3 in M. Consequently (Definition 14.2(a)) WinF succeeds in M at time 3, with the effect that $\mathit{WinnerF}(4)$ is true in M (axioms (1) and (65)).*

For the second part, let M be a preferred$_2$ model of $\Theta_{10} \cup \{\mathit{Occ}(\mathit{Rain})(2)\}$. Then the Rain event succeeds at time 2 in M and so its effect $\neg\mathit{Dry}(3)$ is true in M (Axiom (68), Definition 14.2(b)). As in the proof of the first part, RaceFS and one of the two conflicting secondary events, WinS and WinF, succeed at time 3 in M. As $\neg\mathit{Dry}(3)$ is true in M, it follows (Axiom (70)) that $\mathit{Pref}(\mathit{WinS}, \mathit{WinF})(3)$ is true in M. Moreover, this is the only preference atom with temporal index 3 which is true in M (Definition 14.1). So, as WinS and WinF are both supported at time 3 in M, it follows that WinS has preference rank 1 and WinF has preference rank 2 at time 3 in M. So it follows (Definition 14.2(a)) that WinS succeeds at time 3 in M, with the effect that $\mathit{WinnerS}(4)$ is true in M (axioms (1) and (66)).

More generally, event preferences can be combined with secondary events in order to give a qualitative representation of conditional probabilities. A probability judgment of the form $P(Succ(e)(t)|Succ(e')(t)) = n$ states that the probability of event e succeeding at time t, given that event e' does, is n. Judgments of this kind can be represented in $\mathcal{EL}$ by event atoms of the form $Prob(e, e', n)(t)$. Conditional probabilities can then be translated into event preferences by means of the following axiom:

$$\forall e, e', e'', n, m, t((Prob(e, e', n)(t) \wedge Prob(e'', e', m)(t) \wedge n > m) \rightarrow Pref(e, e'')(t)) \tag{72}$$

together with the appropriate invocations. The use of this technique is illustrated by the final example, which is a probabilistic extension of an example attributed to Reiter (Shanahan, 1997, p. 290).

Example 11 *A chessman is placed haphazardly on a chessboard. It may end up on (within) a single square, but it is more likely that it will overlap four squares, and more likely still that it will overlap just two squares. This situation can be represented by the theory* $\Theta_{11} = \Theta_{Ind} \cup \Theta_{Inv} \cup \Theta_{Pref} \cup \{(72), (73), \ldots, (80)\}$*, where:*

$$\forall t(Pre(Place)(t) \equiv Occ(Place)(t)) \tag{73}$$
$$\forall t(Eff(Place)(t) \equiv (One(t) \mid Two(t) \mid Four(t))) \tag{74}$$
$$\forall t(Pre(Place1)(t) \equiv Pre(Place)(t)) \wedge \forall t(Eff(Place1)(t) \equiv One(t)) \tag{75}$$
$$\forall t(Pre(Place2)(t) \equiv Pre(Place)(t)) \wedge \forall t(Eff(Place2)(t) \equiv Two(t)) \tag{76}$$
$$\forall t(Pre(Place4)(t) \equiv Pre(Place)(t)) \wedge \forall t(Eff(Place4)(t) \equiv Four(t)) \tag{77}$$
$$\forall t(Occ(Place)(t) \rightarrow Inv_1(Place, \{Place1, Place2, Place4\})(t)) \tag{78}$$
$$\forall t(Prob(Place1, Place, 0.2)(t) \wedge Prob(Place2, Place, 0.5)(t) \wedge Prob(Place4, Place, 0.3)(t)) \tag{79}$$
$$U[Place, Place1, Place2, Place4] \wedge Occ(Place)(1) \tag{80}$$

Thus Axiom (74) states the three mutually-exclusive outcomes of placing the chessman, Axiom (78) states that an occurrence of the Place event invokes the three conflicting secondary events, Place1, Place2, and Place4, Axiom (79) associates a conditional probability judgment with each of them, and Axiom (72) translates these judgments into qualitative event preferences.

Proposition $\Theta_{11} \approx_2 Two(2)$.

Proof *Let* M *be a preferred*$_2$ *model of* Θ_{11}*. Then the Place event occurs and invokes the three secondary events Place1, Place2, and Place4 at time* 1 *in* M *(axioms (78), (80)). All four events are supported at time* 1 *in* M *(axioms (17), (73),(75)-(77),(80)). But, in view of their effects they cannot all succeed at time* 1 *in* M *(axioms (1), (74), (75)-(77)). Now the preference atoms* $Pref(Place2, Place4)(1)$, $Pref(Place4, Place1)(1)$, *and* $Pref(Place2, Place1)(1)$ *are all true in* M *(axioms (72), (79)) and these are the only preference atoms with temporal index* 1 *which are defined in* M *(Definition 14.1). Consequently, at time* 1 *in* M*, the three events Place2, Place4, and Place1 have preference ranks* 1, 2 *and* 3 *respectively, and Place is non-preferential. However, as Place invokes the other three* *events and is the only event which does so ((78), Definition 14.1), they can succeed only if it does (Axiom (18)). So it follows that Place succeeds at time* 1 *in* M *(Definition 14.2(b)). The success of Place2 at time* 1 *is consistent with the success of Place (axioms (1), (74), (76)). Consequently, given its preference rank, Place2 succeeds at time* 1 *in* M *(Definition 14.2(a)), with the effect that* $Two(2)$ *is true in* M *(axioms (1), (76)).*

6. Philosophical Justification

In this section the justification of formal theories of prediction is discussed and a justification of the theory of events is given.

