Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Technical
Memo
TM-90-02

The Myth of Domain-Independent Persistence

Jay C. Weber

June 1990

Deutsches Forschungszentrum für Künstliche Intelligenz
GmbH

Postfach 20 80
D-6750 Kaiserslautern
Tel.: (+49 631) 205-3211/13
Fax: (+49 631) 205-3210

Stuhlsatzenhausweg 3
D-6600 Saarbrücken 11
Tel.: (+49 681) 302-5252
Fax: (+49 681) 302-5341

Deutsches Forschungszentrum für Künstliche Intelligenz

The German Research Center for Artificial Intelligence (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI) with sites in Kaiserslautern und Saarbrücken is a non-profit organization which was founded in 1988 by the shareholder companies ADV/Orga, AEG, IBM, Insiders, Fraunhofer Gesellschaft, GMD, Krupp-Atlas, Mannesmann-Kienzle, Nixdorf, Philips and Siemens. Research projects conducted at the DFKI are funded by the German Ministry for Research and Technology, by the shareholder companies, or by other industrial contracts.

The DFKI conducts application-oriented basic research in the field of artificial intelligence and other related subfields of computer science. The overall goal is to construct *systems with technical knowledge and common sense* which - by using AI methods - implement a problem solution for a selected application area. Currently, there are the following research areas at the DFKI:

- ❏ Intelligent Engineering Systems
- ❏ Intelligent User Interfaces
- ❏ Intelligent Communication Networks
- ❏ Intelligent Cooperative Systems.

The DFKI strives at making its research results available to the scientific community. There exist many contacts to domestic and foreign research institutions, both in academy and industry. The DFKI hosts technology transfer workshops for shareholders and other interested groups in order to inform about the current state of research.

From its beginning, the DFKI has provided an attractive working environment for AI researchers from Germany and from all over the world. The goal is to have a staff of about 100 researchers at the end of the building-up phase.

Prof. Dr. Gerhard Barth
Director

The Myth of Domain-Independent Persistence

Jay C. Weber

DFKI-TM-90-02

A revised version of this Technical Memo called "The Myth of Domain-Independent Persistence" was submitted for inclusion in a book "Advances in Human and Machine Cognition, Volume I: The Frame Problem".

The Myth of Domain-Independent Persistence

Jay C. Weber[†]

Abstract

The *frame problem* can be reduced to the problem of inferring the non-existence of causes for change. This paper concerns how these non-existence inferences are made, and shows how many popular approaches lack generality because they rely on a domain-independent assumption of occurrence omniscience. Also, this paper shows how to represent and use appropriate domain-dependent knowledge in three successively more expressive versions, where the causal theories are deductive, non-monotonic, and statistical.

CONTENTS:

[†]The author is currently at the Lockheed Artificial Intelligence Center, Menlo Park, CA, EMail address: lockheed@polya.stanford.edu

1. Introduction

A property can be reasoned to persist over time by reasoning that every action that could cause it to change did not occur. This reduction is the first step of many solutions to the frame problem, including:

Cause-closure axioms [Haas 1987, Schubert 1989]. This approach supplies a disjunction of possible actions that could cause a property to change. For example, using the situation calculus [McCarthy and Hayes 1969], e.g. where the predicate holds relates properties to world situations, and the function do applies an action to a situation to produce a new situation, the following axiom states that if a gun becomes not loaded, then it must have been unloaded or shot:

$$\forall\sigma,\alpha[(\text{holds}(\text{loaded},\sigma) \wedge \neg\text{holds}(\text{loaded},\text{do}(\alpha,\sigma))) \rightarrow (\alpha{=}\text{unload} \vee \alpha{=}\text{shoot})]\ .$$

When combined with a unique names assumption, these axioms are powerful enough to derive the standard frame axioms, e.g. given that "$(\text{wait} \neq \text{unload}) \wedge (\text{wait} \neq \text{shoot})$", then "$\forall\sigma[\text{holds}(\text{loaded},\sigma) \rightarrow \text{holds}(\text{loaded},\text{do}(\text{wait},\sigma))]$" is a theorem. However, cause-closure axioms do not suffer from the historical combinatorial argument against frame axioms, since there is only one axiom per property.

Minimizing potential causes [Lifschitz 1987, Haugh 1987]. This approach generates the above cause-closure axioms indirectly. Causal connections between actions and properties are represented using the predicate causes, e.g.

$$\text{causes}(\text{unload},\text{not}(\text{loaded})) \wedge \text{causes}(\text{shoot},\text{not}(\text{loaded}))\ ,$$

and are related to property change by the following *Law of Inertia*:

$$\forall\pi,\sigma,\alpha[(\text{holds}(\pi,\sigma) \wedge \neg\text{holds}(\pi,\text{do}(\alpha,\sigma))) \rightarrow \text{causes}(\alpha,\text{not}(\pi))]\ .$$

When causes is minimized (i.e. circumscribed [McCarthy 1980]), cause-closure axioms like the one above become theorems (Lifschitz [1987]). This implicit derivation has the advantage of "automatically" modifying the cause-closure axioms when new causal information is added.

