Reasoning about Change over Time: Actions, Events, and their Effects

B Knight, T Peng and J Ma

School of Computing and Mathematical Sciences,
University of Greenwich, London SE18 6PF, U.K.

Abstract

This paper presents a new formalism for reasoning about change over time. The formalism derives a clean separation between the notion of states and situations. It allows more flexible temporal causal relationships than do other formalisms for reasoning about causal change, such as the situation calculus and the event calculus. It includes effects that start during, immediately after, or some time after their causes, and which end before, simultaneously with, or after their causes. A formal distinction between actions, action-types and events is proposed, which allows the expression of common-sense causal laws at high level. It is shown how these laws can be used to deduce state change over time at low level, when events occur under certain preconditions hold. Two problems that beset most interval-based temporal systems, i.e., the so-called *dividing instant problem* and *intermingling problem*, are absent from the formalism.

1. Introduction

Theories of action and change usually involve reasoning about time and causation with regard to the truth of propositions and the occurrences of events. Over the past half century, many approaches have been proposed for reasoning about the dynamic aspects of the world. Amongst these, the two most influential formalisms are probably McCarthy and Hayes' framework of the *situation calculus* [McC63, McH69] and Kowalski and Sergot's *event calculus* [KoS86]. Noticing that temporal reasoning plays an important role in reasoning about actions/events and change, a series of revised formalisms has been introduced to characterise richer temporal features in either the situation calculus or the event calculus (e.g., [Sch90, GLR91, LiS91, PiR93,95, MiS94 and Sha95]). In addition, other researchers have proposed alternative approaches for representing time, action and causation, and change and persistence. Examples of these approaches are those of McDermott [McD82], of Allen and Ferguson [All84, AlF94], of Shoham [Sho88], and of Sandewall [San89], etc. However, there are still some problematic questions that have been left unanswered in the existing approaches.

Generally speaking, in most of these approaches, the world is assumed to persist in a given state until some action is carried out over some specified time to change it into another state. In most versions of the situation calculus, the word "situation" is used just as a kind of context or time label referring to the set of facts that describe the corresponding state of the closed world of discourse; and, sometimes, it is even used interchangeably with the word "state" (e.g., [GeN87]). However, our claim in this paper is that there are indeed some fundamental differences between the concept of situations and the concept of states of the world if temporal issues are taken into account. Hence, in order to fully capture the temporal aspects of the dynamic world, we derive the following separation between the notion of states and situations:

- The state of the world, which may be described by a certain set of facts, is time independent - the state at a given time does not necessarily have to be different from the state at another time.

- Situations are time dependent - a situation is an association of a given state of the world with a particular time over which the world holds in that state. Hence, each situation must be taken as unique. In other words, two situations that refer to the state of the world associated with two different times, respectively, must be taken as different, no matter whether the states of the world at these two times are the same or not. Additionally, since each situation is associated with a particular time, there should be, on intuitive grounds, a constraint that excludes the case that an action may change a situation to an earlier situation.

Another issue concerns how to characterise the notion of actions and events. In the literature, the two words, "action" and "event", are used in many different senses by various researchers; sometimes they are also used interchangeably. Similarly to the separation between states and situations, we also claim that actions and events should be treated as different entities:

- Actions are time independent - they are simply some names denoting certain useful and relevant activities that may be conducted over some time by the agents to accomplish changes of the state of the world. E.g., "Shooting", "Striking", "Running", "Sleeping", and so on. Following Gooday and Galton's idea [GoG96], in this paper, we shall call the association of a given action with a certain temporal duration an action type, e.g., "Striking for 24 hours", "Running for three minutes", and so on. It is important to note that the same action can give various action types. For instance, "Running for three minutes" and "Running for two hours" denote two different action types. Action types are high level entities - they are still time independent: for any given action type, it may perform once, more than once over different times, or may not even perform at all.

- Events are time dependent - each individual performance of an action type over a specified time constitutes an event at low level. E.g., the performance of action type, "Striking for 24 hours", over the whole day of the 16th of September 1994 constitutes the event, e = <Striking, 16/09/1994>.

In most existing formalisms for representing causal relationships between events and their effects, the effect of an event is represented by the result immediately after the occurrence of the event. Expression of flexible temporal relationships between events and their effects, which are in fact quite interesting and complicated, has been mostly neglected. It is true that, in many cases, the effect of an action takes place immediately after the performance of the action; e.g., in the blocks world, as soon as action "moving a block from the top of another block onto the table" completes, the block being moved should be on the table (immediately). However, sometimes there may be a time delay between an action and its effect(s). E.g., *30 seconds after you press the button at the crosswalk, the pedestrian light turns to green* [GLR91]. In [GLR91], Gelfond *et al.* propose an approach to deal with delayed effects, using the idea of appending an action that involves no activity, i.e., *Wait*, to actions with delayed effects. However, as pointed out by Allen and Ferguson [AlF94], this is neither convenient nor intuitive. Also we argue that their method is inadequate to general cases (see section 3 for detail).

The objective of this paper is to propose a new formalism that allows expression of high level common-sense knowledge about action and change, and also supports explicit representation of time and occurrence of events at low level. It preserves general assertions such as "*the beginning of effects cannot precede the beginning of their causes*". The structure of this paper is as follows. Section 2.1 introduces a discrete time structure that, as shown in section 2.2, is useful in dealing with the *Intermingling Problem* and *Dividing Instant Problem* (DIP). The concepts of states and situations are formally characterised in section 2.3. Section 2.4 proposes a clean distinction between actions, action types and events, which enables us to represent common-sense causality at both high-level and low level in section 2.5. Section 3 provides an illustrative example. A brief discussion about the corresponding frame problem is given in section 4. Finally, section 5 summarises and concludes the paper.