Goodman (1954, §III.2) discusses the related problem of the justification of formal theories of enumerative induction, and suggests that we start by considering how we justify a deductive inference. Clearly we can do so by showing that it conforms to a set of valid general rules of deduction.[29] But then the question arises as to how we justify the rules themselves. To suggest that we do so by appealing to some more fundamental "underlying" rules simply postpones the question and so invites a regress.[30] But if we cannot give a foundational justification of deduction, then how can we proceed? Goodman suggests that we can do so by showing that the rules for deduction conform with particular deductive inferences which we actually make and sanction. This is circular, but virtuously so: "The point is that rules and particular inferences alike are justified by being brought into agreement with one another. *A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend.* The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either" (p. 64). Note however that there is no general consensus as to what counts as a valid deductive argument. Classical logicians accept the inference from $\neg\neg\phi$ to ϕ as valid, Intuitionists do not. Similarly, Intuitionists accept the inference from ϕ to $\psi \supset \phi$ as valid, but Relevantists do not. These differences can be explained by the fact that those concerned are attempting to formalize different notions of validity; as there tends to be agreement among those who agree on a given intuitive notion of deduction.[31] This suggests that a justification of a given form of deductive inference will be partly philosophical and partly empirical, as it will consist of an analysis of the concept of deductive validity and the consideration of examples of deductive inference.

Similarly then, in order to justify a formal theory of prediction, we should seek to show, by means of a combination of conceptual analysis and empirical evidence, that its predictions agree with those which we actually make and consider to be reasonable. In the case of a logico-pragmatic theory this amounts to arguing that its pragmatics is appropriate.

29. Or, equally, that it conforms to a notion of entailment defined in terms of a given formal semantics, such as the Tarski semantics for classical predicate logic.

30. Given a formal semantics and accompanying notion of entailment, we can justify a set of rules by proving that they are both sound and complete relative to the semantics. But then again the question arises as to how the semantics and notion of entailment are justified.

31. For example, a valid argument is one in which, if you accept the premises, then you must (on pain of inconsistency) accept the conclusion; a valid argument is one which admits of a constructive proof of the conclusion given the premises; a valid argument is one in which the premises are all required in order to establish the conclusion.

Let Θ be a logico-pragmatic theory of prediction. Then an *extension* of Θ is any theory Θ' which includes Θ and which satisfies certain stated restrictions. For instance, we might require that Θ' is a *semantic extension* of Θ, meaning that the additional axioms in Θ' are intended to be interpreted semantically (rather than pragmatically); thus, for example, Θ_5 is a semantic extension of Θ_{Inv}. We can tentatively define the *intended models* of a given extension Θ' of Θ to be those models of Θ' which accord with our expectations given Θ'.[32] We can then say that a pragmatics is *pragmatically sound* for a class of extensions of Θ if, for every extension Θ' in the class, the pragmatics selects all of the intended models of Θ', and that the pragmatics is *pragmatically complete* for a class of extensions of Θ if, for every extension Θ' in the class, the pragmatics selects only the intended models of Θ'. Thus pragmatic soundness ensures that the theory produces only those predictions which we would consider to be reasonable, and pragmatic completeness ensures that the theory produces all such predictions.

In the case of event theories it seems to be appropriate to concentrate on the pragmatic soundness and completeness of preference$_1$ for semantic extensions of Θ_{Ind}; for, given that preference$_1$ does have these properties, the pragmatic soundness and completeness of preference$_2$ for extensions of Θ_{Ind} which include event preferences but are otherwise semantic seems to be uncontentious.

In order to argue for the pragmatic soundness of preference$_1$ we have to show that the restrictions that it imposes are all necessary. Now the intuitive notion of prediction that event theories endeavour to formalize is that of a context-dependent activity. Prediction takes place at a point in time, "the present", the past is considered to be fixed and the future is considered to be open. Moreover prediction should be based on all and only the available evidence. It then consists of making regularity-based speculations about change given the context, and then (quietly) assuming that facts which are not affected by the changes will persist. These properties are captured by the restrictions imposed by preference$_1$. Location in time and the idea of a closed past and open future are captured by preferring models of an event theory in which it is interpreted chronologically. The restriction of the evidential context and subsequent preference for change over inertia is captured by prioritorized minimaximization at the present time point; first restricting the evidential context to that required by the earlier interpretation of the theory, then assuming whenever possible that events succeed, and finally assuming wherever possible that facts persist. Finally, the physical-theoretical distinction is necessary when a theory contains both kinds of relations, as inertia is a property of the physical world and so should be restricted to physical relations. The need for doing so is illustrated by Example 2.

