SICS/R-89/8913

**On the Applicability of Non-Monotonic Logic to
Formal Reasoning in Continuous Time
by
Manny Rayner**

On the Applicability of Non-Monotonic Logic to Formal Reasoning in Continuous Time¹

Manny Rayner

Swedish Institute of Computer Science
Box 1263, S-164 28 KISTA
Sweden

manny@sics.se

Abstract

The paper criticizes arguments recently advanced by Shoham, McDermott and Sandewall, which purport to demonstrate the relevance of non-monotonic logic to the formalization of reasoning about the evolution of mechanical systems in continuous time. The first half of the paper examines the "Extended Prediction Problem" of Shoham and McDermott; reasons are given to support the claim that the "problem" is the product of a mistaken understanding of the formal basis of Newtonian mechanics, and has no real existence. An example is given showing how, contrary to Shoham and McDermott's arguments, it is possible to formalise reasoning about the evolution of physical systems in continuous time using only classical logic and differential calculus. The second half then reviews Sandewall's non-monotonic logic for almost-continuous systems. Here it is argued that the proposed framework offers only very marginal advantages in compactness of notation, and generally tends to collapse back into classical logic. In summary, I conclude that there is as yet no good reason to believe that non-monotonic logic will be a useful tool in this area.

¹Sections 2, 3 and 4 of this paper appeared in substantially the same form under the title "Did Newton Solve the 'Extended Prediction Problem'?", in R.J. Brachmann, H. Levesque and R. Reiter (eds.) *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann, 1989. The material in section 5 has not been previously published.

1 Introduction

AI researchers have been trying to formalize temporal reasoning in a computationally tractable way for a while, but most attempts so far have assumed a discrete model of time. Recently, however, there has been growing interest in continuous temporal reasoning, partly generated by its intrinsic theoretical interest and partly by the desire to incorporate AI methods in robots, autonomous vehicles, and other systems which carry out intelligent real-time interactions with the exterior world. I shall not attempt to define the meaning of the phrase *continuous time* in the most general possible way (those interested may wish to consult [van Bentham 83]); in the remainder of the paper it will be quite sufficient to interpret it as meaning that time will be isomorphic to the real numbers.

If the formalism is going to be logic-based in some way (and many people seem to agree on this), then it is clearly important to decide what kind of logic should be used. At the moment, considerable attention is being paid to suggestions by Shoham, McDermott and Sandewall, to the effect that some kind of non-monotonic logic is either necessary (Shoham and McDermott), or at least offers clear advantages over classical logic (Sandewall). However, a close examination of the arguments used to justify this point of view reveals that they are in fact anything but convincing. The main purpose of the current paper is to argue the contrary position: namely, that non-monotonic logic may well have nothing to offer here, and that attention should be focussed instead on the much better-understood framework of classical logic.

The rest of the paper is organized as follows. I will first examine Shoham's and MacDermott's claims: in section 2, I summarize their arguments concerning the so-called "Extended Prediction Problem" in continuous-time temporal reasoning, which they claim is not amenable to solution within the framework of classical logic and the infinitesimal calculus. In section 3 I point out a number of logical errors in their analysis, and in section 4 present a concrete counter-example, a simple problem in Newtonian kinematics which I demonstrate can be solved entirely within classical logic. In section 5 I discuss Sandewall's ideas, and once again use a simple problem in kinematics as an illustration of my objections; this time, I show how a few additional axioms force the framework to collapse to the classical one, which I claim will in general be the case. In the last section I sum up my conclusions.

2 Shoham and McDermott: the "Extended Prediction Problem"

I will start by examining the work of Yoav Shoham and Drew McDermott ([Shoham and McDermott 88]; hereafter S&McD). S&McD take as their point of departure the notorious "Frame Problem", and pay particular attention to the question of formulating it in continuous time. They begin by suggesting that it may best be dealt with by dividing it up into two distinct sub-problems, which they refer to as the "qualification problem" and the "extended prediction problem". It is the second of these that will be my primary concern here: I begin by summarizing S&McD's arguments.

S&McD's main claim is that there is a problem, the "extended prediction problem" (EPP). This is supposed to be the fact that it is difficult to formalize the process of making predictions over extended periods of time, if the axioms of a temporal theory are expressed as differential equations; S&McD claim moreover that the problem is not the product of a particular temporal formalism, but is general in nature. To substantiate this statement, they present arguments purporting to show that the problem occurs, not only in a conventional framework, but also in the Hayes "histories" formalism. In a later paper, Shoham [88] then goes on to describe his logic of "Chronological Ignorance" (CI), which (among other things) is supposed to provide a solution to the EPP.

What is the actual problem supposed to be? S&McD are happy to agree that Newtonian mechanics is *in principle* capable of describing the behaviour of dynamic systems in continuous time (p. 52-53)¹; however, they also claim that there is no well-defined associated *computational mechanism* which formalizes the process of making predictions. Considering the concrete problem of predicting the collision of two billiard balls, they write (p. 54):

The "prediction," however, is purely model-theoretic. No attention was paid to the problem of actually *computing* the point of collision. In fact, it is very unclear how to perform the computation, since all axioms refer to time *points*. Somehow we must identify the "interesting" points in time or space, and interpolate between them. The problem seems a little circular, though, since the identity of the interesting points depends on the integration. For example, understanding where the two balls are heading logically precedes the identification of the collision: if we don't know that the two balls are rolling towards each other, there is no reason to expect a something interesting at the actual collision point.

How do people solve such physics problems? The inevitable answer seems to be that they "visualize" the problem, identify a solution in some mysterious ("analog") way, and only then validate the solution through physics... (italics in original)

I will take the two paragraphs just quoted as the kernel of S&McD's claim. Before saying anything else, I think that it is important to point out that it is an extremely *strong* claim: all sorts of people are in the business of doing temporal reasoning using differential equations, and many of them would be prepared to defend themselves against the accusation that they are doing anything that couldn't be formalised. When the accusation is moreover exemplified in the trivial problem of predicting the collision of two billiard balls, the feeling that one is on theoretically secure ground is so strong as to more or less amount to certainty. Although I naturally don't mean to imply that a feeling of certainty *proves* anything, this is worth saying, since it motivates most of the reasoning in the sequel. My counter-claim, then, will be that there is actually nothing mysterious about the process of making predictions about continuous-time processes, and that these can readily be formalized with no more theoretical apparatus than is afforded by classical logic, together with the differential and integral calculus. I will first point out what I regard as several concrete logical errors in S&McD's analysis of the EPP, both in the classical and the "histories" frameworks; later, I will go further to sketch how it is in fact perfectly possible to predict billiard-ball collisions, using only classical logic and well-defined and unmysterious methods of inference.

3 Specific criticism of Shoham and McDermott's arguments

I will start with S&McD's treatment of the classical framework. Firstly, in several places in the argument, it certainly appears as though S&McD are committing the cardinal sin of confusing "infinitesimal" with "very small". Look, for example, at the following passage from page 59:

The most conservative prediction refers to a very short interval of time, in fact an instantaneous one, but that makes it very hard to reason about more lengthy future periods. For example, if on the basis of observing a ball rolling we predict that it will roll just a little bit further, in order to predict that it will roll a long distance we must iterate this process many times (in fact, an infinite number of times). We will call this the *extended prediction problem*.

¹All page references in sections 2, 3 and 4 are to [Shoham & McDermott 88] except where otherwise stated.

Now at risk of stating the obvious, it is not correct to say that a differential equation licences prediction "a little bit forwards", and then talk about doing this "many times - in fact an infinite number of times". Differential equations say things about the instantaneous rate of change of functions; to make predictions about extended periods they must be *integrated*. The integration will hold over a period if the differential equation holds over the same period, but the *length* of the period is completely irrelevant.

This is not mere pedantry; S&McD's lack of precision in expressing themselves is obscuring a crucial point. Since the differential equations don't directly allow forward prediction in the first place, the problem is not one of making an inefficient process more effective, predicting over a long interval rather than a short one. The problem is rather how we can justify prediction over any period at all. This is very much at odds with, for example, the following passage from p. 60 (my italics):

To summarize, the general extended prediction problem is that although we may be able to make predictions about *short* future intervals, we might have to make a *whole lot of them* before we can predict anything about a substantial portion of the future.

All right, so why don't we just integrate the differential equations then? Now S&McD have another argument in reserve; as we have already seen in the passage quoted in section 2 above, they claim that we don't know what interval to integrate over. My second point is, very simply: *This is not a problem*. All that needs to be done is to perform the integration over an interval whose bounds are left unspecified, except by the restriction that the differential equation should hold within them; this is exactly what applied mathematicians normally do in practice. Call our bounds t_1 , t_2 : then what we get is a logical formula of the form

*conditions on what holds at t_1 &
the differential equations hold between t_1 and $t_2 \rightarrow$
conditions on what holds at each point between t_1 and t_2 .*

By using these formulas, together with other facts, we can deduce the maximal t_1 and t_2 over which the integration is valid. In the next section I will illustrate how this is done for the problem with the billiard balls.

My third point concerns the notion of "potential history", which is, I claim, a somewhat misleading concept. Instead of talking about "the way things would turn out if nothing happened" (Shoham's definition of a "potential history"), it will be quite enough to take "the way things *actually* turn out *until* something happens". Then it will be possible to reason that either

- i) nothing ever does "happen"
- ii) there is a first thing that "happens"

In case ii), we will be able to deduce things about when the aforementioned "first thing" occurs. If this sounds cryptic, the example in section 4 should make things clearer.

I now move on to the reformulation of the problem in Hayes's "histories" framework, dealt with by S&McD in their section 1.2. The logical fallacies here are of a similar type to those I have just pointed out, but since the integration has in effect already been performed they are of a more transparent nature.

In one sentence: the way in which S&McD use histories to express the problem isn't the right one. To back up this claim, let's start by reviewing the problem. The initial data is that there are two

ROLLING histories, H11 and H21, of which we are given the prefixes, H11' and H21'. (See figures 1, 2 and 3, adapted from S&McD's diagrams 4, 5 and 6).

Figure 1

S&McD don't actually define exactly what they mean by a "prefix", but I suggest the following: it is the intersection of the history with some suitable given portion S of space-time. A simple way of defining S would be to let it be bounded above and below in time by two closely-spaced instants near the beginning of the period under consideration. Anyway, S&McD now go on to say that we want to predict two "new" ROLLING histories. They then inquire what these new histories should look like: either they will extend up to the collision point, or they won't. I agree with their objection that the second alternative merely postpones the problem one step, but their analysis of the first alternative quite fails to hold water.

Figure 2

S&McD's point here is that what we want to say intuitively is that "the histories persist for as long as possible" - i.e. until they collide with something - and that "there only are two histories". They claim that there is a difficulty with the second part; that there are, actually, a lot more histories lying around, like for example the histories H11" and H21" which follow just after H11' and H21'. But this is just playing with words; obviously, every history contains an infinite number of subhistories, so counting all histories can never get us anywhere. What we are interested in are the *maximal* histories in a given bounded chunk of space-time, which in this domain *are* going to be finite in number. Now we can say that there are exactly two maximal histories in S (the chunk we intersected with in the last paragraph to define our prefixes); these are by construction H11' and H21'.

Figure 3

If we then move on to consider a larger chunk of space-time (call it S'), $H11'$ and $H21'$ are in general no longer going to be maximal. They will, however, be included in two unique maximal histories¹, which in a sufficiently large S' will be precisely $H11$ and $H21$. It is then fairly clear how to express our laws of physics so as to make things work. The rule we need is going to be something like the following:

Let t_1 and t_2 (with $t_1 < t_2$) be two times, and let S be the region of space-time bounded by t_1 and t_2 . Assume that all the maximal histories in S are ROLLING histories, which touch both boundaries and not each other, and that there are exactly N such histories. If there are any collisions after t_2 , call the earliest time at which one occurs T , and call the region on space-time bounded by t_1 and T , S' . Then there are exactly $N+M$ maximal histories in S' , of which M are COLLISION histories and N are ROLLING histories. $M < N$, the intersection of each of the ROLLING histories with S is a distinct maximal history in S , and each of the COLLISION histories occurs at time T .

Together with the rule that COLLISION histories occur precisely when ROLLING histories meet, this will enable us to predict when the first collision occurs, using methods that essentially consist of little more than geometrical calculations in a three-dimensional Euclidean space.

4 Reasoning in continuous time with classical logic: an example

Consider the situation illustrated in figure 4 below. (A propos S&McD's remarks above about "visualization" (p. 54): the diagram is purely for the benefit of human readers, and as will be seen contributes nothing to the proof). We have a billiard table, on which a system of Cartesian coordinates is defined. At time $t=0$ there are three balls: A at $(0,2)$ with velocity vector $(0,-2)$; B at $(1,1)$ with velocity vector $(0,-2)$; and C at $(2,0)$ with velocity vector $(-2,0)$. We will assume that balls are point objects which collide iff their positions coincide, and that the only forces on balls are those obtaining at times of collision. Given these assumptions, we want to know whether there will be any collisions, and if so when the first one will occur. Since we aren't going to reason past this point, we don't need to say anything about how balls behave after a collision, whether they bounce, stick together, smash or whatever. We reason as follows.

¹Proving the truth of this statement would demand a full axiomatization of the histories framework, something that is obviously impractical here. However, the key step would presumably be an axiom to the effect that the union of two connected histories of the same kind was a history.

Figure 4

1) Either no collision will occur, or else there will be a first collision. In this case, one of the following will be true; A and B will be involved in it, B and C will be involved in it, or C and A will be involved in it. We refer to these four possibilities as No_collision, AB_first, BC_first and AC_first, and we refer to the time of the first collision as t_c with the convention that $t_c = \infty$ if No_collision obtains. (This is only to make the proof a little more elegant).

2) For $0 \leq t < t_c$ no force acts on the balls. So by Newton's laws, their velocities will be constant during this period, and thus by performing an elementary integration we have that A's position at time t is $(0, 2-2t)$, B's is $(1, 1-2t)$ and C's is $(2-2t, 0)$, $0 \leq t < t_c$.

3) We want to prove that No_collision doesn't hold, so we use reduction ad absurdum and assume that it does. Thus the positions are as given by 2) for all positive t . We now want to prove that a collision *does* take place to get our contradiction; specifically, we try to prove that A will collide with C. This will be so if we can find a positive t such that

$$(0, 2-2t) = (2-2t, 0)$$

Elementary algebra shows that $t = 1$ is a (in fact, the only) solution. So No_collision doesn't hold.

4) We now prove that AB_first doesn't hold. Again, suppose it did. Then the moment of collision is given by the equation

$$(0, 2-2t_c) = (1, 1-2t_c)$$

This gives us that $0 = 1$, a contradiction. So AB_first doesn't hold either.

5) Now prove that AC_first doesn't hold. Once more, suppose it did. Then just as above, we have that t_c is given by

$$(0, 2-2t_c) = (2-2t_c, 0)$$

Algebra gives us that $t_c = 1$. We must now establish that some other collision occurred at some earlier time, to obtain our contradiction. Specifically, we try to prove that B will collide with C. This will be so if we can find a t' such that

$$(1, 1-2t') = (2-2t', 0)$$

and $0 \leq t' < 1$. But $t = 0.5$ is such a solution, and once again we have a contradiction.

6) We have proved that there will be a collision, so there must be a first collision. Since AB_first and AC_first have been proved impossible, by elimination we have BC_first. Doing the same bit of algebra as in 5) shows that it occurs at $t = 0.5$. QED

Let us consider the structure of this proof. First, in 1), we hypothesize a time-point t_c , which is when the first thing is going to "happen". The important thing to notice here is that we don't yet say what t_c 's *value* is; we define it in terms of its properties, namely that a collision occurs then, and that none occur before. Then in 2) we integrate the differential equations to get the interval-based information that will allow us to make predictions. If we had been working in a histories formalism, the equations would as indicated earlier "already be integrated", and this step would have been superfluous. Having got this far, the remaining steps 3) to 6) are just ordinary monotonic classical logic, and consist of a proof that t_c as described actually does exist, together with a computation of its value. It is clear that the methods used are quite general, and in no way make special use of the billiard-ball scenario. One incidental point is also worth noting explicitly: the collision is shown to have occurred at $t=0.5$, demonstrating that we really have moved outside the integers.

To point the moral, the proof above demonstrates that classical logic and differential calculus are at any rate *sufficient* to solve problems of this kind. Shoham (personal communication) has however advanced another criticism: he claims that, although the approach I have just demonstrated is possible, it is less *efficient* than using CI.

My answer to Shoham's objection is twofold. Firstly, this is not what is being said in the original paper; it is a separate issue, which as far as I can see is nothing to do with the EPP. Secondly, Shoham has still to demonstrate that proofs of this kind can be carried out *at all* in CI. As he admits himself ([Shoham 1988], p. 320), the methods he has so far developed are completely dependant on the use of a discrete model of time; doing CI in continuous time would require the use of different algorithms, the complexity of which is thus completely unknown. Shoham might be right, but he has to present evidence to prove it; to use a metaphor from another game, the ball is now back in his court.

Before concluding this part of the paper, I hope that the reader will pardon a short historical digression. First, I would like to stress that the arguments presented above should not in any way be regarded as startling or unexpected; they are simply the defence of what most mathematicians and physicists would unhesitatingly call the common-sense view, namely that there is no longer anything mysterious about the EPP. If we adopt a broader perspective, however, we can see that the EPP *used* to be a major problem. It is in fact closely related to Zeno's paradox, something that caused philosophers difficulties from Zeno's time until the seventeenth century; until then, nobody even came close to explaining how it was possible to use logic to reason rigorously about continuous change. Indeed, many prominent thinkers went on record as claiming that such things were impossible in principle.

The first person to give a plausible account of mathematical reasoning about continuous processes was Newton, and even *he* was unable to do this in a satisfactorily formal way; this was not achieved until the nineteenth century analysts - people like Bolzano, Dedekind, Riemann and Cauchy - finally managed to put real analysis onto a sound logical footing. Not being an expert on the history of mathematics, I can't say with confidence just when the whole enterprise was completed; but I would be prepared to guess that Russell and Whitehead still had to tie up a few

loose ends in the *Principia Mathematica*. The whole process, in other words, took over two hundred years.

To sum up, then, the EPP is undoubtedly an extremely important and difficult problem. It is, however, a problem that has been *solved*, at least to the extent of transferring it from the province of philosophy to that of science, mathematics and logic: completing this task was exactly the achievement of the program sketched above. None the less, we still have the important question of *efficiency* left to consider; this brings us to the following section.

5 Sandewall: Reasoning about almost continuous systems

We now move on to discuss Sandewall's work [Sandewall 89a, 89b]. In contrast to Shoham and McDermott, Sandewall is quite willing to countenance the use of the infinitesimal calculus to model continuous physical processes. He contends, however, that calculus on its own is not enough, since it cannot easily describe the behaviour of the system at points where the values of parameters become discontinuous. Such discontinuities will in general arise in interesting physical problems, and it seems most natural to model them using some sort of discrete logic. So far, Sandewall's arguments seem eminently reasonable: however, his next step is to suggest that the most suitable logic will be a non-monotonic one, which he refers to as *chronological minimization of discontinuities* (CMD). His argument justifying this decision is as follows:

... a non-monotonic logic is used, since we need to state a default that the left limit value and the right limit value [of a parameter] are equal even in breakpoints, if it is consistent with the axioms for them to be so ... The proposed preference criterion can be viewed as a generalization of frame-problem persistence, from the classical view in AI of a discrete time-axis and discrete properties, to our approach using real-valued time and real-valued, piecewise continuous parameters. (p. 412)¹

From this passage, it is not entirely clear why a non-monotonic logic is deemed to be necessary. Another quote from a later paper expands on the point to some extent:

Of course the criterion for a proposed definition of semantic entailment is not that it in itself obtains the correct model set, since what models are obtained depends also on the axioms. For the conventional "frame problem", for example, it is perfectly possible to obtain the correct model set with standard, monotonic entailment, but the problem is that it may be very cumbersome to write out the axioms. The criterion for a definition of entailment is therefore whether it makes it easy to write out axiomatizations which obtain the right model sets. This is the claim that is tentatively made for chronological minimization of discontinuities. [Sandewall 89b: 896]

Sandewall's claim would thus appear to be roughly the following: that although it might well be *possible* to formulate (and presumably also solve) problems of this kind in classical logic, it is none the less *simpler* to do so in CMD. This claim is far more modest than those made above by Shoham and McDermott, and is consequently by no means as easy to refute. None the less, we shall by examining one of Sandewall's own examples show that he is standing on distinctly shaky ground.

The idealized scenario shown in figure 5 is taken from [Sandewall 89a]. The problem is formulated as follows:

¹ All page references in this section are to [Sandewall 89a] unless otherwise stated.

A ball is moving along a horizontal "plane" (actually a horizontal line) towards a shaft with vertical sides and a horizontal bottom. Idealized physical laws are assumed: the ball has zero size; there is no drag so the ball's horizontal velocity is constant until it reaches an obstacle, and when the ball reaches the pit its vertical acceleration changes instantly from zero to -9.81 . The ball bounces perfectly against the walls of the pit, so that its vertical velocity does not change and its horizontal velocity reverses its sign but keeps the magnitude. Also the bottom of the shaft absorbs impact perfectly, so that the ball stops its vertical movement without bounces when it reaches the bottom. (p. 415)

Figure 5

The problem is to justify the intuitive conclusion that the ball will eventually end up at the bottom of the shaft. Before examining the formal description of the problem, it will first be helpful to look a little more closely at the way in which CMD defines entailment. The basic idea (as the name would suggest) is closely related to Shoham's logic of chronological minimization [Shoham 88b]: as in Shoham's approach there is a *preference relation* between models, and only maximally preferred models are considered when computing entailment. Sandewall gives the following informal definition and motivation of the preference relation (it is defined formally elsewhere in the paper):

... If there is some way of satisfying the axioms without discontinuity then such an interpretation is preferred. If discontinuities are necessary i.e. no interpretation without discontinuity satisfies all the axioms, then the definition prefers to have the discontinuities as late as possible, and secondarily it minimizes the set of discontinuities which do occur.

... For example by this definition of entailment, the axioms characterizing the ball-and-shaft scenario will entail that the temperature of the ball is continuous as the ball begins going into the shaft, since there is no axiom that forces the temperature to be discontinuous there.

The reason for having continuity as a default is not only to deal with other, independent parameters for the same object, such as temperature, but also and perhaps more importantly to deal with multiple objects whose discontinuities occur independently. (p. 417)

This all sounds very good. If Sandewall's system really had the attractive properties claimed here, then there would obviously be a fair case to be made for using non-monotonic logic; but as we shall soon see, things in fact work out rather less smoothly. This will become apparent when we now examine Sandewall's formalization of the problem, which we reproduce below. We have the following axioms:

1.	$\text{Supp}(x_b, y_b) \rightarrow \partial^2 y_b^r = 0$	;The ball has no vertical acceleration when supported
2.	$\text{Wall}(x_b, y_b) \rightarrow \partial x_b^l = -\partial x_b^r$	;The ball bounces perfectly at the wall
3.	$\partial^2 x_b = 0$	;The ball has no horizontal acceleration
4.	$\neg \text{Supp}(x_b, y_b) \rightarrow \partial^2 y_b = -g$	;The ball has vertical acceleration $-g$ when unsupported
5.	$C(y_b) \ \& \ C(y_b)$	;The ball's spatial position is continuous
6.	$\neg \text{Wall}(x_b, y_b) \rightarrow C(\partial x_b)$	;The ball's horizontal velocity is continuous except at
;the wall

(There are also some axioms defining the predicates Supp and Wall)

Now even in Sandewall's tiny example, things are already starting to go wrong. This becomes plain when we consider the justification for including axiom 6: as explained on page 418, it is needed if we are to make the right prediction about what will happen when the ball reaches the edge of the shaft. There has to be a "breakpoint" here: either the ball's horizontal velocity or its vertical acceleration will change discontinuously. However, CMD is unable to prefer to keep the velocity continuous without an additional axiom which explicitly removes the other possibility; this is what axiom 6 is for. Without it, there will be a model where the ball spontaneously reverses direction and starts moving left again on reaching the edge.

This analysis of the rationale behind axiom 6 reveals that the problem it is intended to solve is, unfortunately enough, an extremely general one; in fact, it becomes natural to wonder why there is no need for a further axiom, which will explicitly restrict the situations in which the vertical velocity component can change discontinuously. On closer consideration, it turns out that this is more or less an accident: it just so happens in our particular scenario that vertical motion can never carry the ball over a boundary where some other parameter undergoes discontinuous change.

Figure 6.

The ball is warmed by an infra-red beam on its way down, but its trajectory should be the same.

A small alteration in the problem will be quite enough to remove this fortuitous circumstance. Suppose (just as Sandewall suggests in the passage from p. 417 quoted above) that we also want to think about the ball's temperature. We will assume the modified scenario illustrated in figure 6: when the ball's y -coordinate is greater than e , it loses heat by radiation at a rate proportional to the fourth power of its absolute temperature, but when it is below e it is also warmed by an infra-red

beam which heats it at a constant rate w . Calling the ball's temperature Θ_b , we have the additional axioms

10. $\neg \text{inbeam}(x_b, y_b) \rightarrow \partial \Theta_b = -k\Theta_b^4$;Radiation cooling
11. $\text{inbeam}(x_b, y_b) \rightarrow \partial \Theta_b = w - k\Theta_b^4$;Radiation cooling plus warming from beam
12. $\text{inbeam}(x_b, y_b) \leftrightarrow 0 < y_b < e \ \& \ a \leq x_b \leq b$

Things have thus been set up so that there will be a discontinuous change in $\partial \Theta_b$ when the ball falls into the beam, and now we get exactly the same problem with the vertical velocity component as we did previously with the horizontal one. As things stand, there will be valid models in which the ball, on reaching the beam, spontaneously begins moving upwards so as to stay out of it (figure 7); the only obvious way to block them will be to add another formula analogous to axiom 6, like for example 6a:

- 6a. $\neg \text{Supp}(x_b, y_b) \rightarrow C(\partial y_b)$;The ball's vertical velocity is continuous except on
;hitting the floor

What is really disquieting about the example is that it completely violates the intuitive assumption that the rate of change of the ball's temperature cannot affect its velocity. In general, it seems to show that a *global* analysis of the problem will be necessary if we are to know whether we can dispense with a "continuity" axiom like 6 or 6a: this negates the potential utility of the whole framework, since it would normally be simpler just to add all the continuity axioms immediately and work within classical logic. The obvious moral is that non-monotonic entailment is buying us very little - far less than would justify abandoning the well-understood framework like classical logic.

Figure 7.

An unintended model: the ball reverses direction spontaneously, to stay out of the beam.

6 Conclusions

In summary, I believe I have shown that there is so far no strong case for using non-monotonic logic in the continuous-time temporal reasoning domain. Shoham and McDermott attempt to demonstrate the insufficiency of classical logic, but their arguments are quite simply incorrect. Sandewall, in a rather more restrained vein, suggests that non-monotonic logic is possibly superior in terms of compactness of notation, since it can be used to obviate the necessity of explicitly specifying the regions in which parameters are continuous. However, the drawback with this

approach is that it difficult to ascertain precisely when it is possible to dispense with "continuity" axioms without performing a global classical-logic analysis of the problem; since this is precisely what non-monotonic logic is supposed to avoid, it is hard to see that it will in fact help.

Naturally, it is quite conceivable that some other way of using non-monotonic logic may turn out to offer concrete advantages, and it is indeed impossible in principle to refute such a statement. At the moment, though, it seems clear that the onus of proof is firmly on the non-monotonic logicians.

Acknowledgements

I would very much like to thank Yoav Shoham for his unfailingly courteous and helpful attitude in the face of hostile criticism. The comparison between the EPP and Zeno's paradox at the end of section 4 is due to Rune Gustavsson, and the paper as a whole benefited immensely from numerous discussions with Annika Wærn.

References

- [Sandewall 89a] Erik Sandewall. Combining logic and differential equations for describing real-world systems. In R.J. Brachmann, H. Levesque and R. Reiter (eds.) *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, 412-420, Morgan Kaufman, 1989
- [Sandewall 89b] Erik Sandewall. Filter preferential entailment for the logic of action in almost continuous worlds. *Proc 11th IJCAI*, 894-899, Morgan Kaufman, 1989
- [Shoham & McDermott 88] Yoav Shoham & Drew McDermott. Problems in Formal Temporal Reasoning. *Artificial Intelligence*, 36:49-61, 1988
- [Shoham 88] Yoav Shoham. Chronological Ignorance: Experiments in Nonmonotonic Temporal Reasoning. *Artificial Intelligence*, 36:279-331, 1988
- [van Bentham 83] Johan van Bentham. *The Logic of Time*. D. Reidel, 1983