Explicit occurrence [Georgeff 1986, Morgenstern and Stein 1988, Sandewall 1988, Weber 1988]. These approaches relate properties to times using the predicate true[1], and relate *events* to times with the predicate occurs. Event occurrence at time τ only *partially* constrains the proper-

---

[1] The predicate holds is often used to relate properties to times [Allen 1984, Georgeff 1986, Weber 1988], although in order to avoid confusion this paper reserves that predicate for relating properties to *situations*.

ties over τ' (the successor of t in a discrete time line), as opposed to situation calculus actions that completely determine the resulting situation, making explicit occurrence cause-closure axioms slightly different, e.g.

$$\forall\tau[(\text{true}(\text{loaded},\tau) \wedge \neg\text{true}(\text{loaded},\tau')) \rightarrow (\text{occurs}(\text{Unload},\tau) \vee \text{occurs}(\text{Shoot},\tau))] ,$$

These axioms are given explicitly [Georgeff 1986, Morgenstern and Stein 1988] or derived non-monotonically [Sandewall 1988, Weber 1988] in a manner similar to minimizing potential causes. Note that the above event constants are capitalized words; this is to distinguish them from situation calculus actions, which have a much different meaning.

The remaining technical difficulty is how to make appropriate inferences about the occurrence of actions (or events). The traditional approach, which this paper disputes, is that non-occurrences are inferred even when particular information relevant to those occurrences is lacking. For example, the intuition behind the ubiquitous "shooting" scenario [Hanks and McDermott 1986] says that the reasoner should infer that no unload action occurs, and therefore the loaded gun stays loaded. However, axiomatic specifications of the shooting scenario [Hanks and McDermott 1986, Lifschitz 1987, Morgenstern and Stein 1988, etc.] contain neither knowledge about how guns become unloaded, nor contextual knowledge needed to evaluate whether an unload action actually occurred. Ironically, any approach that resolves the in-herent ambiguity must be unsound and even unreliable, since it substitutes a domain-independent mechanism for the relevant domain-dependent knowledge.

Section two exposes the hidden assumptions that allow many approaches to solve the frame problem despite the lack of the necessary domain-dependent details. Section 2.1 shows that approaches to causal reasoning based on the situation calculus trivialize the inference of non-occurrences because of their treatment of action, which implicitly contains an unrealistic assumption that the reasoner is omniscient with respect to occurrences. Section 2.2 shows that a similar criticism applies to several recent explicit occurrence approaches because of their minimization of the predicate occurs, which embodies a similar assumption about occurrence omniscience.

As an alternative approach, Section three suggests how to make inferences about non-occurrence using reasonable domain-dependent knowledge, rather than an omniscience assumption. Section 3.1 presents a purely deductive approach that disproves necessary preconditions of events, which is powerful enough to solve the "Yale Shooting Problem". Section 3.2 generalizes the deductive approach to use non-monotonicity in a domain-dependent manner, which is powerful enough to solve the "parked car" problem. Finally, Section 3.3 presents a statistical approach, which provides a justification for non-monotonic assumptions, as well as a more expressive way of reasoning about belief in properties and occurrences.

2. Domain-Independent Approaches

Hanks and McDermott [1986] show that McCarthy's [1984] non-monotonic approach to the frame problem exhibits an unintuitive temporal ambiguity, demonstrated by the shooting scenario. The standard exposition involves an agent who loads a gun, waits momentarily, and then fires the gun at a potential victim "Fred". The challenge of this scenario is to supply a domain-independent mechanism which produces the intuitive conclusion that the gun stays loaded during the wait, and therefore Fred dies (this paper concentrates on the key inference that the loadedness of the gun persists, thereby ignoring the actual loading of the gun, and morbid inferences about poor Fred).

This problem has inspired a great deal of discussion. The first solutions use elaborate non-monotonic constructions [Shoham 1986, Kautz 1986, Lifschitz 1986], but have been seriously questioned [Haugh 1987, Goodwin 1987] on the grounds that they produce unintuitive results in some scenarios. The next wave of solutions [Lifschitz 1987, Haugh 1987] use non-monotonicity differently (to complete the causal theory rather than the context), but have been criticized for their functional view of action [Morgenstern and Stein 1988, Weber 1988]; this criticism is refined and extended in Section 2.1. Recent solutions allow for a more natural specification of occurrences [Morgenstern and Stein 1988, Sandewall 1988, Weber 1988], but as shown in Section 2.2, they employ an unnatural minimization of occurrences that leads to problematic inferences.

This state of technical confusion is largely due to the traditional but misguided mandate for a domain-independent mechanism to solve the frame problem. The next two sections discuss how the rationality of several popular solutions to the frame problem suffer from their emphasis on domain-independent mechanisms.

2.1 Implicit Occurrence Omniscience

Approaches that are based on the functional view of action *implicitly* make the often unreasonable assumption that the reasoner is omniscient with regard to the occurrence of actions. These approaches include the situation calculus [McCarthy and Hayes 1969], many deductive or algorithmic planners [Fikes and Nilsson 1971, etc.], and many recent logical frameworks [Lifschitz 1987, Ginsberg and Smith 1987, Haugh 1987]. They all assume that the reasoner knows (or is told) *exactly* what has occurred, trivializing the inference of non-occurrences and therefore the frame problem.

