Combining logic and differential equations for describing real-world systems

Erik Sandewall

Abstract

The paper shows how to combine non-monotonic temporal logic with differential calculus. A temporal logic is defined where time is real-valued and not discrete, and where real-valued, continuous parameters are used with their derivatives. Differential equations can therefore be used directly as axioms, and need not be transformed into confluences. This logic is used for characterizing common-sense physical systems where some parameters, or some of their derivatives, are occasionally discontinuous. Differential calculus then serves for characterizing the parameters during time-segments where they are continuous, and logic is used for characterizing the parameters around the discontinuity points. Models and preferential entailment is defined for this logic. For a simple scenario example (ball falling into shaft) it is shown what geometrical and physical axioms are needed, and how the axioms preferentially entail the desired common-sense conclusion.

This research was supported by the Swedish Board of Technical Development and the IT4 research program.

Appeared in Proceedings of the 1989 International Joint Conference on Knowledge Representation and Reasoning.

IDA Technical Report 1989
LiTH-IDA-R-89-38
ISSN-0281-4250

Department of Computer
and Information Science,
Linköping University,
S-581 83 Linkping, Sweden

1 Introduction and Approach

This paper describes how simple physical systems, of the kinds that we all understand and reason about by common sense, can be characterized and analyzed using a combination of differential equations and formal logic. The proposed method is an alternative to the contemporary standard approach in qualitative reasoning, which uses discrete value spaces[Hay79] and confluences[dKB85].

The method applies to *piecewise continuous* physical systems, i.e. systems whose possible histories over time may contain a number of significant timepoints called *breakpoints*, and where all parameters[1] of the system are assumed to be continuous in the intervals between the breakpoints. At the breakpoints however the parameters may have discontinuities: the *left limit value* $a^l(t)$ and the *right limit value* $a^r(t)$ are defined for the parameter a in the breakpoint t, but need not be equal.

One example where breakpoints arise is when a held object is let go, so that its vertical acceleration (i.e. the second derivative of its vertical coordinate) changes instantly from 0 to -9.81 metric. Another example is when an amount of water is being heated and reaches boiling temperature, at which point the first derivative of its temperature switches from non-zero to zero, in a sufficiently simplified account.

Differential calculus is a fine instrument for characterizing the parameters in their continuous intervals between the breakpoints, but by its very nature the differential calculus does not deal well with discontinuities. The method described here is to use logic formulas (= wffs) in a suitable temporal logic for characterizing the behavior of the quantitative parameters around the breakpoints, and in particular for constraining the left and right limit values of parameters in breakpoints.

Furthermore a non-monotonic logic is used, in order to state as a default that the left limit value and right limit value are equal even in breakpoints if it is consistent with the axioms for them to be so. More precisely, we use preferential entailment[Sho88] under a preference relation which chronologically minimizes the instances of parameter discontinuity. The proposed preference criterium can be seen as a generalization of frame-problem persistence, from the classical view in A.I. of a discrete time axis and discrete properties, to our approach using real-valued time and real-valued, piecewise continuous parameters.

One can think of total histories for such a system as being constructed from parameter segments, each parameter segment being a continuous function over a finite interval on the real time axis. Each available parameter segment is characterized using differential equations. Non-monotonic logic, using appropriate axioms describing the object system and expressing physical laws, is used as the gluing agent which puts the parameter segments together.

[1] The set of parameters is defined so that the derivative of a parameter is also a parameter, recursively. Parameters are real-valued functions of time, corresponding to what is often called fluents in the A.I. literature.

2 What is wrong with the standard approach to qualitative reasoning

The method described here is proposed as an alternative to the present standard approach to qualitative reasoning. The standard approach has evolved through a sequence of significant papers[Hay79, dKB85, For81, Kui82], and has developed a consensus about its basic tenets. Before the new method is described in the next section, we shall briefly discuss the reasons for trying something different from the standard. - The abbreviation QR will be used for qualitative reasoning, and DC for differential calculus.

2.1 Need for a unified theory of temporal reasoning

"Qualitative reasoning", "temporal reasoning", and "knowledge based planning" are now three distinct areas in A.I. with very little technical overlap. However they are all concerned with phenomena over time, and it would be very useful to have a unified theory, presumably based on a temporal logic, which includes stringent accounts of both qualitative reasoning and planning. The research reported here is one step towards such a unified theory.

