THE REPRESENTATION OF EVENTS IN MULTIAGENT DOMAINS

Michael P. Georgeff *

Artificial Intelligence Center
SRI International
Menlo Park, California

Abstract

The purpose of this paper is to construct a model of actions and events suited to reasoning about domains involving multiple agents or dynamic environments. A model is constructed that provides for simultaneous action, and the kind of facts necessary for reasoning about such actions are described. A model-based *law of persistence* is introduced to describe how actions affect the world. No frame axioms or syntactic frame rules are involved in the specification of any given action, thus allowing a proper model-theoretic semantics for the representation. Some serious deficiencies with existing approaches to reasoning about multiple agents are also identified. Finally, it is shown how the law of persistence, together with a notion of causality, makes it possible to retain a simple model of action while avoiding most of the difficulties associated with the frame problem.

1 Introduction

A notion of events and processes is essential for reasoning about problem domains involving one or more agents situated in dynamic environments. While previous papers [3,4,5,6] discussed the importance of the notion of *process*, herein we focus on the representation of events and actions. As we will show, the approach avoids many of the difficulties associated with other models of events and actions.

2 Events

We assume that, at any given instant, the world is in a particular *world state*. Each world state consists of a number of *objects* from a given domain, together with various *relations* and *functions* over those objects. A sequence of world states will be called a *world history*.

A given world state has no duration; the only way the passage of time can be observed is through some change of state. The world changes state by the occurrence of *events*. An event (strictly, an *event type*) is a set of state sequences, representing all possible occurrences of the event *in all possible situations* (see also [1,12]).

In this paper, we will restrict our attention to *atomic events*. Atomic events are those in which the state sequences are of length two, and can be modeled as a transition relation on world states. This transition relation must include all possible state transitions, *including those in which other events occur simultaneously with the given event*. Consequently, the transition relation of an atomic event places restrictions on those world relations that are directly affected by the event, but leaves most others to vary freely (depending upon what else is happening in the world). This is in contrast to the classical approach, which views an event as changing some world relations but leaving most others unaltered.

For example, consider a domain consisting of blocks A and B at possible locations 0 and 1. Assume a world relation that represents the location of each of the blocks, denoted *loc*. Consider two events, $move(A,1)$, which has the effect of moving block A to location 1, and $move(B,1)$, which has a similar effect on block B. Then the classical approach (e.g., see reference [13]) would model these events as follows:

$$\begin{aligned} move(A,1) = \ & \{\langle loc(A,0), loc(B,1)\rangle \rightarrow \langle loc(A,1), loc(B,1)\rangle \\ & \langle loc(A,0), loc(B,0)\rangle \rightarrow \langle loc(A,1), loc(B,0)\rangle\} \end{aligned}$$

and similarly for $move(B,1)$.

Every instance (transition) of $move(A,1)$ leaves the location of B unchanged, and similarly every instance of $move(B,1)$ leaves the location of A unchanged. Consequently, it is impossible to compose these two events to form one that represents the simultaneous performance of both $move(A,1)$ and $move(B,1)$, except by using some interleaving approximation.

In contrast, our model of these events is:

$$\begin{aligned} move(A,1) = \ & \{\langle loc(A,0), loc(B,1)\rangle \rightarrow \langle loc(A,1), loc(B,1)\rangle \\ & \langle loc(A,0), loc(B,1)\rangle \rightarrow \langle loc(A,1), loc(B,0)\rangle \\ & \langle loc(A,0), loc(B,0)\rangle \rightarrow \langle loc(A,1), loc(B,1)\rangle \\ & \langle loc(A,0), loc(B,0)\rangle \rightarrow \langle loc(A,1), loc(B,0)\rangle\} \end{aligned}$$

and similarly for $move(B,1)$.

This model represents all possible occurrences of the event, including its simultaneous execution with other events. For example, if $move(A,1)$ and $move(B,1)$ are performed simultaneously, the resulting event will be the intersection of their possible behaviors:

$$\begin{aligned} & move(A,1) \| move(B,1) = \\ & move(A,1) \cap move(B,1) = \\ & \{\langle loc(A,0), loc(B,0)\rangle \rightarrow \langle loc(A,1), loc(B,1)\rangle\} \end{aligned}$$

Thus, to say that an event has taken place is simply to put constraints on some world relations, and leave most others to vary freely.