As Georgeff [1986] points out, this assumption arises because an action is a (reified) function that maps situations into situations. To prove that some action b did not occur between situation s and situation do(a,s), the reasoner need only prove $a \neq b$. This allows the bizarre inference that any action can be prevented simply by doing something else, i.e. the sun will not set so long as Joshua holds his hands in the air.[2] Despite claims to the contrary [Morgenstern and Stein 1988, Weber 1988], this problem is not an inherent inability to represent simultaneously occurring actions. In fact, it is simple to use a composition operator on actions; e.g. comp(a,b) produces a new action which is the simultaneous performance of a and b. Of course, this new action object must be included in the axioms relating property change and action application (this description can be made simpler by encoding the interactions between a and b). However, even when composite actions are applied to situations, the component actions are necessarily the *only* actions that contribute to the content of the resulting situation. Therefore, when a reasoner applies a ground action term (constant or composite) to a situation, the reasoner claims it knows exactly what occurred. There seem to be few or no domains of interests where this claim of omniscience is justified.

It should be noted that this omniscience assumption is not inherent in the situation calculus, but rather how it is invariably used. The assumption is removed if instead of applying ground actions to situations, the reasoner applies constrained action variables. A constrained variable does not necessarily completely determine the resulting situation, so the omniscience assumption is no longer implicit. Simultaneous events are represented by conjoining constraints. In fact, the use of constrained variables in situation calculus is expressively equivalent to a discrete time line with atomic events [Weber 1989c].

However, removing the omniscience assumption has removed the powerful

[2]The relevance of this biblical reference was discovered by Josh Tenenberg (personal communication).

(albeit inappropriate) mechanism for inferring action non-occurrences, making the frame problem a problem once again. Because of their instrinsic similarity, this is also true of approaches based on time lines and events [Georgeff 1986, McDermott 1982, Morgenstern and Stein 1988, Weber 1988, and in a real sense Allen 1984]. To solve the frame problem, some additional mechanism must be added to infer non-occurrences. The next section examines the popular technique of minimizing the predicate occurs, and shows that like situation calculus solutions, it also imposes an inappropriate assumption of occurrence omniscience.

2.2 Explicit Occurrence Omniscience

McDermott [1982] reluctantly suggested that a non-monotonic implementation of the *sleeping dog strategy*, or "things tend to stay the same", should be added to causal representations based on an explicit time line. A recent adaptation of this suggestion has been to interpret the sleeping dog strategy as "events tend to not occur" and use cause-closure axioms to extend this to properties. This domain-independent assumption about non-occurrence has been implemented by minimizing the predicate occurs, i.e. (modulo multiple minimal models) if it is not known that an event occurs then it is known that it does not occur [Sandewall 1988, Weber 1988]. This approach produces the intuitive reading of the shooting scenario, via the default inference that neither an unload nor a shoot event occurs while the reasoner waits, and therefore by cause-closure, the gun remains loaded.

The approach of Morgenstern and Stein [1988] is similar but with a weaker minimization. They minimize only *unmotivated* occurrences, where an occurrence is motivated when it appears in the consequent of a causal rule whose antecedent is at least partially known. More precisely, a possible element $\langle e,t \rangle$ of occurs is exempt from minimization if and only if in the axiomatic theory there exists a rule "$(\alpha \wedge \beta) \rightarrow occurs(e,t)$", where α is provable. This definition is both semantic and syntactic, requiring special care when constructing axiomatic theories in order to avoid undesirable results. Nevertheless, it also produces the intuitive reading of the shooting scenario because neither unload nor shoot events are motivated.

Minimization of occurrences, however, imposes an assumption of omniscience similar to the one we found objectionable for situation calculus. Whatever events the reasoner cannot prove (or motivate) occur are assumed to not occur by the minimization; this claims that the reasoner knows exactly what occurs at all times. The reasoner should not be able to conclude that an arbitrary gun remains loaded, especially if the gun is in another room, or another country. Valid inferences about occurrences require domain-dependent information about the possibility, feasibility, or probability of particular events at particular times. The next section pursues this domain-dependent paradigm.

3. Domain-Dependent Approaches

The approaches described in the last section were influenced by the traditional conjecture that the frame problem should be solved by some domain-independent mechanism. This stance dates back to the dismissal of frame axioms on combinatorial grounds [McCarthy and Hayes 1969], and seems to be one of the few points of agreement in the causal reasoning literature.

Nevertheless, this section breaks with this tradition by presenting three novel domain-dependent solutions to the frame problem. The first solution demonstrates that some intuitive inferences, in particular the outcome of the shooting scenario, can be derived through the careful specification of a *deductive* theory. The second version demonstrates how to use non-monotonicity for more flexible inference, without exhibiting the problems the previous section described. Finally, the third version shows how the non-monotonic mechanism can be generalized to make inferences about probabilities, based on statistics which capture causal relations.