2.2 Conceptual economy of being consistent with other disciplines

The standard approach to qualitative reasoning involves the reconstruction, for discrete value spaces, of many concepts and results from DC: arithmetic for qualitative values, confluences, a discrete counterpart of the mean-value theorem, etc. This enterprise is by far not yet finished; it remains to map geometry, many remaining physical laws, and so on to the discrete-value domains. A priori it is much more elegant to use the original domains of real-valued time and real-valued parameters as domains for the logic, and to use the original differential equations as axioms.

2.3 Partial information must influence axioms, not interpretations

One of the common arguments for the standard approach to QR is that DC descriptions are over-specific since they assume availability of information we do not have. This is a faulty argument, simply since equality is not the only allowable relation between quantities: inequalities are also a part of the notation! For example the information that a rolling ball moves forward with positive velocity, under the influence of friction, air resistance, and other more or less random factors, can be just as well conveyed by

$$\partial x > 0$$

as by

$$\partial x = \oplus$$

From the model theoretic point of view, the first formula allows us to use interpretations where parameters have real values, whereas the second formula forces us to use interpretations where the parameter values are taken from a discrete set. Of course the set of all models will have much greater cardinality if parameters have real values, but so what?

Scarcity of information about the object system must be reflected in a scarcity of information in the set of axioms describing that system, but need not be reflected in a scarcity of information in each particular interpretation for the axioms.

2.4 Differential calculus is declarative, and therefore multipurpose

Another repeatedly used argument for the standard approach is that DC descriptions can only be used for two purposes, namely for finding their analytic solution, or as a basis for numerical calculation to find the solution in a specific instance. This argument is also incorrect: deduction is a perfectly appropriate use of quantitative descriptions, particularly when inequalities are involved. One obvious example is when it is known that

$$f(a) \leq c \leq f(b)$$

$$a \leq b$$

and that f is continuous in the interval between a and b, one can conclude

$$\exists x[a \leq x \leq b \wedge f(x) = c]$$

Another example is when it is known that

$$d(t_0) < 0$$

where d is a parameter and therefore a function of time, and that

$$\partial d \geq k > 0$$

where k is a constant and is therefore independent of time, to conclude that

$$\exists t_1[t_1 > t_0 \wedge d(t_1) = 0]$$

at least with appropriate assumptions regarding breakpoints and discontinuities.

2.5 Comparing continuous and discrete value spaces for applications without breakpoints

There are of course also systems which do not have breakpoints, or (more precisely) which one chooses to describe as fully continuous and not having breakpoints. For such systems, the standard approach using algebraic operations in discrete value spaces has to compete with logical reasoning about inequalities (and of course also equalities) in DC. It has never been stated explicitly why one should prefer the former over the latter.

One possible reason might be notational elegance and conciseness. Discrete value spaces sometimes win in this respect, since one may write algebraic rules such as

$$\oplus * \ominus = \ominus$$

instead of the more clumsy logical formula

$$a > 0 \wedge b < 0 \Rightarrow a * b < 0$$

The algebraic formulation is particularly advantageous for larger expressions which can be evaluated with discrete values for each sub-expression. This however assumes that operations such as $+$ and $*$ are always defined in the discrete value space, and in particular that the value of a composite expression is a function of the values of its components. That is not the case since for example the value of

$$\oplus - \oplus$$

is undetermined in the discrete value space $\{\oplus, \mathbf{0}, \ominus\}$. The problem is aggravated when other landmark values than zero are admitted.

If discrete value spaces are preferred on the grounds of notational elegance and simplicity, then one would at least need to know that that advantage really holds in a systematic comparison, and that no significant loss of expressiveness occurs as a result. Neither of those issues has been studied (to my knowledge).

2.6 Inference capability

Yet another of the reasons commonly quoted for the standard approach to QR is that there are certain conclusions about the physical systems under consideration, which people can easily obtain by common sense, but which can not be obtained from the DC descriptions of such systems. Therefore, it is argued, one has to develop alternative and A.I.-style descriptions, based on conceptual representations of the types that people are assumed to use.

We have already argued that DC descriptions lend themselves to more kinds of symbolic manipulation than has often been assumed for the sake of the argument. However there is also a deeper issue: since the switch from a quantitative to a qualitative system description involves a *loss* of information, how can it possibly enable one to draw *more* conclusions than would otherwise have been possible? If something is being said in the qualitative formulation which can not be said in the quantitative one, then exactly what is that something?