* Also affiliated with the Center for the Study of Language and Information, Stanford University, Stanford, California.

This research has been made possible in part by a gift from the System Development Foundation and by the Office of Naval Research under Contract N00014-85-C-0251.

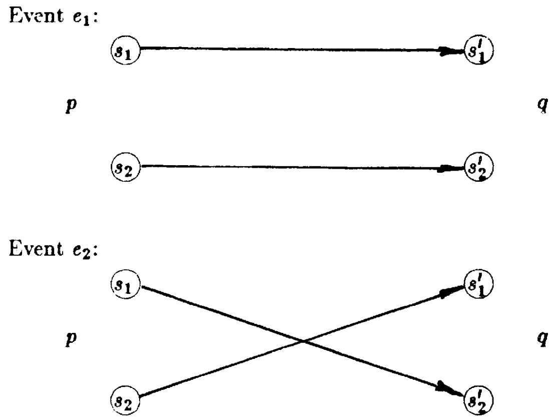

Figure 1: Two Incompatible Events

Of course, to specify events by explicitly listing all the possible transitions would, in general, be infeasible. We therefore need some formalism for describing events and world histories; herein, we will use something similar to situation calculus [11].

Essentially, there are two things we need to say about the possible occurrences of any given event. The first is needed to specify the effects of the event occurring in some given situation. The second is needed to specify under what conditions we consider the event to have occurred, and is essential if we are to reason about the possibility of events occurring simultaneously.

Let ϕ and ψ be conditions on world states (usually called *fluents* [11]), let $occurs(e)(s)$ represent the fact that the event e occurs in state s and, for a given world history w containing state s, let $succ(s)$ be the successor of s. Then we can describe the effects of an occurrence of e with axioms of the following form:[1]

$$\forall w, s \ .\ \phi(s) \wedge occurs(e)(s) \supset \psi(succ(s))$$

This statement is intended to mean that, in all possible world histories, if ϕ is true when the event e occurs, ψ will be true in the resulting state. It has essentially the same meaning as $\phi \supset [e]\psi$ in dynamic logic. Axioms such as these are essential for planning, allowing the determination of the strongest [provable] postconditions and weakest [provable] preconditions of events [15].

At first glance, it appears as if this is *all* we really need for planning and other forms of practical reasoning. For example, assume we have the following axioms describing events e_1 and e_2:

$$\forall w, s \ .\ \phi_1(s) \wedge occurs(e_1)(s) \supset \psi_1(succ(s))$$

$$\forall w, s \ .\ \phi_2(s) \wedge occurs(e_2)(s) \supset \psi_2(succ(s))$$

From this we can infer that

$$\forall w, s \ .\ (\phi_1 \wedge \phi_2)(s) \wedge ((occurs(e_1) \wedge occurs(e_2))(s) \supset (\psi_1 \wedge \psi_2)(succ(s))$$

However, it would be unwise to take this as the basis for a plan to achieve $(\psi_1 \wedge \psi_2)$. The reason is that it may be impossible for the two events to occur simultaneously, even if $(\psi_1 \wedge \psi_2)(succ(s))$ is not provably false.

For example, consider the two events shown in Figure 1. Let's assume that p holds in states s_1 and s_2 and q holds in the successor states. (We have taken some liberties in naming states, but that is not important for this example.) Events e_1 and e_2 satisfy the above axioms, where $\phi_1 = \phi_2 = p$ and $\psi_1 = \psi_2 = q$.

Given these axioms alone, it is quite consistent to assume that both events occur simultaneously, but there is no way to *prove* that they can so occur – in fact, given e_1 and e_2 as shown in Figure 1, such a statement is clearly false. (Given *sufficient* axioms about the effects of these events, we could, of course, prove that such events *could not* occur together.)

To describe what conditions constitute the occurrence of an event e, we need axioms of the form

$$\forall w, s \ .\ \phi(s) \wedge \psi(succ(s)) \supset occurs(e)(s)$$

This statement is intended to capture the fact that, for all world histories, we consider the event e to have occurred if ϕ holds at the beginning of the event and ψ holds afterwards. Facts such as these are critical for reasoning about whether two or more events can proceed simultaneously and *cannot* be inferred from statements of the former kind about the effects of events.