3.1 Deductive Domain Details

As Georgeff [1986] states, "it has been common to use various default rules, non-monotonic operators, or minimal models to constrain the set of possible event occurrences. However, there are many cases in which this is unnecessary - where we can [deductively] *prove* ... that no events (or effects) of interest could possibly occur." This paper advocates the specification of *necessary preconditions* to make these proofs possible. Such axioms specify properties that are necessarily true for the event to occur, as opposed to the usual interpretation of preconditions as sufficient conditions for an action to imply certain effects [Fikes and Nilsson 1971, McCarthy and Hayes 1969, Shoham 1988, Lifschitz 1987[3]]. Thus an event can be proven not to occur by proving that at least one of its necessary preconditions does not hold.

[3]The preconditions of Lifschitz are also necessary for the action to cause any effects; that is, an action with a false precondition will map a situation to itself.

Thus the reduction from change to occurrence coupled with necessary preconditions solves the frame problem. For example, consider the following version of the causal knowledge needed in the shooting scenario:

a. "A gun will remain loaded unless it is unloaded or shot", i.e.

$$\forall\gamma,\tau[(\mathrm{true}(\mathrm{loaded}(\gamma),\tau)\wedge\neg\mathrm{true}(\mathrm{loaded}(\gamma),\tau'))\rightarrow\exists\omega[\mathrm{occurs}(\mathrm{Unload}(\omega,\gamma),\tau)\vee\mathrm{occurs}(\mathrm{Shoot}(\omega,\gamma),\tau)]]\,.$$

This rule is an explicit occurrence cause-closure axiom, although it contains some notational features not previously discussed: functions that produce properties (loaded(γ) produces the property of the gun γ being loaded), functions that produce events (Unload(ω,γ) is the event of agent ω unloading gun γ) and existentially-quantified parameters in the consequent, generating a potentially infinite list of possible causes.

b. "An agent must be grasping a gun to unload or shoot it", i.e.

$$\forall\omega,\gamma,\tau[(\mathrm{occurs}(\mathrm{Unload}(\omega,\gamma),\tau)\vee\mathrm{occurs}(\mathrm{Shoot}(\omega,\gamma),\tau))\rightarrow\mathrm{true}(\mathrm{grasping}(\omega,\gamma),\tau)]\,.$$

This rule specifies a necessary precondition for both Unload(ω,γ) as well as Shoot(ω,γ) events for any ω and γ.

c. "At most one agent can grasp a gun at a given time", i.e.

$$\forall\omega,\theta,\gamma,\tau[(\mathrm{true}(\mathrm{grasping}(\omega,\gamma),\tau)\wedge\mathrm{true}(\mathrm{grasping}(\theta,\gamma),\tau))\rightarrow\omega=\theta]\,.$$

This rule supplies a *domain constraint* [Allen and Koomen 1983] which will be handy to disprove the necessary precondition in rule "b", i.e. to prove that a particular agent is not grasping the object, it suffices to prove that a different agent is grasping the object.

This version of the shooting scenario also needs axioms for details of the scenario in question:

d. "At the onset, the gun is loaded", i.e. "true(loaded(gun1),0)". This axiom appears in all expositions of the shooting scenario, except that in this case the property loaded is parameterized by the gun in question.

e. "At the onset, the reasoning agent grasps the gun", for example, "true(grasping(me,gun1),0)". This axiom is not typically included in expositions of the shooting scenario, but it is strongly suggested by it, since the agent has just loaded the gun and is about to shoot it.

f. "The reasoning agent then waits briefly, *without unloading or shooting the gun.*", i.e.

$$\neg\mathrm{occurs}(\mathrm{Unload}(\mathrm{me},\mathrm{gun1}),0)\wedge\neg\mathrm{occurs}(\mathrm{Shoot}(\mathrm{me},\mathrm{gun1}),0)\,.$$

The italicized clause has not appeared previously in expositions of the shooting scenario, although it seems to follow from the prose description of the problem; whatever "waits" means in the context of gun handling, it must certainly imply that the agent does not unload or shoot the gun.

From these six axioms the reasoner may deductively infer that the gun persists in being loaded. In the following proof, the *reason* for a proof-step consists of the axiom letters and theorem numbers used to derive it:

Step	Theorem	Reason
1	$\forall\omega[\omega\neq me \rightarrow \neg true(grasping(\omega,gun1),0)]$	c,e
2	$\forall\omega[\omega\neq me \rightarrow (\neg occurs(Shoot(\omega,gun1),0) \wedge \neg occurs(Unload(\omega,gun1),0))]$	1,b
3	$\neg\exists\omega[occurs(Shoot(\omega,gun1),0) \vee occurs(Unload(\omega,gun1),0)]$	2,f
4	$\neg(true(loaded(gun1),0) \wedge \neg true(loaded(gun1),0'))$	3,a
5	$true(loaded(gun1),1)$	4,d

This proof demonstrates that it is possible to deductively derive the intuitive outcome of the shooting scenario simply by being explicit about the knowledge behind our intuitions. An algorithmic interpretation is straightforward: to predict the persistence of a property, the reasoner indexes to all events that could change that property, and then investigates the possibility of any of those events occurring. Because this algorithm is based on deduction, we know it is *sound*: if the domain knowledge is correct, then so are the inferences about persistences.