The answer must be that additional deductive principles have been introduced in the QR representation using the qualitative domain, for example in the guise of constraint propagation mechanisms which are commonly used. The exact nature of the new mechanisms is however unclear, or at least has not usually been spelled out in clear logical terms in the QR literature.

Our position is that any additional deductive principle which is needed for common sense reasoning about physical or mechanical systems, should be developed, if at all possible, for use together with quantitative descriptions using DC, and not only for qualitative descriptions.

In particular, the results in the present paper show how frame- problem style persistence methods, can be adapted and generalized so as to also apply for continuous time and continuous parameters.

It is interesting to note that the use of discrete quantity spaces did not predominate as much in the early work on qualitative reasoning. The NEWTON system of de Kleer[dK77], for example, stores mathematical equations in frames. One of the key points with NEWTON is that it contains multiple problem solving techniques, including both qualitative and quantitative methods, which can be applied to the conventional mathematical representation. It is in later papers that de Kleer and other authors introduce the method of "preprocessing" the problem statement into formulations using a discrete quantity space, which does away with the possibility to use problem-solving methods that use the DC representation.

2.7 What is really needed for doing qualitative reasoning

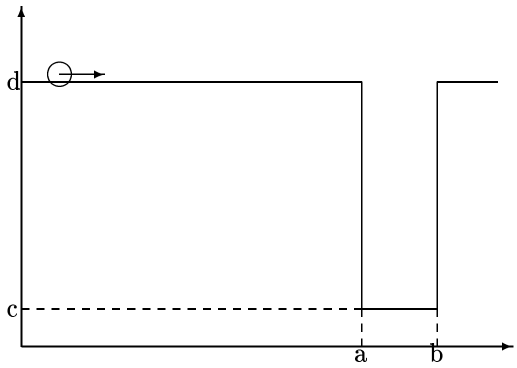
In summary, we claim that the standard approach to QR using discrete value spaces, confluences, etc. is a dead end. It originated from an incorrect analysis of whether one could, and why one could not obtain common sense conclusions from conventional descriptions of physical systems using differential calculus.

For fully continuous systems (without breakpoints i.e. all parameters are continuous at all times), we believe that reasoning about inequalities, equalities, and differential equations is sufficient for obtaining common-sense conclusions.

For object systems which are only piecewise continuous, something however has to be added, namely a descriptive machinery for characterizing - and constraining - the values of the parameters around the breakpoints. Switching from continuous to discrete value spaces is not an adequate response to that need, except if it were the only way to introduce logical constraints at all. In the present paper we show how logical constraints on breakpoint values can be directly combined with descriptions of the object in differential calculus.

3 Example scenario

Consider a simple, vertical two-dimensional mechanical system, such as the one illustrated in the following figure:



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 the ball stops its vertical movement without bounces when it reaches the bottom.

An obvious common sense conclusion from the scenario (scene description + "physical" laws) is that the ball will end up on the bottom of the shaft. We shall show below how that common sense conclusion is logically entailed by axioms expressed directly in real-valued calculus.

It may be objected that there are no discontinuities in actual physical systems. No actual ball starts falling instantly, nor does it make perfect bounces, or stop instantly on the bottom of shafts. This is true but beside the point, since the piecewise continuous description is a good approximation for many physical systems, and in particular can be expected to be good enough as a basis for reasoning.

In general, many interesting mechanical systems can be well characterized as having a number of "modes", "phases", or "states", where differential equations can be used to characterize the relationships between parameters within each of the modes with sufficiently good approximation. At the border between the states those approximations are not sufficient. The simplest possible way of characterizing the system around the shift of state is to assume that some of the parameters are discontinuous there. For common-sense reasoning, such simple approximations are of particular interest.

4 Interpretations and Formulas

We now proceed to the substance of the new method. In pursuing the approach that was outlined in the previous sections, the following three things need to be done first:

1. define the structure of interpretations in the logic
2. define the syntax of logic formulas (= wff)
3. define semantic entailment in this logic

We view the development of an inference system and inference procedure matching the semantic entailment as a secondary issue which can be pursued later on, once semantic entailment has been defined. We also consider it natural to define interpretations first and syntax afterwards.

4.1 Interpretations

An interpretation in our logic is a sixtuple

$$\langle M, U, S, R, Q, W \rangle$$

where M is a set of mode symbols, U is a set of parameter symbols closed under ∂, $S \subseteq \mathbf{R}$ is a "sparse" set of time-points (where $\mathbf{R}$ is the set of real numbers), namely the breakpoints, $R : (\mathbf{R} \times M) \rightarrow \{\mathcal{T}, \mathcal{F}\}$ assigns a truth-value to each mode at each point in time, $Q : (\mathbf{R} \times U) \rightarrow \mathbf{R}$ similarly assigns a value to each parameter at each point in time, and W is a mapping from temporal constant symbols to $\mathbf{R}$.

In addition interpretations must satisfy the following requirements:

1. $Q(t, u)$ as a function of t is continuous over every interval not containing a member of S,
2. if $t \notin S$ then $d/dt\ Q(t, u) = Q(t, \partial u)$,
3. $R(t, m)$ as a function of t is constant over every interval not containing a member of S.

It is in fact possible to get rid of M and R from the definition by encoding modes as parameters whose value is always either of 0 or 1, but only at the expense of understandability.

4.2 Syntax of logic formulas

We introduce the syntax for formulas by examples taken from the ball-and-shaft scenario above:

$$\Box[\partial^2 x_b = 0]$$

where $x_b \in U$ is a parameter or "fluent" representing the horizontal position of the ball, so that $Q(t, x_b)$ is the x-coordinate of the ball at time t. The set U was assumed to be closed under the syntactic operation of prefixing a ∂ symbol, so ∂x_b, $\partial^2 x_b$, etc. are also in U. The value of $Q(t, \partial^2 x_b)$ is the horizontal acceleration of the ball at time t if $t \notin S$, but is not constrained to equal the acceleration if $t \in S$. The operator $\Box$ means that the formula following it is true for all $t \in \mathbf{R}$, and has the weakest possible priority so that it extends to the end of the formula or to the next closing right parenthesis.

$$\Box[Wall(x_b, y_b) \Rightarrow \partial x_b^l = -\partial x_b^r \wedge \partial y_b^l = \partial y_b^r]$$

where $Wall(x, y)$ is true iff the point (x, y) in the vertical plane is a wall point, i.e. x is either a or b, and y is strictly between c and d. Also the l and r "exponents" are used to signify left and right limit value as follows:

$$Q(t, \partial x_b^l) = \lim_{\tau \to t-0} Q(\tau, \partial x_b)$$

is the horizontal velocity of the ball "just before" impact against the wall at time t. The whole formula expresses the idealized physical law that on impact against the wall, the ball reverses the sign of its horizontal velocity component instantly and retains its vertical velocity.

Notice that both the above axioms can be satisfied in the same interpretation. There is an apparent conflict because if the horizontal velocity has a discontinuity, as the second axiom implies, then the horizontal acceleration can not be always zero. However this is not a problem since $Q(t, \partial^2 x_b)$ is only required to equal the horizontal acceleration for those t not in S. If the interpretation is only constructed in such a way that the time of bounce t_0 is included in S, then $Q(t, \partial^2 x_b)$ may equal zero also for $t = t_0$ although $d/dt\ Q(t, \partial x_b)$ is not zero for $t = t_0$, and is in fact not even defined there.

$$\Box[Supp(x_b, y_b) \Rightarrow \partial y_b^r = 0]$$

where $Supp(x, y)$ is true for those points (x, y) located on a horizontal supporting "surface". This axiom says that an object which is already on such a supporting surface stays there, and that an object which falls onto the surface stops without bouncing back up again. (However, as we shall see below this axiom has to be formulated somewhat differently in order to exclude all interpretations where the object "falls through" the surface).

It is straightforward to extrapolate the appropriate syntax for logic formulas from these examples, and to complete the set of axioms characterizing the application. However what we have described so far does not allow one to refer to "observations" i.e. statements about the value of a parameter at a specific point in time. The notation so far is only sufficient for characterizing physical laws (such as the rule relating the parameters involved in a perfect bounce) and for describing the physical structure of the scenario, for example the existence and location of the shaft (assumed to be held fixed over time). We shall refer to this information as the *permanent axioms* and write it as Π.

4.3 Observation axioms

In addition to the permanent axioms there is also a need for axioms describing the state of the system at specific points in time. We shall refer to such axioms as observation axioms.

An appropriate notation for expressing observations may be as

$$[t_1]\ x_b < a \wedge y_b = d$$

for saying that the ball is on the supporting surface and to the left of the shaft at time t_1, or

$$[t_1]\ \partial x_b > 0$$

for saying that the ball is moving towards the right at time t_1. Formulas such as these may be assigned a value by interpretations as defined above, since the W component of interpretations assigns a time-point as value to each time-constant such as t_1.