For example, consider that two events e_1 and e_2 both satisfy[2]

$$\forall w, s \ .\ p(s) \wedge q(succ(s)) \supset occurs(e_j)(s)$$

To prove that these events can occur simultaneously, all we need do is prove that, in some world history, p holds of one state and q holds of its successor.

Often, we may even be able to make stronger statements than these. For example, the event $move(A, 1)$ satisfies

$$\forall w, s \ .\ occurs(move(A, 1))(s) \equiv loc(A, 0)(s) \wedge loc(A, 1)(succ(s))$$

This specification completely characterizes the event $move(A, 1)$ – there is nothing more that can be said about the event. Thus, *at this point of the story,* the frame problem does not arise. Because the event, in and of itself, places no restrictions on the majority of world relations, we do not require (indeed, it would be false to require) a large number of frame axioms stating what relations the performance of the event leaves unchanged. In contrast to the classical approach, we therefore need not introduce any *frame rule* [7] or STRIPS-like assumption [2] regarding the *specification* of events.

3 Actions

When a process brings about an event we will say that the process performs an *action*. For now, we can consider an action and the event it brings about to be the same object – that is, a relation on world states. Later on, we shall have to distinguish the two.

If we are to form plans in multiagent worlds, one of the more important considerations is whether or not any two or more actions can be performed concurrently – it is of little use to form a plan that calls for the simultaneous performance of actions that simply cannot coexist. Thus, to guarantee the validity of a plan containing simultaneous actions, we need to prove that it is indeed possible to perform the actions simultaneously.

Consider two actions a_1 and a_2 that bring about events e_1 and e_2, respectively. In constructing a plan that involves the simultaneous performance of a_1 and a_2, it is not enough that it simply be *consistent* that e_1 and e_2 occur together. The example discussed in the preceding section is a case in point. Of course, this may be the best one can do given incomplete knowledge of the world but, in such cases, there is certainly no guarantee that the plan would ever succeed.

To guarantee the success of such a plan, we need to be able to *prove* that a_1 and a_2 can be performed simultaneously. To

[1] We will assume throughout that, in such axioms, s is an element of w.

[2] The events shown in Figure 1 do not satisfy this axiom!

do this, we need to prove that the intersection of the transition relations corresponding to e_1 and e_2 is nonempty and that its domain includes the states in which the actions a_1 and a_2 might be performed. For example, consider that we have

$$\forall w, s \,.\, \phi_1(s) \wedge \psi_1(succ(s)) \supset occurs(e_1)(s)$$

$$\forall w, s \,.\, \phi_2(s) \wedge \psi_2(succ(s)) \supset occurs(e_2)(s)$$

It is easy to see that, if we are in a state in which ϕ_1 and ϕ_2 hold, both events can occur together if there exists a world history containing a successor state in which ψ_1 and ψ_2 hold. Unfortunately, ascertaining this involves determining the *consistency* of $(\psi_1 \wedge \psi_2)$, which is undecidable in the nonpropositional case. Moreover, determining performability of actions on the basis of consistency arguments can lead to nonmonotonicity – addition of further axioms could invalidate any conclusions drawn.

In fact, a similar problem arises even for single-agent planning – it is not possible to infer from axioms describing the effects of actions that these effects are indeed satisfiable. To get around this problem, it is usual to assume that no action ever fails, i.e., that there is always a transition from any state satisfying the preconditions of the action to some subsequent state (e.g., [15]).

This option is not open to us in the multiagent domain – simultaneous actions are often not performable. What we need is some way to determine whether or not composite actions will fail on the basis of some property of the component actions. To do this, we introduce a notion of *action independence*.

The approach we adopt is to provide additional axioms specifying which relational tuples the action *directly affects*.[3] To do this, for every action a and n-ary predicate symbol P, we introduce a formula $\delta_P(a, \tilde{x})$, called a *direct-effects formula* ($\tilde{x}$ represents an n-tuple of free variables).