3.2 Non-Monotonic Domain Details

Unfortunately, solving the frame problem is not as easy as the previous section suggested. The *qualification problem* [McCarthy 1977] tells us that, in general, it will not be possible to evaluate an exhaustive list of sufficient conditions for property change; surely this lesson also applies to sufficient conditions for lack of change. It will not always (or even rarely) be possible to prove that all potential causes cannot occur, without making some defeasible assumptions. This observation does not contradict the position of this paper stated thus far, so long as the non-monotonic rules are *domain-dependent* and *context-sensitive*.

An example will help clarify this point. Surprisingly, the most convincing example we could find of a reasoning situation that does not require non-monotonicity was the *de facto* benchmark of non-monotonic approaches, the shooting scenario. Therefore, we turn to another common example, where the agent parks its car, goes to work, and then expects to find its car in the same spot at the end of the day [Allen 1984]. Thus the agent has persisted the position of its car, despite that it could have been stolen, borrowed, impounded, or perhaps the parking brake failed and it rolled downhill. None of these actions can typically be proven to not occur, yet the inference that the car will remain is typically very reasonable.

Rather than basing this inference on the general claim that "particular events typically don't happen", we suggest basing the inference on reasonable assumptions about the specific events that could move the car. Such assumptions are inherently domain-dependent. For example, it may be reasonable to assume that the car will not be impounded when parked in the proper lot, as in the following default rule:

$$\forall\tau[\text{true(in-lot},\tau) \Rightarrow \neg\text{occurs(Impounded},\tau)]\ ,$$

where "$\Rightarrow$" means "defeasibly implies". The semantics of such an operator have been carefully investigated elsewhere [Loui 1987, Nute 1986]; for simplicity we will adopt an informal interpretation where a consistent consequence is considered true when the antecedent is considered true. Thus this rule allows the reasoner to exclude the possibility that the car is towed when it is consistent to do so.

Defeasible assumptions must also be sensitive to the acquisition of new relevant evidence. Of course, the nature of defeasible rules provides a weak sensitivity to context, namely that when new information indicates that the consequence is no longer *consistent*, it is no longer asserted. However, this mechanism does not handle new information that suggests the consequent is no longer *reasonable*; for example, if the car is parked in a fire lane or a handicapped space, it may be more appropriate to invoke the following rule:

$$\forall\tau[(\text{true(in-lot},\tau) \wedge \text{true(illegally-parked},\tau)) \Rightarrow \text{occurs(Impounded},\tau)]\ ,$$

which stipulates the opposite conclusion. Along the same lines, if hazard flashers allow for temporary parking in an illegal space, the reasoner may also have the following rule:

$$\forall\tau[(\text{true(in-lot},\tau) \wedge \text{true(illegally-parked},\tau) \wedge \text{true(flashing},\tau)) \Rightarrow \neg\text{occurs(Impounded},\tau)]\ ,$$

and so on. These rules show how different contextual knowledge can motivate quite different assumptions. This context-sensitivity is only possible in a domain-dependent approach to specifying causal knowledge.

The coexistence of the above rules will lead to a conflict between their invocations, requiring that the reasoner has some mechanism to resolve the conflict. This is simple for the above rules, because there is a clear intuition that the last rule should have priority due to its more *specific* antecedent [Etherington 1987]. When the relationship between rules is not so clear, their interactions can be represented explicitly using defeaters [Loui 1987, Nute 1986]; defeaters also have the advantage that they support a weaker interaction, where an assumption is blocked without having to assume its negation, e.g. the reasoner makes no assumption about the car being towed.

There is an argument for domain-independent causal assumptions that should be mentioned here. If a causal scenario is interpreted as a *story*[4], then certain communication conventions become involved. In particular, the storyteller is expected to supply all details necessary for the reader to correctly make the relevant inferences [Grice 1975]. Therefore, if a relevant question is unanswered, the reader may assume that the storyteller believes that the answer is the most typical one given the story. For example, if the standard shooting scenario is a story, then the reader would answer the question "was the gun unloaded during the wait?" with "of course not, or I would have been told so". These assumptions can be drawn for any story, so they are in that sense domain-independent. However, note that the inference involves the reader ascertaining what value is typically based on the supplied details, which is a domain-dependent inference like discussed above. Regardless, the encoding of communication conventions should not be considered fundamental to causal reasoning, since there are applications where these inferences are not generally valid (e.g. in planning, the agent may not have sensed some relevant property, and he will not have any guarantees that the property's value is typical).

This domain-dependent non-monotonic approach follows the course that Georgeff [1986] suggests when he says:

> ... it seems reasonable, at first, to assume that my car is still where I left it this morning, unless I have information that is inconsistent with that assumption. However, this premise gets less and less reasonable as hours turn into days, weeks, months, years, and centuries – even if it is quite *consistent* to make such a premise. This puts the [frame] problem where it should be – namely, in the area of making reasonable assumptions, not in the area of *defining* the effects of actions, the persistency of facts, or causal laws.

This passage also (perhaps unintentionally) exposes a limitation of the

[4]It appears that architects of non-monotonic approaches to causal reasoning often base their intuitions on this interpretation, although this is rarely made explicit (an exception being by McCarthy [1986]).

approach thus far, since it has no way of capturing the notion "less and less reasonable". Indeed, it is even unclear what makes an assumption "reasonable" at all; must it be true in some large percentage of possible worlds, or merely more likely than its negation? Must the risks and benefits of adopting the assumption be considered? The next section provides an approach from which to address these questions, by generalizing causal knowledge from non-monotonic to *statistical*.