The need to maximize inertia assumptions chronologically when attempting to represent inductive reasoning about inertia was already clear to Hanks and McDermott (1987) in their discussion of the Yale Shooting Problem. Moreover, the inductive extension of this example (in which the events are inductive rather than deductive), given as Example 1, illustrates the need to prefer change to inertia at any given time point. However, it might be objected that it is not clear that earlier events should succeed in favour of later ones; that is, that success

32. The question of what counts as an intended model of a theory can be a vexed one; see, for example, the discussions by Sandewall (1994, pp. 68-69) and Collins, Hall, and Paul (2004, pp. 32-39). However, in the following discussion, it is sufficient to focus on cases where there is general agreement on expected outcomes.

assumptions should be made chronologically. But if we consider the extension of Example 1 in which a second, later, shot occurs (as in the discussion of Θ'_1 following the example), then it seems clear that we expect that the first shot will prove fatal, with the consequence that the second shot fails because Fred is already dead when it occurs; although, of course, if the first shot were to fail for some independent reason, then we would expect the second shot to succeed. When predicting that the first shot will succeed we do not consider that its success may be jeopardized by a later shot. To allow later events to influence earlier events in this way would be to allow a mysterious form of backwards causation.

A long-standing objection to the chronological assumption of inertia is that this does not fit with the generation of explanations. The standard example of this is Kautz's (1986) Stolen Car Problem; where a car is parked, left unattended, and discovered to be stolen at some later point in time. It has since become part of the folklore that the chronological assumption of inertia is pragmatically unsound because it results in the conclusion that the car was stolen just before this was discovered, when, intuitively, it seems reasonable to conclude that the car could have been stolen at any earlier point at which it was unattended. In consequence, proponents of the chronological maximization of inertia have been prompted to qualify its application; for example, Sandewall (1994) suggests that certain *occluded* relations should be exempt from the law of inertia, at least for certain intervals of time.[33] However it seems to me that there is a better response, namely to argue that examples involving explanation are irrelevant when considering the pragmatic soundness of a theory of prediction because prediction and explanation are different forms of reasoning (Bell, 1998).[34]

Another long-standing objection to the chronological assumption of inertia involves non-determinism. Thus it is claimed, as illustrated by Example 6, that the chronological minimization of inertia can determine the outcome of non-deterministic events, thereby eliminating intended models. However, rather than seek to weaken the preference criterion, I suggest that we should take care to represent the causal structure of non-deterministic events (the fact that they introduce branching histories) correctly; as illustrated by Example 7.

On the basis of these considerations it seems reasonable to conjecture that preference$_1$ is pragmatically sound.

33. This idea is different from the physical-theoretical distinction. To say that a fact is occluded at a point in time is, indeed, to say that it is not subject to the law of inertia at that point in time. However, the occluded fact could naturally be classified as a physical one, such as those represented by the *At* relation in Example 2.

34. Prediction is a form of inductive reasoning; given an epistemic context, the task is to produce reasonable conclusions on the basis of it. By contrast, explanation is a form of abductive reasoning; given a conclusion and an epistemic context which does not imply it, the task is to generate appropriate explanations of the conclusion. Doing so when reasoning about events and their effects involves extending the epistemic context in appropriate ways, so that the conclusion can be induced (predicted) from each of the extensions. In order to ensure that an explanation is appropriate it is reasonable to require that an extension should be such that some event therein causes the conclusion. Consequently I suggest (Bell, 2001, 'Causal counterfactuals') that explanations are best dealt with counterfactually in the more comprehensive setting provided by the theory of causation sketched in Footnote 4. On this view, an event occurrence ϵ together with conditions ϕ explain ψ at a world w iff ϵ causes ψ at all closest $\epsilon \wedge \phi$-worlds to w. Thus, in the case of the Stolen Car Problem, the occurrence of an additional steal event at any time point in the interval during which the car is unattended is a cause of, and so explains, the absence of the car at the next time point.

In order to argue for the pragmatic completeness of preference$_1$ we have to show that the restrictions that it imposes are sufficient. Examples involving ramifications, such as Lifschitz's (1990) lamp-circuit and Baker's (1991) ice-cream eating pedestrian appear to show otherwise. The point of these examples is illustrated by Example 3 where unintended models are selected by the pragmatics. However, I suggest that the problem lies not with the definition of preference$_1$ but with the fact that we need to represent the causal nature of ramifications (as arising from events which are contemporaneously caused (ultimately) by primary events) correctly; as illustrated by Example 4.[35]

In the absence of counterexamples it seems reasonable to conjecture that preference$_1$ is pragmatically complete; although, of course, it is always possible that some new example will show that preference$_1$ is too liberal. If so, then this will not be disastrous for the theory proposed here. On analysis, the example will reveal some further property of prediction which is not captured by preference$_1$, and preference$_1$ can be refined accordingly. In Goodman's terms, this would simply be part of the delicate process of bringing theory and practice into agreement.[36]

Rather than seeking for justifications of this kind, Sandewall (1994) proposes a radically different methodology. He begins by suggesting a series of *ontological characteristics* of instances of predictive reasoning. These include context-free inertia (**I**), alternative results (**A**), ramification (**D**), concurrency (**C**), surprises (**S**), and normality (**N**). Examples of reasoning which include several of these characteristics can be classified as belonging to the corresponding family; for example, the inductive version of the Yale Shooting Problem (Example 1 above) belongs to the family **IN**, as it involves reasoning about inertia and about the normal outcome of the shoot event. Sandewall then proposes a formal pragmatics, the *trajectory semantics*, which *defines* the class of *intended* models for any given example of the family **IAD**, and uses this to prove the correctness of (to provide *validations* for) various formal pragmatics for theories expressed in appropriate formal languages. For example, a simple form of chronological minimization called PCM is proved to be valid for the family **IAD**; by showing that, for any given example of the **IAD** family, PCM selects exactly the same models as the trajectory semantics does. Thus the *range of applicability* of PCM is established; that is, PCM is proved to be applicable for all instances of prediction which have the characteristics of the **IAD** family.