2. The formalism

The formalism will be described in terms of a many-sorted reified logic with equality, including variables for **prime times** (p, p_1, p_2, ...), **propositional fluents** (f, f_1, f_2, ...), **states** (s, s_1, s_2, ...), and **actions** (a, a_1, a_2, ...). Also we adopt the conventional theories of reals and integers.

2.1 The Time Structure

The time structure described below is in fact an extension of the time structure previously proposed by Knight and Ma in [KnM92], which can be seen as a specified model of the general time theory introduced by Ma and Knight [MaK94].

We assume there is a function, *Dur*, assigns to each prime time a non-negative real number, called the duration of that prime time. We shall call p a moment if $Dur(p) > 0$, otherwise, p is called a point. Additionally, we assume that the collection of prime times is *similar to* the set of integers. That is, there exists a *similar function* from the set of integers to prime times, i.e., a one-to-one mapping that preserves the order relation [Lip64]. Hence, we have:

- The collection of prime times form a discrete set, denoted by **P**, which is well-ordered by an immediate predecessor relation. We shall use $Meets(p_1, p_2)$ to denote that prime time p_1 is the immediate predecessor of prime time p_2.
- **P** has no *limit elements* [Lip64];
- The fundamental time structure is *discrete*, *linear* and *unbounded*.

N.B. The fundamental temporal structure we assume is a simple linear model of time. Branching time models have been proposed as a useful means of handling possible worlds, uncertainty about the past or the future and the effects of alternative actions when planning. It is argued that it is not necessary to have a theory that assumes that time itself branches; notions of possibility would be handled by introducing a separate model operator to represent possibility explicitly [AlF94, KnM94].

It is important to note that prime times have no internal structure. In other words, they are all non-decomposable, even some of them, i.e., those moments, may have a positive duration, e.g., $Dur(p) = 2.5$. We impose the following axiom:

(P1) $Meets(p_1, p_2) \Rightarrow Dur(p_1) > 0 \vee Dur(p_2) > 0$

which preserves that two points cannot meet each other. This is different from that proposed by Allen and Hayes' [AlH89], which alternatively forbids a moment to meet another moment.

Based on the fundamental time structure, we extend the concept of prime times to general times, denoted by t or T (possibly indexed). The collection of (general) times, **T**, is the smallest set, such that:

- Each prime time is a time (i.e., $\mathbf{P} \subset \mathbf{T}$);
- For any two times, t_1 and t_2, such that $Meets(t_1, t_2)$, the *adjacent union* of t_1 and t_2, denoted as $t_1 \oplus t_2$, is a time:

(T1) $Meets(t_1, t_2) \Rightarrow$
$\forall t',t''(Meets(t', t_1) \wedge Meets(t_2, t'') \Rightarrow Meets(t', t_1 \oplus t_2) \wedge Meets(t_1 \oplus t_2, t'')$

(T2) $Meets(t_1,t_2) \Rightarrow Dur(t_1 \oplus t_2) = Dur(t_1) + Dur(t_2)$

(T3) $Meets(t_1', t_1) \wedge Meets(t_1', t_2) \wedge Meets(t_1, t_2') \wedge Meets(t_2, t_2') \Leftrightarrow t_1 = t_2$

N.B. (T1) and (T2) preserve the extension of *Meets* relation and *Dur* function to general times, where (T3) ensures that the time that connects two times separated by a sequence of times is unique. In what follows, $t_1 \oplus t_2$ always implies that $Meets(t_1, t_2)$.

In fact, **T** is the closure of **P,** under operator $\oplus$. Elements in **T** are not necessary non-decomposable. We shall call decomposable times *intervals*. From the definition, for any interval t, there exist time t_1 and time t_2, such that $t = t_1 \oplus t_2$; and by (P1) and (T2), $Dur(t) > 0$.

For the convenience of expression, we define the following temporal relations between any two times, t_1 and t_2:

$Equal(t_1, t_2) \Leftrightarrow t_1 = t_2$
$Before(t_1, t_2) \Leftrightarrow \exists t(Meets(t_1, t) \wedge Meets(t, t_2))$
$Overlaps(t_1, t_2) \Leftrightarrow \exists t, t',t'' (t_1 = t' \oplus t \wedge t_2 = t \oplus t'')$
$Starts(t_1, t_2) \Leftrightarrow \exists t(t_2 = t_1 \oplus t)$
$During(t_1, t_2) \Leftrightarrow \exists t',t''(t_2 = t' \oplus t_1 \oplus t'')$
$Finishes(t_1, t_2) \Leftrightarrow \exists t(t_2 = t \oplus t_1)$
$Met\text{-}by(t_1, t_2) \Leftrightarrow Meets(t_2, t_1)$
$After(t_1, t_2) \Leftrightarrow Before(t_2, t_1)$
$Overlapped(t_1, t_2) \Leftrightarrow Overlaps(t_2, t_1)$
$Started\text{-}by(t_1, t_2) \Leftrightarrow Starts(t_2, t_1)$
$Contains(t_1, t_2) \Leftrightarrow During(t_2, t_1)$
$Finished\text{-}by(t_1, t_2) \Leftrightarrow Finishes(t_2, t_1)$
$In(t_1, t_2) \Leftrightarrow Starts(t_1, t_2) \vee During(t_1, t_2) \vee Finishes(t_1, t_2)$
$Sub(t_1, t_2) \Leftrightarrow Equal(t_1, t_2) \vee Starts(t_1, t_2) \vee During(t_1, t_2) \vee Finishes(t_1, t_2)$

In the above, while $In(t_1, t_2)$ denotes that time t_1 is a proper part of time t_2, $Sub(t_1, t_2)$ states that time t_1 is a proper part of time t_2 or is t_2 itself.