3.3 Statistical Domain Details

The similarity between non-monotonic and statistical inference has often been noted [Loui 1987, Pearl 1988], since they both concern the assignment of belief to sentences based on global properties of a body of knowledge. This section shows how to employ basic techniques from statistical inference [Kyburg 1974] to assign numerical beliefs to assertions about the persistence of properties. The basic statistical persistence inference is the derivation of sentences of the form:

$$B(true(p,t') \mid true(p,t)) = x ,$$

for *particular* p, t and x. This equation should be read as "x is the belief that p is true immediately after t, given that p is true at t." This sentence is also often called an assignment of *conditional probability*. Under the statistical interpretation of probabilistic belief, this sentence can be derived from a *conditional statistic* which generalizes over objects in the domain of discourse, e.g. the following:[5]

$$\%\pi[true(\pi,t') \mid true(\pi,t)] = x ,$$

which should be read as "x (by convention, from the real interval $[0,1]$) is the proportion of properties true at t which are also true at time t'." Note that the symbol % plays a role similar to a quantifier in ordinary logic. Thus, given this statistic, and the epistemic conviction that property p *is a random element* [Kyburg 1987] of properties true at time t, then we may use the number x as the belief "$B(true(p,t') \mid true(p,t))$". Intuitively, if we know that "very little changed today", then some random property is likely to not have changed.

Many causal reasoning inferences are based on a generalization over temporal arguments. For example, "the car almost always stays in the parking

[5]The notation (and semantics) for statistics is from Bacchus [1988], except that the variables appear as a prefix instead of a subscripted suffix. A simpler set-notation for causal statistics appears elsewhere [Tenenberg and Weber 1989, Weber 1989a, Weber 1989b], but the current notation is more appropriate for this paper because of its similarity to the standard representation of properties and occurrences.

lot" is analogous to the statistical assertion:

$$\%\tau[\text{true(in-lot},\tau') \mid \text{true(in-lot},\tau)] \approx 1 .$$

The generalization over times seems to be very basic to human appreciation of causality, as evidenced by common phrases like "almost always" and "most of the time". The popular intuition "properties tend to not change" can be captured by generalizing over both properties and times, i.e.

$$\%\pi,\tau[\text{true}(\pi,\tau') \mid \text{true}(\pi,\tau)] \approx 1 .$$

However, this statistic is only applicable *when the properties and times of interest are random.* This is the essence of the domain-independent/dependent issue in this statistical setting: a reasoner will often have contextual information that invalidates the use of the above generalization, making it necessary for the reasoner to have and use more specific statistics, just as a default reasoner is obliged to use more specific non-monotonic rules when more contextual information is considered.

Contextual knowledge is handled by asserting statistics with more specific conditions on the right-hand side of the conditional bar. For example, when the car is known to be illegally parked, it will be more appropriate to use the statistic:

$$\%\tau[\text{true(in-lot},\tau') \mid \text{true(in-lot},\tau) \wedge \text{true(illegally-parked},\tau)] = .8 ,$$

which inspires a much different numerical persistence belief. Thus the interactions between the application of different statistics exhibits the "non-monotonic" behavior necessary for causal reasoning.

Actions may be causally related to property change in a manner similar to cause-closure axioms. The value of "$B(\text{true}(p,t') \mid c(p,t))$" for some contextual knowledge $c(p,t)$ can be related to a list of potential causes $a_1,\ldots,a_n$ by repeated applications of the additivity theorem:

$$B(P \mid Q) = B(P \wedge R \mid Q) + B(P \wedge \neg R \mid Q) ,$$

which produces the following series:

$$B(\text{true}(p,t') \wedge \text{occurs}(a_1,t) \mid c(p,t)) + B(\text{true}(p,t') \wedge \neg\text{occurs}(a_1,t) \wedge \text{occurs}(a_2,t) \mid c(p,t)) + \ldots + B(\text{true}(p,t') \wedge \neg\text{occurs}(a_1,t) \wedge \ldots \wedge \neg\text{occurs}(a_n,t) \mid c(p,t)) .$$

To simplify this series, assume that the actions $a_1,\ldots,a_n$ cannot occur pairwise simultaneously (by definition), and that the last term is insignificant (which assumes that the reasoner knows all significant potential causes). This leaves the following series:

$$\sum_i B(\text{true}(p,t') \wedge \text{occurs}(a_i,t) \mid c(p,t)) .$$

This can be simplified further if the actions unconditionally imply the effect (by definition[6]), making the first conjunct redundant, thus producing the following equality:

$$B(\text{true}(p,t') \mid c(p,t)) = \sum_i B(\text{occurs}(a_i,t) \mid c(p,t)) .$$

This equality justifies the following algorithm: to derive a probabilistic belief in a property being true in some context, sum up the beliefs in each possible cause, relative to the current context. Note that this applies to either the belief in the persistence of p, when c(p,t) implies true(p,t), or the change of p, when c(p,t) implies ¬true(p,t). Also, remember that the values of these beliefs follow from appropriate statistics as described above.