The meaning of this formula is that, for all $\tilde{x}$, if $\delta_P(a, \tilde{x})$ holds in some state s, *only* those relational tuples denoted by $P(\tilde{x})$ may be affected by the performance of action a; any relational tuple that is not a direct effect of action a is thus free to vary *independently* of the occurrence of a. Thus, $P(\tilde{x})$ may be *forced* to take on some particular truth value in any state resulting from the performance of a; conversely, all other atoms involving P are free to take on any truth value.

For example, $\delta_{loc}(move(A, 1), x, y) \equiv (x = A)$. This means that the action $move(A, 1)$ could affect any tuple denoted by $loc(A, y)$, for any y; on the other hand, it would not affect any other tuples in the relation denoted by loc. There are two important points to note here: (1) this does *not* mean that the other tuples of loc remain *unchanged* – some other action could occur simultaneously that affected these tuples also; and (2) if we wish to infer that $loc(B, y)$ does not change for some y, we need to know that A and B denote different objects.

Given such formulae, it follows that two actions a_1 and a_2 can occur simultaneously in a state s if s is in the domain of each action and, for each n-ary predicate P, $(\neg\exists\tilde{x}.(\delta_P(a_1, \tilde{x}) \wedge \delta_P(a_2, \tilde{x})))$ holds in s – that is, both actions don't directly affect the same relational tuple. For example, assuming unique names, we can infer that the location of B is unaffected by $move(A, 1)$ and that $move(A, 1)$ could be performed simultaneously with any action a' that changed the location of B, provided that, conversely, a' did not affect the location of A.

In the case that the same relational tuples *are* affected, it might be that each relational tuple is changed by each action in the same way, and simultaneity would still be possible. But we then get forced back to considering consistency of formulae. There is no difficulty with this if consistency can be determined and does not involve any nonmonotonicity (such as when one condition (say, ψ_1) implies the other (ψ_2), and we know that ψ_1 is satisfiable). However, if this is not the case, any conclusions drawn must be subject to retraction and thus should be treated as *assumptions* about the problem domain.

Note that all the direct effects of an action need not be involved in any single occurrence of that action – they represent only *possible* effects. Also, the direct effects of an action do not define the possible state transitions – this is given, as before, by the state transition relation associated with the action.

There are some problems with this representation, not the least being that, in many cases of interest, we still have to check consistency of formulae. However, knowledge about the relational tuples that actions may affect, and reasoning about interactions on the basis of this knowledge, seems to be an important part of commonsense reasoning. As we will shortly see, such knowledge also plays an important role in determining the effects of actions performed in isolation.

[3] In the general case, we would also have to specify which functional values and constants were directly affected by the action. This is a straightforward extension of the described approach, and we will not consider it further (see reference [13]).

4 The Law of Persistence

We have been viewing atomic actions or events as imposing certain constraints on the way the world changes while leaving other aspects of the situation free to vary as the environment chooses. That is, each action transition relation describes all the potential changes of world state that could take place during the performance of the action. Which transition actually occurs in a given situation depends, in part, on the actions and events that take place in the environment. However, if we cannot reason about what happens when some subset of all possible actions occurs – in particular, when only one action occurs – we could predict very little about the future and any useful planning would be impossible.

What we need is some notion of *persistence* that specifies that, in general, world relations only change when forced to [12]. For example, because the action $move(A, 1)$ defined in the previous section places no constraints on the location of B, we would not expect the location of B to change when $move(A, 1)$ was performed in isolation from other environmental actions.

One possibility is to introduce the following *law of persistence*:

$$\forall w, s, \tilde{x} \,.\, \phi_P(\tilde{x})(s) \wedge (\neg\exists a \,.\, (occurs(a) \wedge \delta_P(a, \tilde{x})))(s)$$
$$\supset \phi_P(\tilde{x})(succ(s))$$

where $\phi_P(\tilde{x})$ is either $P(\tilde{x})$ or $\neg P(\tilde{x})$.

This rule states that, provided no action occurs that directly affects the relational tuple denoted by $P(\tilde{x})$, the truth value of $P(\tilde{x})$ is preserved from one state to the next. It can be viewed as a generalization of the rule used by Pednault for describing the effects of actions in single-agent worlds [13]. For example, we could use this rule to infer that, if $move(A, 1)$ were the only action to occur in some state s, the location of B would be the same in the resulting state as it was in state s.