2.2 Fluents

A fluent is a proposition whose truth value is dependent on times. It can take different values over different times. We introduce the predicate, *Holds*, so that for each pair of a fluent f and a prime time p, the formula $Holds(f, p)$ represents that fluent f holds true with respect to prime time p.

Corresponding to the extension from **P** to its closure **T** under operator $\oplus$, we also extend predicate *Holds*, and hence the formula $Holds(f, t)$, for each pair of fluent f and time t.

It is interesting to note that, in most interval-based temporal theories, there exist two major problematic questions, i.e., the so-called *dividing instant problem* and the *intermingling problem*. The dividing instant problem involves the question of whether time intervals should include their ending-points or not, which can lead to some difficulties in determining the truth values of fluents at the ending-points [Van83, Vil94]. The intermingling problem arises when a fluent may change its truth value infinitely often in an interval with a finite duration [Ham71, Gal96]. This will lead to some difficulties in characterising the relationships between the negation of fluents and the negation of involved sentences.

However, in the time structure described above, each time is simply defined as an ordered union of prime points/moments that are non-decomposable, there is no definition of ending points for intervals. Therefore, the so-called *dividing instant problem* doesn't exist. In addition, since the time structure excludes any limit element, the *intermingling problem* is overcome.

In our formalism, we take the follwing interpretation of negation as the constraint imposed on the *Holds* predicate with respect to decomposable intervals: $\neg Holds(f, t)$ is true if and only if it is not the case that f holds true throughout t. In othre words, $\neg Holds(f, t)$ is true if and only if f changes its truth-value over time t. Such a weak interpretation seems appropriate for the standard definition of implication and preserves a simple two-valued logic [AlF94]. The following axiom shows the relation between the truth of a proposition over an interval and its truth over parts of that interval:

(2.2.1) $t = t_1 \oplus t_2 \Rightarrow Holds(f, t) \Leftrightarrow \forall t'(In(t', t) \Rightarrow Holds(f, t'))$

It is important to note that, (2.2.1) associates sub-interval/moments/points property just to decomposable intervals, rather than non-decomposable moments or points. This is different from Allen's approach [All84], which may lead to some dubious inferences (e.g. any property holds true with respect to any moment) [MaK94].

By (2.2.1), it is straightforward to infer that, for any fluent f and any time t_1, t_2:

(Th2.2.1) $Holds(f, t_1) \wedge Holds(f, t_2) \wedge Meets(t_1, t_2) \Rightarrow Holds(f, t_1 \oplus t_2)$.

Following Galton's suggestion [Gal90], we shall use not(f) to represent the negation of fluent f, to be kept distinct from ordinary sentence-negation, symbolised by "$\neg$".

Intuitively, with respect to any time t, any fluent and its negation can not both hold true. In other words, they are in conflict with each other over any time. However, it is important to note that, for a given fluent, say f, it may be the case that its negation not(f) is not the only fluent that conflicts with it. That is, there may be some fluents other than not(f) that can not be true together with fluent f.

In this paper, we shall use Conf(f_1, f_2) to denote that fluent f_1 is in conflict with fluent f_2:

(2.2.2) $Conf(f_1, f_2) \Rightarrow \forall t(\neg Holds(f_1, t) \vee \neg Holds(f_2, t))$

2.3 States and Situations

The state of the world is described by a collection of fluents. We use *Belongs*(f, s) to denote that fluent f belongs to the collection of descriptors for state s. Following Shanahan's approach [Sha95], the predicate *Belongs* is defined as follows:

(2.3.1) $s_1 = s_2 \Leftrightarrow \forall f(Belongs(f, s_1) \Leftrightarrow Belongs(f, s_2))$

(2.3.2) $\forall s_1, f_1 \exists s_2(\forall f_2(Belongs(f_2, s_2) \Leftrightarrow Belongs(f_2, s_1) \vee f_1 = f_2))$

(2.3.3) $\exists s \forall f(\neg\ Belongs(f, s))$

A state is called *consistent* if and only if

$$Belongs(f_1, s) \wedge Belongs(f_2, s) \Rightarrow \neg Conf(f_1, f_2)$$

otherwise, s is called an *inconsistent* state. In what follows, all the states we will consider are assumed as consistent.

We also use *Holds*(s, t) to denote that s is the state (of the world) with respect to time t:

(2.3.4) $Holds(s, t) \Leftrightarrow \forall f(Belongs(f, s) \Rightarrow Holds(f, t)) \wedge (\neg Belongs(f, s) \Rightarrow Holds(\text{not}(f), t)))$

By (2.3.1) and (2.3.4), it is straightforward to infer that:

(Th2.3.1) $Holds(s_1, t) \wedge Holds(s_2, t) \Rightarrow s_1 = s_2$

That is, with respect to a given time, the state of the world is unique. However, there is nothing to stop a temporally contiguous time having the same state. In order to capture the relationships between states and times, we introduce the concept of situations. A situation is defined as the state of the world associated with a particular time over which the world holds in that state. We shall denote situations as sit, sit_1, sit_2, ... All the situations form a collection of pairs of states and times, such that:

(2.3.5) $\forall sit \exists s \exists t(sit = \langle s, t \rangle \wedge Holds(s, t))$
(2.3.6) $sit = \langle s_1, t_1 \rangle \wedge sit = \langle s_2, t_2 \rangle \Rightarrow s_1 = s_2 \wedge t_1 = t_2$

Hence, the representation of any situation is unique. In what follows, if $sit = \langle s, t \rangle$ is a situation, we shall call s and t its *reference state* and *reference time*, denoted as *State*(sit) and *Time*(sit), respectively. In fact, *State* and *Time* can be seen as two functions from the set of situations to the set of states, and the set of times, respectively. By (2.3.1) - (2.3.6), it is easy to see that:

(Th2.3.2) $Time(sit_1) = Time(sit_2) \Rightarrow State(sit_1) = State(sit_2)$

Without confusion, we shall use $Holds(f, sit)$ to denote that fluent f holds true in situation sit, providing that:

(2.3.4) $Holds(f, sit) \Leftrightarrow Holds(f, Time(sit))$

By axioms (2.2.1), and (2.3.1) - (2.3.4), one can infer that:

(Th2.3.3) $Holds(\text{not}(f), sit) \Leftrightarrow \neg Holds(f, sit)$

2.4 Actions and Events

As discussed in the introduction, in this paper, we treat actions and events as different entities. *Actions* are simply some names denoting certain useful and relevant activities. An *action type* is defined as a pair of an action and a non-negative real number denoting a temporal duration, i.e., $<a, d>$, representing the scheme of performing action a for duration d. Both actions and action types are time independent high level entities. An action type may perform once, more than once over different times, or may not even perform at all. Each individual performance of an action type over the time constitutes an event at low level. We shall use notation $<a, t>$ to represent the event that action a performs over time t, and denote events as e, e_1, e_2, ...

We introduce a binary predicate $Performs(a, t)$ to represent that action a performs over time t. The following axiom preserves that if an action performs over two adjacent times respectively, then it performs over the ordered union of the two times:

(2.4.1) $Performs(a, t_1) \wedge Performs(a, t_2) \wedge Meets(t_1, t_2) \Rightarrow Performs(a, t_1 \oplus t_2)$

and the uniqueness of the representation of an event is guaranteed by:

(2.4.2) $\forall e \exists a \exists t(e = <a, t> \wedge Performs(a, t))$
(2.4.3) $e = <a_1, t_1> \wedge e = <a_2, t_2> \Rightarrow a_1 = a_2 \wedge t_1 = t_2$

Analogous to the definitions of the reference state and the reference time of a given situation, for event $e = <a, t>$, we shall call a and t the *reference action* and the *reference time* of e, denoted as $a = Action(e)$ and $t = Time(e)$, respectively.

2.5 Causality

Whereas the intuition behind the notion of situations is persistence, the intuition behind the notion of events is change. The world is supposed to persist in a given state over some certain time until some event occurs to change it into another state. In what follows, we shall show how to use the approach proposed in this paper to represent common-sense causality at both high level and low level.

As argued by Gooday and Galton [GoG96], many action and change problems encountered in AI may be formulated without any low level temporal reference. In particular, many common-sense causal relationships are actually expressed at high level, making no reference to any

specific times. Although these are relationships involving time, they are time independent in that they hold for all time. For example, we might know that if a kettle is filled with cold water and switched on, it will start to boil after five minutes. This causal knowledge applies whenever the kettle is switched on, and specifies that the action type, i.e. <*Heating*, 5>, causes the world to change from state {*Water_Is_Cold*} to state {*Water_Is_Boiling*}. Of course, a common-sense statement such as this does not really mean that the water will always be boiling after being switched on, regardless of other possible actions. Rather it means that it will be boiling in the absence of other events occurring during the 5 minutes, such as countermanding switching or electrical power failures.

To capture this common sense meaning, we define a predicate *Causes,* as follows:

(2.5.1) $Causes(s_1, \langle a, d_a \rangle, d, s_2) \Leftrightarrow$
$\forall t_1,t_a,t(\quad Holds(s_1, t_1) \wedge Dur(t_a) = d_a \wedge Performs(a, t_a) \wedge Dur(t) = d$
$\wedge Meets(t_1, t_a) \wedge Meets(t_1, t) \wedge \forall t'\forall a'(Sub(t',t) \wedge a' \neq a \Rightarrow \neg Performs(a',t')))$
$\Rightarrow \exists t_2(Meets(t, t_2) \wedge Holds(s_2, t_2)))$

Predicate *Causes* relates an initial state s_1, an action type $\langle a, d_a \rangle$, a temporal duration d and the result state s_2. It states that starting from state s_1, after a time duration d, the action type $\langle a, d_a \rangle$ will change the world into state s_2 as long as no other event has occurred. Predicate *Causes* is independent of time, so that it expresses high level knowledge in the similar sense of Gooday and Galton [GoG96].

By (2.2.1), (2.5.1) and (Th2.3.1), we can infer the following theorem which states that the result state is unique:

(Th2.5.1) $Causes(s_1, \langle a, d_a \rangle, d, s_2) \wedge Causes(s_1, \langle a, d_a \rangle, d, s_2') \Rightarrow s_2 = s_2'$

Formula $Causes(s_1, \langle a, d_a \rangle, d\ s_2)$ characterised above denotes a high level causal law that demonstrates the temporal duration and order in which, the performance of an action causes the world to change from one state to another state. It makes no reference to any absolute times at low level. For convenience of expression at low level, we introduce a predicate $E_Causes(sit_1, e, t, sit_2)$ which expresses an instance of the causal law. This predicate has the meaning that the world changes from situation sit_1 into situation sit_2 after time t, and the change is caused by the occurrence of event e, where no other disturbing events occur.