The derivation of probabilistic beliefs improves on the non-monotonic approach in several ways. The notion of becoming "less reasonable" has a straightforward interpretation as declining numeric belief. Dean and Kanazawa [1988] model the decay of belief in the persistence of a property by using *survivor functions*, which express belief in a property as an (exponential) function over elapsed time. Also, the notion of what assumptions are reasonable can be defined by an acceptance threshold, or better yet, numeric beliefs provide a basic for constructing a *decision theory* [Chernoff and Moses 1959], in order to appreciate the influences of utility and decision. Also, causal statistics provide a more powerful approach to the qualification problem [Tenenberg and Weber 1989], and allow for efficient numerical approximations based on the examination of statistical impacts [Weber 1989b].

A full exposition of statistical causal reasoning is beyond the scope of this paper. We refer the interested reader to the work of Weber [1989a, 1989b], Tenenberg and Weber [1989], Dean and Kanazawa [1988], Hanks [1988], and Pearl [1988] for more details and areas of current research.

4. Conclusions

The search for a domain-independent solution to the frame problem has not been successful. The optimistic insistence on generating reasonable inferences about persistence without the requisite causal and contextual

[6]This assumption about the definition of actions does not preclude contingent effects, since these actions can be contingently generated from other actions. For example, the action "shoot" will always imply that the gun is no longer loaded, yet "shoot" is generated from "pull-trigger" only when the gun was loaded.

knowledge has produced a series of solutions that contain an unreasonable assumption of occurrence omniscience. It is time for the field to concentrate on the competence and performance of representations that capture the domain-dependent knowledge that a reasoner can use to derive the plausibility of properties persisting. The most promising such approach uses statistical assertions to capture causal relationships, employing well-established techniques from statistical inference to derive probabilistic beliefs.

Acknowledgements

This work was supported partially by the Air Force Systems Command, Rome Air Development Center, and the Air Force Office of Scientific Research under Contract No. F30602-85-C-0008, which supports the Northeast Artificial Intelligence Consortium (NAIC). This work was also supported partially by U.S. Army contract no. DAAB10-87-K-022, and by the German Research Institute for Artificial Intelligence (DFKI).

During the many revisions of this paper, I benefitted greatly from discussions with Josh Tenenberg, Henry Kyburg, Len Schubert, Brian Haugh, and James Allen. Special thanks to Rainer Bleisinger for his careful reading of this latest version.

References

[Allen 1984] James F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, **23**(2):123–154.

[Allen and Koomen 1983] James F. Allen and Johannes Koomen. Planning with a temporal world model. In the *Proceedings of IJCAI-83*.

[Bacchus 1988] Fahiem Bacchus. *Representing and Reasoning with Probabilistic Knowledge.* PhD thesis, University of Alberta.

[Chernoff and Moses 1959] H. Chernoff and L. Moses. *Elementary Decision Theory.* Wiley.

[Dean and Kanazawa 1988] Thomas Dean and Keiji Kanazawa. Probabilistic temporal reasoning. In the *Proceedings of AAAI-88*, pages 524-528.

[Etherington 1987] David W. Etherington. More on inheritance hierarchies with exceptions: Default theories and inferential distance. *Proceedings of AAAI-87*, pages 352–357.

[Fikes and Nilsson 1971] Richard E. Fikes and Nils J. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, **2**:198–208.

[Georgeff 1986] Michael P. Georgeff. Actions, processes, and causality. In the *Proceedings of the 1986 Workshop on Reasoning about Actions and Plans.*

[Ginsberg 1987] Matthew Ginsberg. Possible worlds planning. In the *Proceedings of the 1986 Workshop on Reasoning about Actions and Plans.*

[Goldman 1970] Alvin Goldman. *A Theory of Human Action.* Princeton University Press, Princeton, NJ.

[Goodwin 1987] Scott D. Goodwin. Reasoning in temporal domains: Dealing with independence and unexpected results. In the *Proceedings of CS/CSI*, pages 46–52.

[Grice 1975] H. Paul Grice. Logic and conversation. In Peter Cole and Jerry L. Morgan, editors, *Syntax and Semantics, Volume 3: Speech Acts*, pages 41–58. Academic Press.

[Haas 1987] Andrew Haas. The case for domain-specific frame axioms. In Frank M. Brown, editor, *The Frame Problem in Artificial Intelligence: Proceedings of the 1987 Workshop.* Morgan Kaufman.

[Hanks and McDermott 1986] Steve Hanks and Drew McDermott. Default reasoning, non-monotonic logics, and the frame problem. In the *Proceedings of AAAI-86.*

[Hanks 1988] Steve Hanks. Probabilistic projection. In the *Proceedings of AAAI-88.*

[Haugh 1987] Brian Haugh. Simple causal minimizations for temporal persistence and projection. In *Proceedings of IJCAI-87*, pages 218–223.

[Kautz 1986] Henry A. Kautz. The logic of persistence. In *Proceedings of AAAI-86.*

[Kyburg 1974] Henry E. Kyburg, Jr. *The Logical Foundations of Statistical Inference.* Reidel.