35. The discussion of the gears example, culminating in Example 5, shows that the theory of events is not restricted by the presence of symmetric constraints, but rather that it can be used to represent our reasoning about the direction of causation among symmetrically constrained events.

36. A sceptical reader might object that all I have done is to show that the theory works for a few "toy" examples. However, this is to overlook the care taken to represent prediction accurately, and to misunderstand the motivation behind the choice of examples. The examples referred to in the justification given above were not chosen because they are easily represented, but because they are well-know benchmark examples which are specifically designed to probe for weaknesses; in Goodman's (1954, p. 18) terms, they represent "clinically pure cases that . . . display to best advantage the symptoms of a widespread and destructive malady". So the fact that a theory represents them correctly provides significant empirical evidence in its favour. If the representations are also intuitively convincing, then they provide significant evidence in favour of the conceptual basis of the theory.

A related objection is that the theory has only been shown to work for small-scale examples, and there is no guarantee that it will "scale up" easily to larger examples. But I know of no inherent limitations of scale. In particular, the definitions of events (their preconditions and effects) can be extended in a modular way, and the fact that events are inductive, rather than deductive, means that a given theory can be extended such that additional events occur without fear of contradiction.

However this methodology is limited in two ways.

Firstly, there is the question of how the trajectory semantics itself is justified. It provides a formal definition of the intended models of any given example (of the **IAD** family), but how can we be sure that the models that *it* selects for any given example correspond to the ones that *we* would select? It is not sufficient to simply *stipulate* that this is the case. It is possible to justify a formal pragmatics by proving it to be equivalent to another (as in the case of PCM and trajectory semantics), however at some point a formal pragmatics has to be squared with our intuitions by means of an argument of the kind employed above. Thus the best that we can hope for is a *thesis* relating a formal theory of prediction and our intuitive notion of it.[37] As Kripke once remarked, there is no mathematical substitute for philosophy.

Secondly, as the trajectory semantics is restricted to the **IAD** family (essentially to STRIPS events), it would need to be extended to the **IADCSN** family before it could be applied to event theories.

Nevertheless, it would be worthwhile to attempt to undertake a mathematical assessment of event theories relative to some other formal theory; such as the above extension of trajectory semantics. If possible, mathematical investigations of this kind provide an additional means of justifying formal theories. If the attempt to prove an equivalence between two theories fails, then this will typically highlight inadequacies in one or both of them, and so will suggest that the intuitions behind them need to be refined. Alternatively, if it is possible to prove that the theories are equivalent, then this would provide mutually-supporting evidence for the robustness of the intuitions underlying each of them; as it would suggest that, despite appearances to the contrary, the two formal theories capture the same properties of our intuitive notion of predictive causal reasoning.[38] There is no philosophical substitute for mathematics.

7. Related Work

The theory of natural events presented in this paper has been developed over many years, and earlier versions of parts of it have appeared elsewhere. These earlier fragments have been revised and combined here into a unified whole.

An earlier version of the theory of inductive events (Section 3) was suggested in previous research (Bell, 1998), and its model-building implementation was discussed by White, Bell, and Hodges (1998). The theory has its intellectual origins in the work of McCarthy (1980, 1986), who suggested that circumscription could be used to approach the qualification problem and the frame problem. This proposal was developed by Shoham (1988), who introduced the notion of *chronological minimization* in a classical, temporal, modal language. Shoham's theory offers the promise of a simple and intuitive approach to the two problems. However, his theory is limited in many ways. In particular, it is propositional and has no fact-event distinction, so it is not possible to state general axioms for change and

37. Similarly, the Chuch-Turing thesis is a thesis, rather than a theorem, as it claims an equivalence between an informal intuitive notion, effective computability, and a formal theory of computability (recursive functions, Turing machines).

38. Just as the equivalence results between the rival formalizations of effective computability (recursive functions, Turing Machines, etc.) provide mutually-supportive evidence of the soundness of the intuitions which underlie each of them.

inertia in it. Also, Shoham requires that his theories meet a number of syntactic restrictions, including the restriction that no two *causal rules* conflict (see his Definition 4.7(7)). His reason for doing so is to ensure that his theories are deterministic, so that the process of chronological minimization interprets them correctly. But this makes it impossible to express problems involving both (inductive) change and inertia, such as the inductive version of the Yale Shooting Problem (discussed in Section 3 above) in his theory. Lifting the restriction would allow the problem to be expressed, but the chronological minimization of the theory would have the counter-intuitive result that Fred remained alive after the shooting in one class of models, and died in another. In short, chronological minimization is too simple. The theory of inductive events can thus be thought of as a generalization and refinement of Shoham's theory which fulfills its promise.