However, at this point we encounter a serious deficiency in the action model we have been using and, incidentally, in all others that represent actions and events as the set of all their possible behaviors (e.g., [1,12]). Consider, for example, a seesaw, with ends A and B and fulcrum F. We shall assume there are no other entities in the world, that the only possible locations for

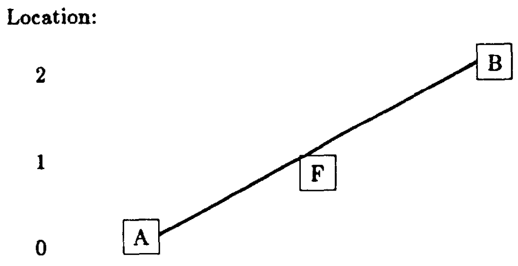

Figure 2: Possible Seesaw State after $move_F$

A, F, and B are 0, 1, and 2, and that these are always colinear. Assume that initially A, F, and B are at location 0, and consider an action $move_F$ that moves F to location 1 (see Figure 2), while allowing all possible movements of A and B, depending on what other actions are occurring at the same time (such as someone lifting B). Of course, the objects must always remain colinear.

The possible transitions for $move_F$ are to one of the states $\langle loc(A,1), loc(F,1), loc(B,1)\rangle$, $\langle loc(A,0), loc(F,1), loc(B,2)\rangle$, or $\langle loc(A,2), loc(F,1), loc(B,0)\rangle$. Furthermore, because the movement of F places constraints on both the locations of A and B, the direct affects of the action will include the locations of all objects:

$\delta_{loc}(move_F, x, y) \equiv (x = A) \vee (x = F) \vee (x = B)$.

Thus, the effect of $move_F$, in addition to changing the location of F, will be to change either the location of A or the location of B or both. The question is, if no other action occurs simultaneously with $move_F$, *which* of the possible transitions can occur?

Let's assume that, because of the squareness of the fulcrum F, the action $move_F$ always moves A and B to location 1 at the same time, unless some parallel action forces either A or B to behave differently. Unfortunately, using our current action model there is simply no way to represent this. We cannot restrict the transition relation so that it always yields the state in which A, F, and B are all at location 1, because that would prevent A or B from being moved simultaneously with A. Furthermore, the constraint on locations is a contingent fact about the world, not an analytic one – thus, we cannot sensibly escape the dilema by considering any of the relations *derived* from the others (as many philosphers have pointed out).

From a purely behavioral point of view this is how things should be. To an external observer, it would appear that $move(A,1)$ sometimes changed the location of A and not B (when some simultaneous action occurred that raised A to location 2), sometimes changed the location of B and not A (when some simultaneous action raised B), and sometimes affected the locations of both A and B. (Of course, the action would always change the location of F). As there is no *observation* that could allow the observer to detect whether or not another action was occurring simultaneously, there is no way the action $move_F$ could be distinguished from any other that had the same transition relation. For example, there would be no way to distinguish $move_F$ from an action $move'_F$ that exhibited the same set of possible behaviors but, when performed in isolation, left A where it was and moved B to location 2.

On the other hand, when reasoning about processes, we *do* want to be able to make this distinction. For example, there may be two different ways of moving F, one corresponding to $move_F$ and the other to $move'_F$. In other cases, while an action like $move_F$ might be appropriate to seesaws, an action analgous to $move'_F$ might be needed for describing object movements in other situations. For example, consider the situation where, instead of being parts of a seesaw, A is a source of light and B is F's shadow.

We therefore make a distinction between actions and events – one that is critical for reasoning about processes and plans. That is, an event is simply identified with all its possible occurrences; in particular, two atomic events having the same transition relation are considered identical. However, actions with the same transition relation (such as $move_F$ and $move'_F$) are not necessarily identical – they may behave differently when performed in isolation and may play different causal roles in a theory of the world.