[Kyburg 1987] Henry E. Kyburg, Jr. Full beliefs. *Theory and Decision.*

[Lifschitz 1986] Vladimir Lifschitz. Pointwise circumscription: Preliminary report. In *Proceedings of AAAI-86.*

[Lifschitz 1987] Vladimir Lifschitz. In Frank M. Brown, editor, *The Frame Problem in Artificial Intelligence: Proceedings of the 1987 Workshop.* Morgan Kaufman.

[Loui 1987] Ronald P. Loui. *Theory and Computation of Uncertain Inference and Decision.* Ph.D. thesis, University of Rochester Computer Science Department.

[McCarthy 1977] John McCarthy. Epistemological problems of artificial intelligence. In the *Proceedings of IJCAI-77*, pages 1034–1044.

[McCarthy 1980] John McCarthy. Circumscription - a form of non-monotonic reasoning. *Artificial Intelligence* **13**(1,2):27-39.

[McCarthy 1984] John McCarthy. Applications of circumscription to formalizing common sense knowledge. In the *Proceedings of the AAAI Non-Monotonic Workshop*, pages 295–324.

[McCarthy and Hayes 1969] John McCarthy and Patrick J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In *Machine Intelligence 4*, pages 463-502.

[McDermott 1982] Drew McDermott. A temporal logic for reasoning about processes and plans. *Cognitive Science 6:101-155.*

[Morgenstern and Stein 1988] Leora Morgenstern and Lynn Andrea Stein. Why things go wrong: A formal theory of causal reasoning. In the *Proceedings of AAAI-88*, pages 518–523.

[Nute 1986] Donald Nute. LDR: A logic for defeasible reasoning. ACMC Research report 01–0013, Advanced Computational Methods Center, University of Georgia.

[Pearl 1988] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufman.

[Sandewall 1988] Erik Sandewall. Non-monotonic entailment for reasoning about time and action - part I: Sequential actions. Working paper.

[Schubert 1989] Lenhart Schubert. Solving the original frame problem without frame axioms or non-monotonicity. In Henry Kyburg and Ron Loui, editors, *Selected Papers from the 1988 Society for Exact Philosophy Conference.*

[Shoham 1986] Yoav Shoham. Chronological Ignorance. In the *Proceedings of AAAI-86.*

[Tenenberg and Weber 1989] Josh D. Tenenberg and Jay C. Weber. A statistical approach to the qualification problem (and how it also addresses the frame problem). Technical report, University of Rochester Computer Science Department.

[Weber 1988] Jay C. Weber. A versatile approach to action reasoning. Technical report 237, University of Rochester Computer Science Department.

[Weber 1989a] Jay C. Weber. *Principles and Algorithms for Causal Reasoning with Uncertainty.* Ph.D. thesis, University of Rochester Computer Science Department.

[Weber 1989b] Jay C. Weber. A parallel algorithm for statistical belief refinement and its use in causal reasoning. *Proceedings of IJCAI-89.*

[Weber 1989c] Jay C. Weber. The comparative expressiveness of situation calculus and temporal logic, in theory and practice. working paper.

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

DFKI
-Bibliothek-
Stuhlsatzenhausweg 3
6600 Saarbrücken 11
FRG

DFKI Publikationen

Die folgenden DFKI Veröffentlichungen oder die aktuelle Liste von erhältlichen Publikationen können bezogen werden von der oben angegebenen Adresse.

DFKI Publications

The following DFKI publications or the list of currently available publications can be ordered from the above address.

DFKI Research Reports

RR-90-01
Franz Baader
Terminological Cycles in KL-ONE-based Knowledge Representation Languages
33 pages

RR-90-02
Hans-Jürgen Bürckert
A Resolution Principle for Clauses with Constraints
25 pages

RR-90-03
Andreas Dengel & Nelson M. Mattos
Integration of Document Representation, Processing and Management
18 pages

RR-90-04
Bernhard Hollunder & Werner Nutt
Subsumption Algorithms for Concept Languages
34 pages

RR-90-05
Franz Baader
A Formal Definition for the Expressive Power of Knowledge Representation Languages
22 pages

RR-90-06
Bernhard Hollunder
Hybrid Inferences in KL-ONE-based Knowledge Representation Systems
21 pages

RR-90-07
Elisabeth André, Thomas Rist
Wissensbasierte Informationspräsentation: Zwei Beiträge zum Fachgespräch Graphik und KI:

1. Ein planbasierter Ansatz zur Synthese illustrierter Dokumente
2. Wissensbasierte Perspektivenwahl für die automatische Erzeugung von 3D-Objektdarstellungen

24 pages

RR-90-08
Andreas Dengel
A Step Towards Understanding Paper Documents
25 pages

DFKI Technical Memos

TM-89-01
Susan Holbach-Weber
Connectionist Models and Figurative Speech
27 pages

TM-90-01
Som Bandyopadhyay
Towards an Understanding of Coherence in Multimodal Discourse
18 pages

TM-90-02
Jay C. Weber
The Myth of Domain-Independent Persistence
18 pages

DFKI Documents

D-89-01
Michael H. Malburg & Rainer Bleisinger
HYPERBIS: ein betriebliches Hypermedia-Informationssystem
43 Seiten

D-90-01
DFKI Wissenschaftlich-Technischer Jahresbericht 1989
45 pages

The Myth of Domain-Independent Persistence

Jay C. Weber

TM-90-02

Technical Memo