The first general common sense theory of change and inertia was proposed by Lifschitz (1987). The axioms that he defines (in the Situation Calculus) are substantially different from those of the theory of inductive events. However an important similarity is his restriction of his inertia axiom, on the basis of a distinction between primitive and defined fluents (the Situation Calculus counterparts of object relations). He later (Lifschitz, 1990, p. 371) says that this distinction should be regarded as a technical trick, and suggests an alternative, more principled, distinction based on frames (McCarthy & Hayes, 1969). However it seems that his primitive-defined distinction can be justified by identifying it with the physical-theoretical distinction introduced in Section 3.[39]

Secondary events (Section 4) were suggested in previous research (Bell, 1999, 2000). Their use in the representation of ramifications should be compared with Thielscher's (1997) treatment. Thielscher views the problem of ramifications as a logical one, which arises (as discussed in Section 4) because of the lack of "causal directedness" in material conditionals. His solution starts with a deductive STRIPS-like representation of events. The ramifications of an event are then brought about by applying a series of *causal constraints* until a stable state is reached. Causal constraints can be thought of as directed conditionals between two single effects, stating the circumstances under which the first *causes* the second. Thus, the problem posed by Example 3 is solved by having Ollie at $L2$ as a direct effect of moving from $L1$ to $L2$, and having the fact that he is at $L2$ and is holding the block *cause* the additional effect of its being at $L2$ also. While this may have the same effect as the invocation of secondary events in some cases, the two approaches are radically different. Thielscher's causal constraints are deductive in nature, so, once begun, their application runs to its conclusion without possible interruption. There is thus no possibility of representing failure at any stage of the indirect-effect-propagation process. So, as it stands, his solution is limited to deterministic deductive events, which either occur in isolation or which do not (directly or indirectly) conflict with each other. Sandewall's (1996) Causal Propagation Semantics is similar to Thielscher's approach and suffers from the same limitations. By contrast, on my approach, success is propagated along an invocation chain, producing the associated effects, but at some point an event may fail, in which case the propagation terminates; as illustrated

39. Readers familiar with Goodman's (1954, Ch. 3) paradox will know that the logical complexity of predicates is relative to a choice of language. But here, like Quine (Footnote 12), we can appeal to ordinary language and its scientific refinements, and refrain from venturing into the fly-bottle. (Wittgenstein, 1953, §309: "What is your aim in philosophy?—To shew the fly the way out of the fly-bottle".)

by Example 4 above. This representation of ramifications can thus be freely combined with non-deterministic events and with conflicting events.

Several formal theories of events employ a primitive causation relation; notable examples are the $\mathcal{A}$-language family originating with Gelfond and Lifschitz (1993) and Lin's (1995) extension of (Toronto) Situation Calculus (Lin & Reiter, 1994). The appearance of an unanalyzed causation relation in a formal theory of events seems to beg the question; as an appeal is made to a more complex notion (causation) in order to give an analysis of a simpler one (change). However, rather than view the causation relation as an appeal to full-blown causation (see Footnote 4), it is better to view it as a means of encoding detailed causal knowledge. In Lin's theory the ternary relation $\mathit{Caused}(p, v, s)$ "is true if the fluent p is caused (by something unspecified) to have the truth value v in the situation s" (p. 1986). In keeping with this interpretation, two axioms are given:

$$\mathit{Caused}(p, \mathit{true}, s) \supset \mathit{Holds}(p, s) \ ,$$
$$\mathit{Caused}(p, \mathit{false}, s) \supset \neg \mathit{Holds}(p, s) \ .$$

Thus if fluent p is caused to have the value true (false) in situation s, then p holds (does not hold) in situation s. By way of illustration, Lin discusses the example of a spring loaded suitcase. If both of its locks are up, the suitcase opens. Initially the suitcase is closed, one lock is up, the other is down, and the down-lock is flipped. Naturally we expect that the suitcase will open as a result. However, attempts to formalize this example using an ordinary domain axiom and inertia fail; as in the lamp circuit example, the other lock remains up in the intended models, but there are also unintended models in which the suitcase remains closed and the other lock changes position as a result. Consequently Lin proposes a causal domain axiom:

$$\mathit{up}(L1, s) \wedge \mathit{up}(L2, s) \supset \mathit{Caused}(\mathit{open}, \mathit{true}, s) \ ; \tag{81}$$

which states that if locks $L1$ and $L2$ are both up in situation s, then the suitcase is *Caused* to be open in s. The intention is that axioms such as this are interpreted "causally" (positively), and this effect is achieved by circumscribing the *Caused* relation and by adding an inertia axiom which states that fluents which are not *Caused* to change persist. Thus according to Lin's account, ramifications arise as a result of causally-directed domain axioms such as (81), in which facts cause other facts to change. Indeed, Lin (p. 1986) says of (81) that:

> [T]he physical spring loaded mechanism behind the causal rule has been abstracted away. For all we care, it may just as well be that the device is not made of spring, but of bombs that will blow open the suitcase each time the two locks are in the up position. It then seems natural to say that the fluent *open* is caused to be true by the fact that the two locks are both in the up position.