Clearly, therefore, we cannot determine which action we intend from knowledge (even complete knowledge) of all the possible state transitions (event occurrences) which constitute performance of the action. In particular, we *cannot use any general default rule or minimality criteria* to determine the intended effects of an action when performed in isolation. Indeed, in the case of $move_F$, note that we do *not* minimize the changes to world relations or maximize their persistence: *both* A and B change location along with F.

It appears, then, that the only thing we can do is to specify what happens when the action occurs in isolation *in addition to* specifying what happens when other actions occur in parallel. This is certainly possible, but the representation would be cumbersome and unnatural.

5 Causality

One way to solve this problem is by introducing a notion of causality. As used herein, if an action a_1 is stated to *cause* an action a_2, we require that a_1 always occur simultaneously with a_2. Thus, in this case, a_1 could never be performed in isolation – a_2 would always occur simultaneously with every occurrence of a_1.

For example, we might have a causal law to express the fact that whenever a block x is moved, any block on top of x and not somehow restrained (e.g., by a string tied to a door) will also move. We could write this as

$$\forall w, s, x, y, l \,.\, (occurs(move(x,l)) \wedge on(y,x) \wedge \neg restrained(y))(s) \\ \supset occurs(move(y,l))(s)$$

The notion of causality used by us is actually more general than that described above, and is fully described elsewhere [5]. We use the term in a purely technical sense, and while it has many similarities to commonsense useage, we don't propose it as a fully-fledged theory of causality. Essentially, we view causality as a relation between atomic actions that is conditional on the state of the world. We also relate causation to the temporal ordering of events, and assume that an action cannot cause another action that precedes it. However, we do allow an event to cause another that occurs simultaneously (as in this paper). This differs from most formal models of causality [8,12,16].

But how does this relate to the problem of persistence and the specification of the effects of actions performed in isolation? The answer is that we can thereby provide axioms that explicity describe how an action affects the world in the context of other actions either occurring *or not.*

For example, consider the action $move_F$ described in the previous section. We begin by modifying the definition of this action so that its only direct effect is the location of the fulcrum F itself. This means that the transition relation for $move_F$ will have to include world states in which A, F, and B are not colinear, but this is no problem from a technical point of view. Indeed, at least in this case, there is also an intuitive meaning to such worlds; namely, those in which the seesaw is broken. However, there is no problem with requiring all possible world histories (*not* all world states!) to satisfy the linearity constraint.

We then add causal laws which force the simultaneous movement of either A or B or both. For example, we might have the following causal law:

$$\forall w, s \,.\, occurs(move_F)(s) \land (\neg\exists a' \,.\, (occurs(a') \land interferes(a', move(A,1))))(s) \supset occurs(move(A,1)(s)$$

where $interferes(a_1, a_2)(s)$ means that it is not possible to perform actions a_1 and a_2 simultaneously in state s (see Section 3).

The intended meaning of this causal law is that, if we perform the action $move_F$, $move(A,1)$ is caused to occur simultaneously with $move_F$ unless another action occurs that forces A to occupy a location different from 1. A similar causal law would describe the movement of B. Both laws could be made conditional on the seesaw being intact, if that was desired.

There are a number of things to be observed about this approach. First, it would appear that we should add further causal laws requiring the movement of at least one of A or B in the case that both could not move to location 1. However, this is not necessary. For example, let us assume that, at the moment we perform $move_F$, some other action occurs simultaneously that moves B to location 2 (without directly affecting the location of A). As the direct effects of neither this action nor the action $move_F$ include the location of A, we might expect application of the above causal law to yield a resulting state in which A is at location 1. However, this is clearly inconsistent with the constraint that A, F, and B must remain colinear.

If we examine this more carefully, however, the impossibility of such a world state simply implies that the antecedent of the above causal law must, in this case, be false. That is, there must exist an action that occurs in state s and that cannot be performed simultaneously with $move(A,1)$. Indeed, this is exactly the action that would have appeared in any causal laws that forced the colinearity constraint to be maintained. The point of this example is that in many cases we do not need to include causal laws to maintain invariant world conditions – we can, instead, use the constraints on world state to infer the existence of the appropriate actions.

Second, the application of causal laws need not yield a unique set of caused actions – it could be that one causal law requires the location of A to change and B not, while another requires the location of B to change and A not. Given only this knowledge of the world, the most we could infer would be that one but not both of the actions occurs – but *which* one would be unknown. (Interestingly, this bears a strong similarity to the different possible extensions of a theory under certain kinds of default rules [14].)