Actually observations have to be recorded as logical formulas of the form

$$[t_1]\ \partial x_b > 0 \wedge C(\partial x_b)$$

where C is a predicate saying that the parameter appearing as its argument, is continuous at the present time-point. We will soon come to the reasons why this is necessary.

If Π is a set of permanent axioms, and Γ is a set of observation axioms for *one* time-point t_1, then we should expect the model set for $\Pi \cup \Gamma$ to be the set of all "histories" i.e. temporal developments which are compatible with the observations in Γ. The types of reasoning which are commonly called prediction (temporal projection) and postdiction (temporal explanation) can therefore be understood as identifying formulas α which satisfy

$$\Pi \cup \Gamma \models \exists t_2[t_1 < t_2 \wedge [t_2]\ \alpha]$$

and

$$\Pi \cup \Gamma \models \exists t_2[t_1 > t_2 \wedge [t_2]\ \alpha]$$

respectively.

Also if Γ is a set of observations at *two* points in time, the model set and the entailed conclusions can be interpreted as an explanation of observations (given that the facts in Γ have actually been observed, conclude what must have happened in between). If the logic is extended so that actions are also admitted, then the conclusions from $\Pi \cup \Gamma$ in this case may alternatively be interpreted as a plan (or more precisely, as a disjunction of all available plans) for getting from the first "observed" state to the second "observed" state. In this case the second state is of course the goal statement for the plan.

The case where Γ is the empty set, finally, corresponds to the reasoning process called *envisioning* by de Kleer[dK77]. In other words, the model set for Π alone should be the set of all possible histories in the world, as constrained by the laws of "physics" and physical structure, but not by any particular observations of the state of the system.

5 Preferred models and entailment

We have now defined the structure of interpretations, and (outlined) the syntax of logical formulae. We proceed to the third issue that was specified initially, namely defining the criteria for the model set, and therefore the entailment relation for the present type of logic.

The result is as follows. It is not adequate to use ordinary, monotonic entailment, where the set of models for the axioms is simply the set of those interpretations wherein all the axioms are true. The reason is that the axioms may specify under what circumstances a parameter must have a discontinuity, e.g. by specifying the relationship between its old and new value (like in the axiom for the perfect bounce above), but it is difficult or impossible to specify through separate axioms all the cases where a parameter should not have a discontinuity in those time points belonging to S. There is therefore a need for a non-monotonic principle saying that parameters are continuous at break-points whenever that is consistent with the axioms.

Non-monotonic inference can be characterized in a model theoretic framework using preference relations over interpretations, as has been demonstrated by Shoham[Sho88]. A preference relation $\ll$ over interpretations is defined as follows.

Definition. If $J = \langle M, U, S, R, Q, W \rangle$ is an interpretation, and t is a time point, then the *breakset* i.e. the set of discontinuities at time t in J is

$$breaks(J,t) = \{m \mid R^l(t,m) \neq R^r(t,m)\} \cup$$

$$\{u \mid Q^l(t,u) \neq Q^r(t,u)\}$$

Clearly if t is not in S then the breakset of J at t is empty.

Definition. If $J = \langle M, U, S, R, Q, W \rangle$ and $J' = \langle M', U', S', R', Q', W' \rangle$ are interpretations, then we write $J \ll J'$ and say that J is preferred over J' iff there is some time-point t_0 such that

1) for all $t < t_0$, all $m \in M$ and all $u \in U$, $R(t, m) = R'(t, m)$ and $Q(t, u) = Q'(t, u)$

2) $breaks(J, t) \subset breaks(J', t)$.

Definition. If Π is a set of permanent axioms i.e. axioms not containing any explicit references to time-points, and if α is another logic formula, we say that Π preferentially entails α and write $\Pi \models_{\ll} \alpha$ iff

$$Min(\ll, Mod(\Pi)) \subseteq Mod(\alpha)$$

where $Mod(\phi)$ is the set of all interpretations wherein ϕ resp. all members of ϕ are true, and $Min(\ll, s)$ is the set of all $\ll$-minimal members of the set s.

This is of course similar to the principle of chronological minimization as introduced by Shoham[Sho88].

The way this definition works is that 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 discontinuities 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.

The set S of breakpoints is itself not minimized. In particular if two interpretations are identical except that one of them has some extra breakpoints in S, but all the parameters are continuous in the extra breakpoints, then those two interpretations are equally preferred. Breakpoints without any discontinuity in them are insignificant.

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 in order to deal with other, independent parameters for the same object, such as temperature, but also and maybe more importantly for dealing with multiple objects whose discontinuities occur independently. In an extended ball-and-shaft scenario there may be several shafts, and several balls falling asynchronously into those shafts. One should not need to have axioms saying that one ball moves continuously when another ball bounces against the wall in a shaft somewhere else. This is very similar to the "frame problem" but for a continuous domain.

Notice also that the preference relation only compares interpretations which are identical up to the point of comparison. One might think of having a stronger preference relation which minimizes breaksets chronologically regardless of whether parameter values are equal. However with such a stronger preference relation, the axioms for the ball and shaft scenario would logically entail that the ball starts as far to the left as possible, and moves as slowly as possible, since that postpones as much as possible the time-point for the first discontinuity. With the definition given above that problem does not arise.

6 Axioms and conclusions in the ball and shaft example

Let us illustrate the proposed method for the ball and shaft example that was introduced in section 2. The following axioms express the idealized physical laws that are being used:

1. $\Box[Supp(x_b, y_b) \Rightarrow \partial^2 y_b^r = 0]$
2. $\Box[Wall(x_b, y_b) \Rightarrow \partial x_b^l = -\partial x_b^r]$
3. $\Box[\partial^2 x_b = 0]$
4. $\Box[\neg Supp(x_b, y_b) \Rightarrow \partial^2 y_b = -9.81]$
5. $\Box[C(x_b) \wedge C(y_b)]$
6. $\Box[\neg Wall(x_b, y_b) \Rightarrow C(\partial x_b)]$

The following axioms characterize the physical configuration of the surface and the shaft:

7. $\Box[a < b \wedge c < d]$
8. $\Box[Supp(x, y) \leftrightarrow (x < a \vee x > b) \wedge y = d \vee a \leq x \leq b \wedge y = c]$
9. $\Box[Wall(x, y) \leftrightarrow (x = a \vee x = b) \wedge c \leq y < d]$

Finally the following axioms characterize the initial state of the system at time t_1:

10. $[t_1]\ x_b < a \wedge \partial x_b > 0 \wedge y_b = d \wedge C(\partial x_b)$

Theorems of differential calculus are not needed for determining semantic entailment, because of the chosen definition of interpretations, where parameters are assumed to be continuous with all their derivatives except in the break points.

The geometric quantities a, b, c and d are assumed to be constants with respect to time. If one does not want to have a separate syntactic type for such constants, this condition may be expressed by

11. $\Box[C(a) \wedge \partial a = 0]$

and similarly for b, c and d.

Section 4 above contained some of these axioms as examples of the syntax. However the common sense versions of the axioms in section 4 have been revised slightly. In axiom 2 the consequent $\partial y_b^l = \partial y_b^r$ has been dropped since the same result is obtained by default. Also axiom 1 is now expressed in terms of the second derivative in order to eliminate some non-intended model histories.

Let us now see how these axioms preferentially entail that the ball will end up on the bottom of the shaft. Consider the set of all interpretations which satisy the scenario and observation axioms (7) - (11), which also assign the same value to t_1, and which have identical histories up to then for all the parameters. This set of interpretations will then be further reduced by the requirement to satisfy the first six axioms, and to be minimal with respect to the preference relation. The remaining set is the set of (preferred) *model histories.*

In every model history the ball moves right at constant speed to the shaft, i.e. until $x_b = a$. It is clearly possible to satisfy all the axioms in such an interpretation. Therefore, although there may be other interpretations which introduce a breakpoint and a discontinuity before the ball reaches the shaft, and although such interpretations may also satisfy all the axioms, they can not be minimal with respect to the preference relation.

Also, in every model history it must be the case that y_b is constantly $= d$ as long as the ball moves along the support. For, any interpretation where the ball changes its vertical position obtains a contradiction between (1) and (4) at the timepoint where the ball leaves the support.

According to DC there exists some timepoint t_2 where $x_b = a$. Necessarily t_2 must be a breakpoint, and some parameter must be discontinuous there, since otherwise some axiom is violated. Two possibilities are of interest:

a) ∂x_b has a discontinuity and becomes negative, all other parameters are continuous.

b) $\partial^2 y_b$ has a discontinuity and changes to -9.81, all other parameters are continuous.

However the first alternative is excluded by axiom (6), so only the second possibility remains. In fact, axiom (6) only serves the purpose of ruling out the first alternative in cases like this one, and is somewhat contrary to the spirit of our general approach since we would prefer to obtain parameter continuity by default.

Anyway, in all the model histories the ball continues with constant ∂x_b. It also has continuous ∂y_b (by default) when it starts going into the shaft, but of course discontinuous $\partial^2 y_b$.

The next breakpoint will be either at the right wall or on the bottom of the shaft, whichever comes first. Either of those conditions must occur within finite time, according to DC. Also, any interpretations with breakpoints and discontinuities along the way to the right wall or bottom, will be deselected by the preference condition. Consider those interpretations where the ball hits the right wall of the shaft first.

Axiom (2) can only be satisfied if there is a breakpoint at the time-point when the ball's position satisfies the predicate for the walls of the shaft, i.e. when the ball hits the (right) wall. At that breakpoint x_b and y_b are continuous (by 5), ∂x_b reverses its sign (by 2), ∂y_b is continuous (by default), and $\partial^2 x_b$ and $\partial^2 y_b$ are constant (by 3 and 4).

The ball then bounces back and forth between the walls of the shaft, with decreasing y_b. Eventually y_b reaches the value c, at which point there must again be a breakpoint. There x_b and y_b are again continuous, ∂x_b is continuous (by default), $\partial^2 x_b$ remains 0 (by 3), and y_b becomes constantly equal to c by the same argument about second derivatives as when the ball was travelling along the upper support.

In every model history there is therefore some time after which the ball is forever at the bottom of the shaft, and in fact it bounces forever back and forth there. This was the desired "common sense" conclusion, which we however obtained using non-monotonic logic and differential equations, and while retaining full mathematical precision.

7 Discussion

The semantics and semantic entailment that have been defined here does not contain the notion of actions. In a separate paper [San88] we describe an extension of the present logic that also contains actions. The extension is non-trivial, since it requires an entailment criterium which turns out to be significantly more complex than the one used here.

The axiomatization of the ball and shaft scenario uses idealized physical laws where e.g. the ball has constant horizontal velocity and constant vertical acceleration. A more realistic axiomatization would give bounds for those parameters, instead of giving their exact values. For example one of the "laws" could be that when the ball is unsupported, either its vertical acceleration is < -9, or its vertical velocity is negative, with a magnitude greater than some constant. Another axiom could say that when the ball bounces against the wall, its horizontal velocity has its sign reversed and its magnitude decreased. A change to such axioms makes the deductions a bit more cumbersome, but does not change the essence of what models are obtained and what is logically entailed by the axioms.

7.1 Observations and continuity

Observation axioms i.e. axioms describing the state of the system at a particular point in time, are crucial for example for stating the system's initial state, or for specifying the goal state(s) in the extension to planning. Observation axioms must always be written with a statement that the observed parameter is continuous at the time of observation, if it is, for example as in the ball-and-shaft scenario

$$[t_1]\ \partial x_b > 0 \wedge C(\partial x_b)$$

since otherwise the observation would not be enforced. Consider an interpretation where the history before t_1 arrives continously to a state such that $\partial x_b \leq 0$ at t_1. No interpretation of that form should be a model. However if it has t_1 as a breakpoint, and assigns a positive value to ∂x_b in t_1, then such an interpretation satisfies all the axioms except the one requiring ∂x_b to be continuous in t_1.

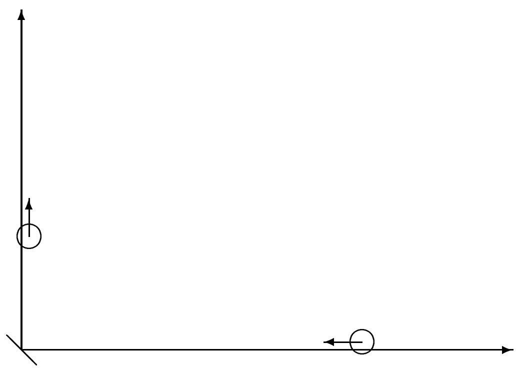
Therefore also the definition of the continuity predicate $C(\lambda)$ is that λ^l, λ, and λ^r must all be equal in the given time-point. It is not sufficient to only require that the left and right limit values be equal.

Of course if observations are recorded at a time of discontinuity for the parameter, they must specify whether they refer to the left or the right limit value.

7.2 Open and closed intervals

The main purpose of logic in the tandem of logic and DC, is to characterize the behavior of parameters around their discontinuities. To achieve that purpose we need the separate notation for referring to the left and right limit values a^l and a^r for a parameter a. This notation may seem a bit clumsy. However the problem is not a new one, and in particular one has the same problem when discrete value spaces are used. For example, if one wants to characterize the temperature of water which is brought to boil at time t_1, using a discrete value space consisting of < 100, $= 100$, and > 100 degrees Celsius, one must take care that the "value" is < 100 in the *open* interval ending in t_1, since presumably it is $= 100$ at the time t_1 itself.

Also in order to characterize with a discrete value space the x and y coordinates of a particle that travels along the x axis, and bounces in origo in order to continue along the y axis, as in the following figure:



,then it is not sufficient to only have two time intervals for before and after the bounce-time t_0. The x coordinate of the ball, x_b, is $\oplus$ at time $< t_0$, and is $\mathbf{0}$ at time $\geq t_0$. However y_b is $\mathbf{0}$ at time $\leq t_0$ and $\oplus$ at time $> t_0$. Therefore, in the standard approach to QR using discrete value spaces, one has to subdivide the time axis differently for characterizing x and y, or at least divide the time axis into the time-point of reflection, plus the two open intervals before and after that time-point.

Several previous authors have commented on the intricacies of dealing with parameters around the end-points of intervals. Allen[All84] ascribed such problems to the use of time-points, and quotes it as one of the reasons why a temporal logic should be based on intervals rather than time-points.

In the perspective of earlier work, it therefore seems that the notation proposed here, i.e. using time-points but with a formula syntax that allows explicit reference to the left and right limit values of a parameter at a time-point, represents a reasonable trade-off between notational precision and notational convenience.

7.3 Naturalness of the axioms

In declarative and logic based approaches, one ideally wishes that an axiomatization should simply be a set of statements each of which comes naturally as a description of the application domain. In practice, of course one must often compromise the ideal already in order to obtain a complete (or less incomplete) axiomatization, as well as for performance reasons in an implementation when one comes to that point.

In the axioms for the ball and shaft scenario only the former consideration was involved. Most of the axioms are the "natural" ones, but the ideal has not been attained completely. Axiom (1) had to be formulated in terms of the second derivative instead of the natural formulation using the first derivative, in order to rule out certain unintended model histories as we have seen. Also the extra axiom (6) had to be introduced for that same reason.

In other words, we have approached to, but not fully arrived at the ideal situation that one could take differential equations straight from the physics textbook and use them as axioms in the logic.

7.4 Remaining mathematical groundwork

In this paper we have skipped over certain basic mathematical issues which would have to be covered in a fully rigorous account. In particular, we have assumed that minimal models always exist (no infinite sequences of more and more preferred models, except if the sequence has a limit which is even more preferred). We have also ignored whether problems could arise if there are infinitely many breakpoints.

8 Conclusion and acknowledgements

Our grand goal is to develop a coherent theory for temporal reasoning, knowledge based planning, and qualitative reasoning. The present paper has shown how the basic intentions of qualitative reasoning can be achieved within a framework of non-monotonic temporal logic. The logical next steps are to proceed to planning, if possible using the same approach or an extension of it, and also to try to represent additional physical scenarios using the same method as for the ball and shaft scenario.

The work reported here has been catalyzed by regular discussions with Lennart Ljung and his colleagues in the Division of Automatic Control of our university. The significance of our group meetings is hereby acknowledged.

References

[All84] James F. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2):123–154, 1984.

[dK77] Johan de Kleer. Multiple representations of knowledge in a mechanics problem-solver. In *Proc. Fifth International Joint Conference on Artificial Intelligence*, pages 299–304, 1977.

[dKB85] Johan de Kleer and John Seeley Brown. A qualitative physics based on confluences. In Daniel G. Bobrow, editor, *Qualitative Reasoning about Physical Systems*, pages 7–83. MIT press, 1985.

[For81] Kenneth D. Forbus. Qualitative reasoning about physical processes. In *International Joint Conference on Artificial Intelligence*, 1981.

[Hay79] Patrick Hayes. The naive physics manifesto. In Donald Michie, editor, *Expert Systems in the Microelectronic Age*. Edinburgh University Press, 1979.

[Kui82] Benjamin Kuipers. Getting the envisionment right. In *Proc. AAAI -82*, pages 209–212, 1982.

[San88] Erik Sandewall. Filter preferential entailment for the logic of action in almost continuous worlds. Technical Report LAIC-IDA-88-TR29, IDA, 1988. Preliminary version.

[Sho88] Yoav Shoham. *Reasoning about Change*. MIT Press, 1988.