However this seems odd from a common sense point of view, which has it that events cause change and that facts are otherwise inert. Moreover, as the actual cause has been abstracted away, it is difficult to see how Lin's account could be extended in order to include the treatment of qualifications; for example, the suitcase may fail to open when both locks are up because the mechanism is rusty, someone is sitting on the suitcase, etc. By contrast, it is obvious how to represent the problem using primary and secondary events; in a context in which one lock is up and the other is down, flipping the down-lock invokes a secondary event which, if it succeeds, has the additional effect that the suitcase is open. Moreover, as noted above, the introduction of secondary events makes it possible to represent more complex examples of ramifications, such as those of Example 4.

Secondary events should also be compared with the *natural actions* proposed by Reiter (1996), Lin (1998), and Pinto (1998). In addition to actions initiated by agents, they suggest that there are also natural actions which arise due to the nature of the world (the Laws of Nature, etc.), and which are guaranteed to succeed. Their use is illustrated by Pinto's treatment of Lifschitz's (1990) lamp-circuit example. If the switches in a circuit are in the same position (both up or both down) then the lamp is on, otherwise it is off. Initially the switches are in opposite positions and one of them is flipped. Clearly we expect the lamp to be on as a result. However attempts to formalize this example using domain axioms with an inertia axiom fail. In the intended models the other switch retains its position by inertia and so the lamp comes on as a result of the flip event. However there are also unintended models in which the lamp remains off by inertia and consequently the other switch mysteriously changes position as a result; this problem is essentially the same as that posed by Example 3. In Pinto's treatment of the problem, the agent flipping a switch results in the natural action of current flowing in the circuit, and this in turn results in the lamp being on. The natural action of the current flowing resolves the conflict between alternative uses of the inertia axiom by bringing about an intermediary state in which both switches are up (as the flowing current does not affect their position) and in which the inertia axiom can only be applied in the intended way (as the flowing current is guaranteed to have the effect that the lamp is on). As this proposal aims to solve the causal-directedness problem by introducing additional events it can, perhaps, be seen as lying somewhere between Thielscher's "logical" approach and my "representational" approach. Unlike natural events, natural actions are deductive. So it is difficult to see how natural actions could be used to represent qualifications. For example, the natural action of the current flowing is guaranteed to succeed (turning the lamp on), but in reality the current might not flow if the wire loses its conductivity, is cut, etc. But clearly this complication can be represented using natural events. In a context in which the switches are in opposite positions, flipping a switch invokes a secondary event which, if it succeeds, results in the lamp being on; the lamp's coming on is thus a defeasible secondary effect of flipping the switch. Moreover, while natural actions are independent of the actions of agents, secondary events are causally dependent on (at least one of) the events which invoke them.

Event preferences (Section 5) were introduced in previous research (Bell, 2001, 'Simultaneous events'). The presentation has been substantially improved, and, as the examples show, the combination of event preferences and secondary events provides interesting new possibilities.

Recently Vo and Foo (2005) have suggested an inductive theory of reasoning about action which, they suggest, provides the basis for a unified solution to the frame, qualification, and ramification problems. Their theory is based on the theory of argumentation developed by Bondarenko, Dung, Kowalski, and Toni (1997), and so it differs radically from mine in terms of its technical realization. At the conceptual level there are interesting similarities. Like me, Vo and Foo suggest that event occurrences should be minimized and that the inertia of fluents (their counterpart of object-relations) should be maximized (p. 448). Furthermore, in order to integrate their solution to the frame and qualification problems, they, in effect (p. 493), adopt the principle that change is preferred to inertia (Bell, 1998). However, in their theory there is no suggestion that the minimization and maximization should be done chronologically. Indeed, they suggest that not doing so is a strength of their approach as it enables them to provide explanations by "reasoning backwards" (or antichronologically) in examples such as the Stolen Car Problem. However, as reasoning chronologically is an essential feature of predictive reasoning, I suspect that this will prove to be a problem for Vo and Foo's theory. Instead I suggest (Footnote 34) that explanations should be treated as counterfactual causes.

8. Conclusion

I began by arguing that Deductionism is a logical mistake, and have made a case for Inductionism. This began with the basic theory of inductive events, which provides the basis for an integrated solution to the qualification and frame problems. I then introduced the distinction between primary and secondary events in order to represent the causal structure of natural events, thereby providing the basis for a solution to the ramification problem and to the problem of representing non-determinism. Finally, I introduced event preferences, which can be used to express defeasible preferences over the outcomes of conflicting simultaneous events. The development of the theory illustrates the benefit of starting off on the right foot, with inductive, rather than deductive, events. The basic theory is simple and intuitive, and its extensions require no more than the addition of a few axioms and the refinement of one clause of the formal pragmatics.