Third, actions are clearly distinct from events (cf. [1,12,16]). In particular, actions with the *same* transition relation – i.e., exhibiting the same set of possible behaviors – may play different causal roles. For example, with no outside interference, $move_F$ causes the movement of both A and B, whereas $move'_F$ causes the movement of A alone. This is *not* the same distinction that is made between actions and events in the philosophical literature, but it does have some similarities.

Finally, we may not be able to *prove* that no interference arises, which, in the above example, would prevent us from inferring that the action $move(A,1)$ occurs. However, this is not a serious problem – if we cannot prove that the action either occurs or does not, we simply will not know the resulting location of A (unless, of course, we make some additional *assumptions* about what events are occurring).

Causal laws can be quite complex, and may depend on whether *or not* other actions occur as well as on conditions that hold in the world. It is the introduction of such laws that allows us to represent what happens when only a subset of all possible actions occur. We gain by having simpler descriptions of actions but, in return, require more complex causal laws. On the other hand, it is now easy to introduce other causal laws, such as ones that describe what happens when a block is moved with a cup on top of it, when the cup is stuck with glue, or tied with a string to a door, or when other blocks are in the path of the movement.

Some predicates are better considered as *defined* predicates, which avoids overpopulating the world with causal laws. For example, the distance between two objects may be considered a defined predicate. Instead of introducing various causal laws stating how this relation is altered by various move actions, we can simply work with the basic entities of the problem domain and infer the value of the predicate from its definiens when needed.

6 The Frame Problem

The frame problem, as Hayes [7] describes it, is dealt with in our approach by means of the law of persistence. This has a number of advantages. First, because this law is a property of our action model, and not of our action specification language, we avoid all of the semantic difficulties usually associated with the frame problem.

Second, we avoid the problem of having to state a vast number of uninteresting frame axioms by means of direct-effects formulae, which describe *all* those relational tuples (and, in the general case, functional values and constants) that can possibly change.

Third, we avoid having unduly complex direct-effects formulae and action representations by introducing causal laws that describe how actions bring about (cause) others. Of course, the causal laws can themselves be complex (just as is the physics of the real world), but the representation and specification of actions is thereby kept simple.

There are also important implementation considerations. The approach outlined here is at least tractable, as the relations and functions that can be affected by the occurrence of an action require, at most, provability of the formulae of interest. Interestingly, one of the most efficient action representations so far employed in AI planning systems – the STRIPS representation [2,10] – is essentially the special case in which (1) the transition relation for each action can be represented by a single precondition-postcondition pair; (2) the postcondition is a conjunction of literals; (3) the direct effects (which correspond to the elements in the delete list) include all the literals mentioned in the postcondition; and (4) no actions ever occur simultaneously with any other. The approach used by Pednault [13] can also be considered the special case in which there are no simultaneous actions.

Some researchers take a more general view of the frame problem, seeing it as the problem of reasoning about the effects of actions and events with *incomplete information* about what other actions or processes (usually the environment) may be occurring simultaneously. Unfortunately, this problem is often confused with the *representation* of actions, with the result that there is usually no clear model-theoretic semantics for the representation.

For example, one of the major problems in reasoning about actions and plans is in determining which actions and events can possibly occur at any given moment. Based on the relative infrequency of "relevant" actions or events, or that one would "know about" these if they occurred, it has been common to use various default rules (e.g., [12]) or minimal models (e.g., [9,16]) to constrain the set of possible action occurrences. However, there are many cases where this is unnecessary – where we can *prove*, on the basis of axioms such as those appearing in this paper, that no actions of interest occur. We may even have axioms that allow one to avoid consideration of whole classes of actions, such as when one knows that certain actions are external to a given process. Thus, in many cases, there is simply no need to use default rules or minimality principles – reasoning about plans and actions need not be nonmonotonic.

In the case that we *do* need to make assumptions about action occurrences, the use of default rules and circumscription can be very useful. For example, by minimizing the extension of the *occurs* predicate we can obtain a theory in which the only action occurrences are those that are *causally necessary*. However, there is no need to limit oneself to such default rules or minimality criteria. There may be domain-specific rules defining what assumptions are reasonable, or one may wish to use a more complicated approach based on information theory. We may be able to make reasonable assumptions about freedom from interference; to assume, for example, that a certain relational tuple will not be influenced by actions in other processes.

It is not our intention to consider herein the problem of making useful assumptions about actions and freedom from interference – it is, of course, not a simple problem. However, it is important to keep this problem separate from the issue of action representation. For example, it at first seems reasonable 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 assumption gets less and less reasonable as hours turn into days, weeks, months, years, and centuries. This puts the problem where it should be – in the area of making reasonable assumptions, not in the area of *defining* the effects of actions [2,7], the persistency of facts [12], or causal laws [16].

7 Conclusions

We have constructed a model of atomic actions and events that allows for simultaneity, and described the kind of facts required for reasoning about such actions. We introduced a *law of persistence* that allows the effects of actions to be determined and, most importantly, have shown how the representation of actions and their effects involves no frame axioms or syntactic frame rules. We also pointed out some deficiencies in existing approaches to reasoning about multiagent domains: for example, that consistency of predications over states or intervals cannot be taken as proof that actions can proceed concurrently, and that models that represent actions simply as the set of all their possible behaviors cannot make certain distinctions critical for planning in multiagent domains. Finally, we showed how the law of persistence, together with the notion of causation, makes it possible to retain a simple model of action while avoiding most of the difficulties associated with the frame problem.

Acknowledgments

I wish to thank especially Amy Lansky and Ed Pednault, both of whom helped greatly in clarifying many of the ideas presented in this paper.

References

[1] Allen, J. F., "A General Model of Action and Time," Computer Science Report TR 97, University of Rochester, Rochester, New York (1981).

[2] Fikes, R. E., and Nilsson, N. J., "STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving," *Artificial Intelligence*, 2, pp. 189–208 (1971).

[3] Georgeff, M. P., "A Theory of Action for Multiagent Planning," *Proc. AAAI-84*, Austin, Texas (1984).

[4] Georgeff, M. P. "A Theory of Process" Workshop on Distributed AI, Sea Ranch, California (1985).

[5] Georgeff, M. P. "Process, Action, and Causality," Workshop on Planning and Reasoning about Action, Timberline Lodge, Mount Hood, Oregon (1986).

[6] Georgeff, M. P., and Lansky, A. L., "Procedural Knowledge," *Proc. IEEE*, Special Issue on Knowledge Representation (1986).

[7] Hayes, P. J., "The Frame Problem and Related Problems in Artificial Intelligence," in *Artificial and Human Thinking*, A. Elithorn and D. Jones (eds.), Jossey-Bass (1973).

[8] Lansky, A. L., "Behavioral Specification and Planning for Multiagent Domains," Tech. Note 360, Artificial Intelligence Center, SRI International, Menlo Park, California (1985).

[9] Lifschitz, V. "Circumscription in the Blocks World," Computer Science Working Memo, Stanford University, Stanford, California (1985).

[10] Lifschitz, V. "On the Semantics of STRIPS," Workshop on Planning and Reasoning about Action, Timberline Lodge, Mount Hood, Oregon (1986).

[11] McCarthy, J., and Hayes, P. J., "Some Philosophical Problems from the Standpoint of Artificial Intelligence, in *Machine Intelligence 4*, pp. 463–502 (1969).

[12] McDermott, D., "A Temporal Logic for Reasoning about Processes and Plans," *Cognitive Science*, 6, pp. 101–155 (1982).

[13] Pednault, E. P. D., "Toward a Mathematical Theory of Plan Synthesis," Ph.D. thesis, Department of Electrical Engineering, Stanford University, Stanford, California (1986).

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

[15] Rosenschein, S. J., "Plan Synthesis: A Logical Perspective," *Proc. IJCAI-81*, Vancouver, British Columbia (1981).

[16] Shoham, Y. "Chronological Ignorance: Time, Knowledge, Nonmonotonicity, and Causation," Workshop on Planning and Reasoning about Action, Timberline Lodge, Mount Hood, Oregon (1986).