(2.5.2) $E_Causes(sit_1, e, t, sit_2)$
$\Leftrightarrow \quad Meets(Time(sit_1), Time(e)) \wedge Meets(Time(sit_1), t) \wedge Meets(t, Time(sit_2)))$
$\wedge \forall t'\forall a'(Sub(t', t) \wedge a' \neq Action(e) \Rightarrow \neg Performs(a', t'))$

N.B. $E_Causes(sit_1, e, t, sit_2)$ requires that no other events except e occur over any part of time t. However, in many applications, such a piece of knowledge may not be explicitly available, and deduction has to be done by means of reasoning by default. Therefore, in this case, one has to handle the frame problem. (see chapter 4 for the discussion)

Axioms (2.5.1) and (2.5.2) together characterise an important causal relation at distinct levels. In fact, in the case where the corresponding knowledge is available, we can use the following procedure to deduce low level changes from high level causal laws:

(Th2.5.2) Given causal axiom:

(CA) $auses(s_1, <a, d_a>, d, s_2) \Rightarrow \psi$

where Φ and ψ are conjunctions of predicates *Belongs* and/or $\neg$*Belongs*

and

Temporal constraints:

(TC) $Meets(Time(sit_1), Time(e)) \wedge Meets(Time(sit_1), t)$

Observations:

(Ob1) $State(sit_1) = s_1$
(Ob2) $Action(e) = a$
(Ob3) $Dur(Time(e)) = d_a$

Disturbing action constraint:

(DC) $Sub(t', t) \wedge a' \neq a \Rightarrow \neg Performs(a', t')$

We can deduce that there exists a situation sit_2, such that:

$E_Causes(sit_1, e, t, sit_2)$.

In most existing versions of the situation calculus or event calculus, e.g., [McH69], [Sch90], [LiS91], [PiR93,95], [MiS94] and [Sha95], the effect of an event is represented by the effect immediately after the occurrence of the event (i.e., Case (B) in figure 2). However, axiom (2.5.2) presented above allows other cases besides that of an effect immediately following the cause. In fact, the temporal constraint imposed in (2.5.2) is consistent with the so-called (most) general temporal constraint (*GTC*) (see [McD82], [All84], [Sho88] and [TeT95]), which guarantees the common-sense assertion that "*the beginning of the effect cannot precede the beginning of the cause*". There are in total 5 possible qualitative temporal relationships between $Time(sit_1)$, $Time(e)$, t and $Time(sit_2)$ that satisfy (2.5.2). These are illustrated in Figure 2, including:

Case (A) where the end of event e coincides with the end of the effect sit_2. In the extreme circumstance, when time t is a point with duration of zero, the duration of event e equals to the duration of effect sit_2; that is, the effect becomes true immediately after the beginning of its causal event, and only holds true while the event is in progress;

Case (B) where the effect becomes true immediately after the end of the event and remains true for some time (i.e., $Time(sit_2)$) after the event;

Case (C) where the effect becomes true during the progress of the event and remains true for some time after the event.

Case (D) where the effect only hold over some time during the progress of the event (Similar to case (A), cases (C) and (D) include the extreme circumstance where the effect becomes true immediately after the beginning of the event, i.e., time t is a point with zero duration);

Case (E) where there is a time delay, i.e., t_d, between the event and its effect.

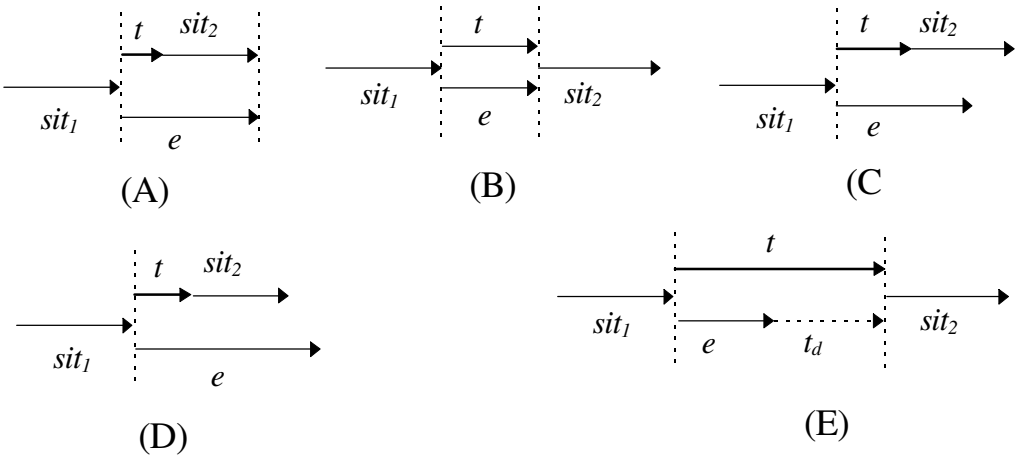

Figure 2

By (2.5.2) and (Th2.3.2), it is easy to infer that:

(Th2.5.3) $E_Causes(sit_1, e, t, sit_2) \wedge E_Causes(sit_1, e, t, sit_2') \Rightarrow State(sit_2) = State(sit_2')$

3. Example

As discussed in the above, temporal relation between events and their effects is complex. The formalism proposed in this paper combines the advantages of temporal logic by the way of allowing explicit time representation and the benefit of state-based approaches such as the situation calculus which can easily represent the worlds that remain static except when the agent acts, and where nothing important happens while actions are being executed [AlF94]. In this section we will use an example to show the application of the new formalism.

Consider the following example:

> *25 seconds after a pedestrian starts pressing the button at the crosswalk, the pedestrian crossing light turns to yellow from red, and after another 5 seconds it turns to green.*

As discussed in introduction, it is difficult to use Gelfond *et al*'s approach to express this example. In fact, since there are no situations defined during the time over which an action is performing, one cannot assert anything about the world during an action performance. Intuitively, fluents other than *Time*, such as the locations and colours of objects, do not change their values while you wait [GLR91]. In fact, in this example, action *PressButton* has two delayed effects: the first is the light turns to yellow 25 seconds later, and the second is the light turns to green 30 seconds after pressing the button. Therefore, there is a waiting action, say $Wait_1$ (with duration of 25 seconds), between the action and its first effect, and another waiting action, say $Wait_2$ (with duration of 30 seconds), between the action and its second effect. Here, action $Wait_2$ may be expressed as $Wait_1;Wait_1'$ as in Gelfond *et al*'s approach, where the duration of $Wait_1'$ is 5 seconds. However, in this case, the fluent denoting the colour of the pedestrian light does change its value during the waiting action.

To express this scenario in the formalism proposed here, we use three fluents to describe the state of the pedestrian crossing light:

RedOn: the light at the crosswalk is red;
YellowOn: the light at the crosswalk is yellow;
GreenOn: the light at the crosswalk is green.

Also, we shall use one action type:

<*PressButton*, 0>: pressing the button punctually that takes no duration;

and one event:

$E_{PressButton}$: pressing the button at time point T_{Press}.

We assume that in any situation there is exactly one of the three fluents that holds true:

(3.1) $Holds(RedOn, sit) \wedge \neg Holds(YellowOn, sit) \wedge \neg Holds(GreenOn, sit)$
$\vee \neg Holds(RedOn, sit) \wedge Holds(YellowOn, sit) \wedge \neg Holds(GreenOn, sit)$
$\vee \neg Holds(RedOn, sit) \wedge \neg Holds(YellowOn, sit) \wedge Holds(GreenOn, sit)$

The high level causal axioms in this case are as below:

(3.2) $Belongs(RedOn, s_1) \wedge Causes(s_1, <Pressbutton, 0>, 25, s_2) \Rightarrow Belongs(YellowOn, s_2)$

(3.3) $Belongs(RedOn, s_1) \wedge Causes(s_1, <Pressbutton, 0>, 30, s_2) \Rightarrow Belongs(GreenOn, s_2)$

Let Sit_0 denote a situation in which the red light is on, and the yellow and green lights are off. By axiom (3.1), this knowledge can be expressed as:

(3.4) $Holds(RedOn, Sit_0)$

Assuming in situation Sit_0 event $E_{PressButton}$ occurs, by (3.1)-(3.4) together with the knowledge that for time T_2 (see Fig 3)

$$Sub(t, T_2) \wedge a \neq Action(E_{PressButton}) \Rightarrow \neg Performs(a, t)$$

then by (Th2.5.2), we deduce that there exist situations, say Sit_1 and Sit_2, 25 seconds and 30 seconds after the beginning of event $E_{PressButton}$, respectively, such that:

(3.5) $E_Causes(Sit_0, E_{PressButton}, T_1, Sit_1)$
where $Dur(T_1) = 25 \wedge Holds(YellowOn, Sit_1) \wedge Dur(Time(Sit_1)) \leq 5$

(3.6) $E_Causes(Sit_0, E_{PressButton}, T_2, Sit_2)$
where $Dur(T_2) = 30 \wedge Holds(GreenOn, Sit_2)$

Here we successfully express the fact that there is a delayed time, say T_{D1}, standing between the reference time of event $E_{PressButton}$ and the reference time of the situation Sit_1, that is:

$$Meets(Time(E_{PressButton}), T_{D1}) \wedge Meets(T_{D1}, Time(Sit_1))$$

Similarly, there is a delayed time, say T_{D2}, standing between $Time(E_{PressButton})$ and $Time(Sit_2)$, that is:

$$Meets(Time(E_{PressButton}), T_{D2}) \wedge Meets(T_{D2}, Time(Sit_2))$$

The above knowledge can be graphically presented as Figure 3:

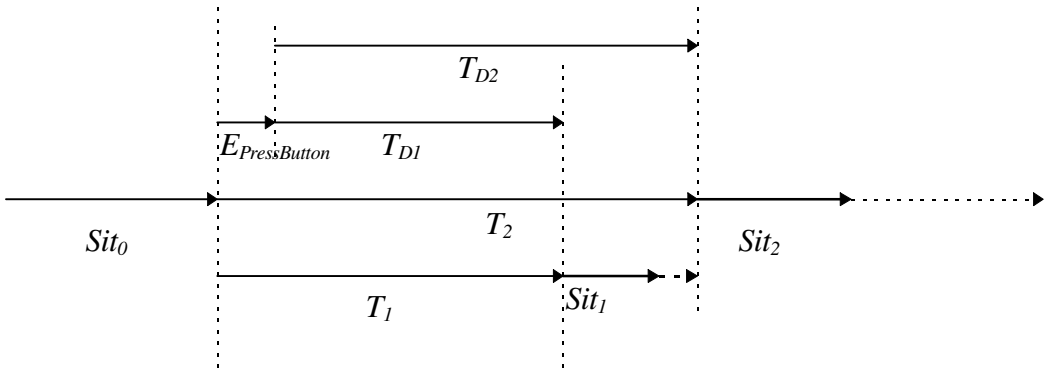

Figure 3

N.B. In this example, we assumed that no other events except $E_{PressButton}$ occur over time T_2. However, as discussed in section (2.5), we have in fact got no knowledge about change of the truth value of fluents over time T_{D1}. What we exactly know is "immediately after time T_1, the pedestrian crossing light turns from red to yellow". Therefore, the frame problem arises: do all the fluents hold their truth value over time T_{D1}?

4. The frame problem

As mentioned above, for the formalism presented in this paper, the frame problem has an extra dimension of time. Not only do we have to specify the values of fluents that are completely unaffected by an action, but also we have to specify the values of fluents on time elements where they are unaffected. That is we have to specify times over which fluents change their truth values as the result of performing an action, as well as time periods over which they persist. To avoid the problem of exhaustive specification, we propose a solution to the corresponding frame problem by extending the standard circumscription technique to include time.

Since predicate *E_Causes* in our formalism has got some extra arguments compared to the *Result* function as used in conventional situation calculus, we need to re-define the predicate *Ab* for the sake of applying the circumscriptive method. In the conventional way, $Ab(f, a, sit)$ simply states that, as the result of performing action a starting from situation sit, fluent f will change its truth value. In our model we shall use $Ab(f, sit_1, e, t, sit_2)$ instead of $Ab(f, a, sit)$. Correspondingly, we have the following modified form for the *common-sense law of inertia*:

(4.1) $\neg Ab(f, sit_1, e, t, sit_2) \Rightarrow Holds(f, sit_1) \Leftrightarrow Holds(f, sit_2)$

The intuitive meaning of axiom (4.1) is that fluent *f* has the same truth value in both situation sit_1 and situation sit_2 unless something is abnormal. It is important to note that, fluent *f* is not abnormal with respect to sit_1, *e*, *t* and sit_2 just ensures that *f* has the same truth value over $Time(sit_1)$ and $Time(sit_2)$, but does not preserve that the truth value of fluent *f* persists throughout time *t*.

Since formula $E_Causes(sit_1, e, t, sit_2)$ tells us not just what a state will hold in the result situation, but also the difference among the result situations as the variation of argument *t*, our circumscription policy should be circumscribing *Ab* with both sit_2 and *t* vary. The main idea of this policy is, if time argument *t* and situation sit_2 are allowed vary, the result situations will range over all possible situations as the variation of argument *t* and sit_2. In this way, we can obtain the expected minimised model, and get rid of all unexpected models.

For simple expression, we shall denote the conjunction of all domain constraint axioms by *D*, the conjunction of all observation facts by *O*, the conjunction of all the causal axioms by *CA*, and the *common-sense law of inertia*, i.e. (4.1), by *LI*. In what follows, we use the pedestrian crossing light example to show that the circumscription

$$\text{CIRC}(D \wedge O \wedge CA \wedge LI; Ab; t, sit_2)$$

provides a solution to the corresponding frame problem. Here sit_2 denotes the result situation.

The actual question with this example can be described as: Does fluent *RedOn* persist over the next 25 seconds after performing action *PressBotton* in the original situation Sit_0?

To deal with this problem by using the above circumscription policy, a key component of the State-Based Minimisation technique is to take the existence-of-situations axiom as a domain constraint axiom [Sha97]. With respect to axiom (3.2.1), such a situation existence axiom for the addressed question is:

(4.2) $\exists sit(Holds(RedOn, sit) \wedge \neg Holds(YellowOn, sit) \wedge \neg Holds(GreenOn, sit)) \wedge$
$\exists sit(\neg Holds(RedOn, sit) \wedge Holds(YellowOn, sit) \wedge \neg Holds(GreenOn, sit)) \wedge$
$\exists sit(\neg Holds(RedOn, sit) \wedge \neg Holds(YellowOn, sit) \wedge Holds(GreenOn, sit))$

In addition, we have the following axioms for uniqueness of event names, and the domain closure for fluents:

(4.3) $\forall e(e = E_{PressButton})$
(4.4) $\forall f(f = RedOn \vee f = YellowOn \vee f = GreenOn)$

Let Y_0 be the conjunction of axioms (3.2.1)-(3.2.6) and (4.1)-(4.4), and let Y_1 be the circumscription of *Ab* in Y_0 with *t* and *sit* varied:

$$Y_1 \equiv CIRC(Y_0; Ab; t, sit)$$

We can prove that Y_1 ensures the following conclusions:

(Con1) $E_Causes(Sit_0, E_{PressButton}, T_1, Sit_1)$
$Holds(YellowOn, Sit_1) \wedge Holds(RedOn, T_1)$

(Con2) $E_Causes(Sit_0, E_{PressButton}, T_1, Sit_1) \wedge E_Causes(Sit_0, E_{PressButton}, T_2, Sit_2)$
$Holds(YellowOn, T_3) \wedge Holds(GreenOn, Sit_2)$

where Sit_0, Sit_1, Sit_2, T_1 and T_2 are interpreted as in Figure 3, and T_3 is the time connecting T_1 and $Time(Sit_2)$.

5. Conclusion and Future work

In this paper, we present a new formalism for reasoning about actions, events and their effects. We argued that, for the purpose of representing both the snapshot and the persistence of the state of the world, a clean separation between the notion of states and situations is needed. Also, to express both high level and low level common-sense causality, we propose a clean distinction between actions, action types and events. The formalism allows flexible temporal causal relationships and supports general assertions such as "*the beginning of effects cannot precede the beginning of their causes*". The introduction of the concept of prime times, from which time intervals are constructed, bypasses/overcomes the so-called *dividing instant problem* and the *intermingling problem*, which bother most interval-based time theories.

By extending the circumscription policy to include time, a nonmonotonic solution to the corresponding frame problem is briefly presented. However, due to the limitation of the length of the paper, details are not included. An interesting topic for further research is to extend this framework to include representing and reasoning about concurrent actions/events and their effects. In addition, since the formalism presented here is based on a discrete time model, it would be inadequate for reasoning about continuous change. It seems reasonable to extend the fundamental time structure to a dense one in order to express continuous change. However, in this case, the intermingling problem may arise and hence needs some careful treatments.

References

[All84] Allen J. F. Towards a General Theory of Action and Time, *Artificial Intelligence*, **23**, pp.123-154, 1984.

[AlF94] Allen J. F., Ferguson G. Actions and Events in Interval Temporal Logic, *the Journal of Logic and Computation*, Vol.4, Iss.5, pp.531-579, 1994.

[AlH89] Allen J. F. and Hayes P.J. Moments and Points in an Interval-based Temporal-based Logic, *Computational Intelligence*, **5**, pp.225-238, 1989.

[Gal90] Galton A. A Critical Examination of Allen's Theory of Action and Time, *Artificial Intelligence*, **42**, pp.159-188, 1990.

[Gal96] Galton A. P. An Investigation of Non-intermingling Principles in Temporal Logic, *Journal of Logic and Computation*, Vol.6(2), 267-290.

[GeN87] Genesereth M. and Nilsson N *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, 1987.

[GLR91] Gelfond M., Lifschitz V. and Rabinov A. What are the Limitations of the Situation Calculus? In *Working Notes of AAAI Spring Symposium Series. Symposium: Logical Formalization of Commonsense Reasoning*, pp.59-69, 1991.

[GoG96] Gooday J. and Galton A. The Transition Calculus: a High-level Formalism for Reasoning about Action and Change, *Journal of Experimental and Theoretical Artificial Intelligence*, **9**, pp.67-95, 1996.

[Ham71] Hamblin C. L. Instants and Intervals, *Studium Generale*, **24**, pp127-134, 1971.

[KnM92] Knight B. and Ma J. A General Temporal Model Supporting Duration Reasoning, *AICOM* Vol.5, No.2 pp.75-84, 1992.

[KnM94] Knight B. and Ma J. Time Representation: A Taxonomy of Temporal Models, *Artificial Intelligence Review* Vol.7, pp.401-419, 1994.

[KoS86] Kowalski R. and Sergot M. A Logic-based Calculus of Events, *New Generation Computing,* 4. pp67-95, 1986.

[Lip64] Lipschutz S. Theory and Problems of Set Theory and Related Topics, Schaum Publishing Co., 257 Park Avenue South, New York, 1964.

[LiS91] Lin F. and Shoham Y. Provably Correct Theories of Action, in *Proceedings of AAAI-91*, pp.349-354, 1991.

[MaK94] Ma J. and Knight B. A General Temporal Theory, *The Computer Journal*, Vol.37, No.2, pp.114-123,1994.

[McC63] McCarthy J. Situations, actions and causal laws, *Stanford Artificial Intelligence Project: Memo 2*, 1963.

[McD82] McDermott D.V. A Temporal Logic for Reasoning about Processes and Plans, *Cognitive Science*, Vol.6, pp.101-155, 1982.

[McH69] McCarthy J. and Hayes P. J. Some philosophical problems from the standpoint of artificial intelligence, in *Machine Intelligence*, Vol.4, (Eds.) Meltzer B. and Michie D., Edinburgh University Press, pp.463-502, 1969.

[MiS94] Miller R. S. and Shanahan M. P. Narratives in the Situation Calculus, *the Journal of Logic and Computation*, Vol.4, No.5, pp.513-530, 1994.

[PiR93] Pinto J. and Reiter R. Temporal Reasoning in Logic Programming: A Case for the Situation Calculus, *In Proceedings of 10th Int. Conf. on Logic Programming*, Budapest, Hungary, June 21-24, 1993.

[PiR95] Pinto J. and Reiter R. Reasoning about Time in the Situation Calculus, Annals of Mathematics and Artificial Intelligence, Vol:14, Iss:2-4, pp251-268, 1995.

[San89] Sandewall E. Filter preferential entailment for the logic of action in almost continuous worlds. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-89)*, pp.894-899, 1989.

[Sch90] Schubert L. Monotonic Solution of the Frame Problem in the Situation Calculus: an Efficient Method for Worlds with Fully Specified Actions, in: H.E. Kyburg, R. Loui and G. Carlson, eds., *Knowledge Representation and Defeasible Reasoning* pp.23-67, Kluwer Academic Press, 1990.

[Sha95] Shanahan M. A Circumscriptive Calculus of Events, *Artificial Intelligence*, **77**, pp29-384, 1995.

[Sha97] Shanahan M. *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*, The MIT Press, 1997.

[Sho88] Shoham Y. *Reasoning about Change*: *Time and Causation from the Standpoint of Artificial Intelligence*, MIT Press, 1988.

[TeT95] Terenziani P. and Torasso P. Time, Action-Types, and Causation: an Integrated Analysis, *Computational Intelligence*, Vol.11, No. 3, pp.529-552, 1995.

[Van83] Van Benthem J. *The Logic of Time*, Kluwer Academic, 1983.

[Vil94] Vila L. A Survey on Temporal Reasoning in Artificial Intelligence, *AICOM* Vol7, No.1, pp4-28, 1994.