In simple cases there may be little to choose between deductive and inductive theories of events; as both can produce predictions which are wrong. However the inductive representation of natural events is more accurate because it reflects their defeasibility. The representation of inductive events makes it possible to define primary and secondary events, and defeasible event preferences. These in turn make it possible to give accurate representations of ramifications, non-determinism, and conflicting events. And this in turn makes it possible to represent complex cases accurately. For instance, in Example 8, if any two of the stooges attempt to move to a location, then one is expected to succeed, however if all three do, then none succeeds. As always with inductive events, the example can be elaborated. Thus if two stooges attempt the move and the preferred one fails for some independent reason (he slips, say), then the other stooge normally succeeds (unless he also slips). Or in the case where all three attempt the move, if Ollie slips, then Stan is expected to succeed. But if he also slips, then Charlie is expected to succeed; but may also slip. Moreover, the example can readily be combined with others. For example, each stooge could carry a stack of blocks; where each block moves only if the block beneath it, or the stooge holding it, moves. I know of no other theory of events which can represent reasoning of this subtlety and complexity.

In future work, the model-building implementation of event theories will be investigated. The general idea was outlined in previous research (Bell, 1995), and White et al. (1998) describe the implementation of (an earlier version of) the theory of inductive events. Essentially the idea is to build finite initial parts of the representative preferred models of a given event theory chronologically, as suggested by the informal discussion of the Yale Shooting Problem in Section 3.

As suggested in the introduction (and in Footnote 4), the theory of natural events forms part of a larger theory of causation (Bell, 2004, 2006, 2008). Event theories are used to represent detailed, regularity-based, causal knowledge about events; expressed in the form of preconditions and effects, invocations, and event preferences. This is used as the basis of a general definition of sufficient causation, which is combined with a refinement of Lewis's (1986, Ch. 21) counterfactual-dependence condition to give the definition of causation.

Acknowledgments

I am grateful to everyone who has commented on this work, especially the anonymous referees, Wilfrid Hodges, Robert Miller, Edmund Robinson, Murray Shanahan, and Graham White.

References

Baker, A. B. (1991). Nonmonotonic reasoning in the framework of the situation calculus. *Artificial Intelligence*, *49*, 5–23.

Barringer, H., Cheng, J., & Jones, C. (1984). A logic covering undefinedness in program proofs. *Acta Informatica*, *21*, 251–269.

Bell, J. (2001). Causal counterfactuals. In *Working Notes of Common Sense 2001.* Available at: www.cs.nyu.edu/cs/faculty/davise/commonsense01/.

Bell, J. (1995). Pragmatic reasoning; a model-based theory. In Pólos, L., & Masuch, M. (Eds.), *Applied Logic; How, What and Why*, pp. 1–27. Kluwer. A selection of papers from the *Applied Logic Conference*, Amsterdam, 1992.

Bell, J. (1996). A model-based approach to predictive causal reasoning. In Doherty (Doherty, 1996), pp. 169–195.

Bell, J. (1998). Chronological minimization and explanation. In Miller, R., & Shanahan, M. (Eds.), *Working papers of Common Sense '98; the fourth symposium on logical formalizations of common sense reasoning.*

Bell, J. (1999). Primary and secondary events. In Thielscher, M. (Ed.), *Working notes of the 4th workshop on Nonmonotonic Reasoning, Action and Change*, pp. 65–72.

Bell, J. (2000). Primary and secondary events. *Electronic Transactions on Artificial Intelligence* discussion paper. Available at: www.ida.liu.se/ext/etai/rac/. Subsequent version available at: www.dcs.qmul.ac.uk/~jb, 2001.

Bell, J. (2001). Simultaneous events: Conflicts and preferences. In *Proceedings of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2001)*, pp. 714–725. Springer.

Bell, J. (2004). Causation and causal conditionals. In *Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning (KR 2004)*, pp. 2–11. AAAI Press.

Bell, J. (2006). Causation as production. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI 2006)*, pp. 327–331. IOS Press.

Bell, J. (2008). A common sense theory of causation. Forthcoming.

Blackburne, S. (1994). *The Oxford Dictionary of Philosophy.* Oxford University Press, Oxford.

Bochvar, D. A. (1939). On a three-valued logical calculus and its application to the analysis of contradictories. *Matematiceskij sbornik*, *4*.

Bondarenko, A., Dung, P. M., Kowalski, R. A., & Toni, F. (1997). An abstract argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, *93*, 63–101.

Collins, J., Hall, N., & Paul, L. A. (Eds.). (2004). *Causation and Counterfactuals.* MIT Press, Cambridge, Massachusetts.

Davidson, D. (1980). *Essays on Actions and Events.* Oxford University Press, Oxford.

Denecker, M., Dupré, D. T., & Belleghem, K. V. (1998). An inductive definition approach to ramifications. *Electronic Transactions on Artificial Intelligence*, *2*, 25–67.

Doherty, P. (Ed.). (1996). *Partiality, Modality, and Nonmonotonicity.* CSLI Publications, CSLI, Stanford University, Palo Alto.

Fikes, R. E., & Nilsson, N. J. (1971). STRIPS, a new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, *2*, 189–208.

Gelfond, M., & Lifschitz, V. (1993). Representing action and change by logic programs. *Journal of Logic Programming*, *17*, 301–322.

Gelfond, M., Lifschitz, V., & Rabinov, A. (1991). What are the limitations of the situation calculus?. In Boyer, R. (Ed.), *Essays in Honour of Woody Bledsoe*, pp. 167–179. Kluwer Academic Publishers.

Ginsberg, M. L., & Smith, D. E. (1988). Reasoning about action II: The qualification problem. *Artificial Intelligence*, *35*, 311–342.

Girard, J.-Y. (1987). Linear logic. *Theoretical Computer Science*, *50:1*, 1–102.

Goodman, N. (1954). *Fact, Fiction, and Forecast.* Harvard University Press, Cambridge Mass. References are to the 4th edition, 1983.

Hanks, S., & McDermott, D. (1987). Nonmonotonic logic and temporal projection. *Artificial Intelligence*, *33 (3)*, 379–412.

Hume, D. (1739). *A Treatise of Human Nature.* Clarendon Press, Oxford. References are to the Clarendon edition, 1978.

Hume, D. (1777). *An Enquiry Concerning Human Understanding.* Clarendon Press, Oxford. References are to the Clarendon edition, 1975.

Kautz, H. (1986). The logic of persistence. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI-86)*, pp. 401–405.

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

Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, *72*, 690–716.

Lewis, D. (1986). *Philosophical Papers*, Vol. II. Oxford University Press, Oxford.

Lifschitz, V. (1987). Formal theories of action. In Brown, F. (Ed.), *The Frame Problem in Artificial Intelligence*, pp. 35–58. Morgan Kaufmann.

Lifschitz, V. (1990). Frames in the space of situations. *Artificial Intelligence*, *46*, 365–376.

Lin, F., & Reiter, R. (1994). State constraints revisited. *Journal on Logic and Computation*, *4(5)*, 655–678.

Lin, F. (1995). Embracing causality in specifying the indirect effects of actions. In *Proceedings of the 14th International Joint Conference Artificial Intelligence (IJCAI '95)*, pp. 1985–1991.

Lin, F. (1998). On the relationships between static and dynamic causal rules in the situation calculus. In Ortiz, C. (Ed.), *Working Papers of the AAAI'98 Spring Symposium on Prospects for a Commonsense Theory of Causation*, pp. 38–43.

Mackie, J. L. (1974). *The Cement of the Universe*. Clarendon Press, Oxford.

Mackie, J. L. (1975). Causes and conditions. In Sosa, E. (Ed.), *Causes and Conditionals*, pp. 15–38. Oxford University Press, Oxford. First published in *American Philosophical Quarterly* 2, 1965, pages 245-264.

McCarthy, J. (1977). Epistemological problems of artificial intelligence. In *Proceedings of the 5th International Joint Conference Artificial Intelligence (IJCAI '95)*, pp. 1038–1044.

McCarthy, J. (1980). Circumscription – a form of nonmonotonic reasoning. *Artificial Intelligence*, *13*, 27– 39.

McCarthy, J. (1986). Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, *28*, 89–116.

McCarthy, J., & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In Michie, D., & Meltzer, B. (Eds.), *Machine Intelligence 4*, pp. 463–502. Edinburgh University Press.

Mill, J. S. (1898). *A System of Logic*. Longmans, London. References are to the 8th edition, 1941.

Pinto, J. (1998). Causality in theories of action. In Miller, R., & Shanahan, M. (Eds.), *Working papers of the 4th Symposium on Logical Formalizations of Commonsense Reasoning*, pp. 349–364.

Quine, W. V. (1995). *From Stimulus to Science*. Harvard University Press, Cambridge, Massachusetts.

Reiter, R. (1996). Natural actions, concurrency and continuous time in the situation calculus. In *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR'96.)*, pp. 2–13.

Reiter, R. (2001). *Knowledge in Action; Logical Foundations for Specifying and Implementing Dynamical Systems*. MIT Press, Cambridge, Mass.

Rosser, J., & Turquette, A. (1952). *Many-valued Logics*. North-Holland, Amsterdam.

Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelian Society*, *13*, 1–26.

Russell, S., & Norvig, P. (2003). *Artificial Intelligence; A Modern Approach*. Prentice Hall, New Jersey. 2nd Edition.

Sandewall, E. (1994). *Features and Fluents; The Representation of Knowledge About Dynamical Systems*. Oxford University Press, Oxford.

Sandewall, E. (1996). Assessments of ramification methods that use domain constraints. In *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR'96.)*, pp. 99–110.

Shanahan, M. (1997). *Solving the Frame Problem; A Mathematical Investigation of the Common Sense Law of Inertia*. MIT Press, Cambridge, Mass.

Shoham, Y. (1988). *Reasoning About Change*. MIT Press, Cambridge Mass.

Taylor, R. (1975). The metaphysics of causation. In Sosa, E. (Ed.), *Causation and Conditionals*, pp. 39–43. Oxford University Press.

Thielscher, M. (1997). Ramification and causality. *Artificial Intelligence*, *98*, 317–364.

Vo, Q. B., & Foo, N. Y. (2005). Reasoning about action: An argumentation-theoretic approach. *Journal of Artificial Intelligence Research*, *24*, 465–518.

White, G., Bell, J., & Hodges, W. (1998). Building models of prediction theories. In Cohn, A., Schubert, L., & Shapiro, S. (Eds.), *Proc. KR'98*, pp. 557–568, San Francisco. Morgan Kaufmann.

Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford.