Contemplating the Addition of Stochastic Behaviour to Hybrid Event-B

Richard Banach
School of Computer Science,
University of Manchester, Manchester, U.K.
E-mail: banach@cs.man.ac.uk

***Abstract*—In real hybrid and cyberphysical systems, noise is a constant accompaniment to (and distraction from) the deterministic behaviour that is ideally desired. Nevertheless, most formalisms for such systems restrict to the deterministic realm. This also includes Hybrid Event-B, an extension of Event-B that caters for continuous behaviour as first class citizen. The incorporation of stochastic behaviour into Hybrid Event-B is investigated. Some essential elements of this enhancement are discussed, and a small case study is explored.**

I. Introduction

In today's ever-increasing interaction between digital devices and the physical world, formalisms are needed to express the more complex behaviours that this allows. Furthermore, these days, it is no longer sufficient to focus on isolated systems, as it is more and more the case that families of such systems are coupled together using communication networks, and can thus influence each others' working. So today *Cyber-Physical Systems* [1], [2], [3], [4] are the primary focus of attention. These new kinds of system throw up novel challenges in terms of design technique, as it is proving more and more difficult to ignore the continuous characteristics in their behaviours.

As soon as one contemplates including continuous behaviour in an essential way in the modelling of hybrid and cyber-physical systems, the problem of noise rears its head. And it does so in a manner that is qualitatively different to the role of noise in purely discrete systems. In discrete systems, the discrete state values are represented (physically) by well separated physical configurations of the implementation; it is the role of the designers of the implementation platform to ensure that this is so. Hence, although there is always some noise in any implementation, in a discrete system, its role *by the definition of what we mean by a discrete system* is sufficiently insignificant that it can be neglected.

By contrast, continuous behaviour proceeds by 'infinitesimally small steps' (a slogan that must be taken with a pinch of salt in the context of normal engineering descriptions). No matter how small the intrusion of noise into such a continuous process, in principle, the infinitesimally small increments of continuous behaviour become small enough that the noise fluctuations are not negligible in comparison with them. That said, it may nevertheless be the case that the macroscopic consequences of noise can be neglected in specific cases. Our point though, is that this is not true *a priori*, but must be established on a case by case basis.

Our focus in this paper is to explore the impact of these observations on the popular discretely based B-Method modelling and verification framework [5], [6]. Increasingly, applications of the B-Method's more modern incarnation, Event-B [6], involve continuous behaviour in some form, leading to the impingement of the issues mentioned on the development activity.

Hybrid Event-B [7] has been introduced to bring continuous capabilities to the traditionally based discrete Event-B, in order to address some of the challenges referred to. For applications of this formalism see [8], [9], [10]. As described in the next section, traditional discrete Event-B events serve as the 'mode events' that interleave the 'pliant events' of Hybrid Event-B. The latter express the continuously varying behaviour of a hybrid formalism that includes both kinds of event. In this manner, a rigorous link can be made between continuous and discrete update, as needed in contemporary applications.

In [7], and in the applications of Hybrid Event-B to date, continuous behaviour is idealised and noise-free, and so is restricted to situations in which such a perspective is legitimate. In this paper we contemplate extending the framework of [7] to allow noise to occur as a first class citizen.

In the kinds of scenario that the Hybrid Event-B formalism is most useful for, as exemplified by the kinds of automotive and similar applications in [8], [9], [10], the mode transitions are (almost) all brought about by stimuli from the environment, such as specific user actions. Accordingly, in this initial study, mode transitions remain deterministic, outside the stochastic domain, and we confine the stochastic aspects of our work to noise processes acting during the pliant transitions, even though there is no theoretical necessity for such a restriction.

The rest of the paper is as follows. Section II gives a brief description of non-stochastic Hybrid Event-B that is sufficient for the remainder. Section III introduces the stochastic perspective, and with that background, the impact on the Hybrid Event-B is described. Section IV covers a small case study. Section V concludes.

II. An Outline of Conventional Hybrid Event-B

In this section we give an overview of conventional, i.e. non-stochastic, Hybrid Event-B. In Fig. 1 we see a bare bones Hybrid Event-B machine, *HyEvBMch*. To save space later, we also include in the syntactic template of Fig. 1, those additional elements necessitated by the incorporation of the stochastic

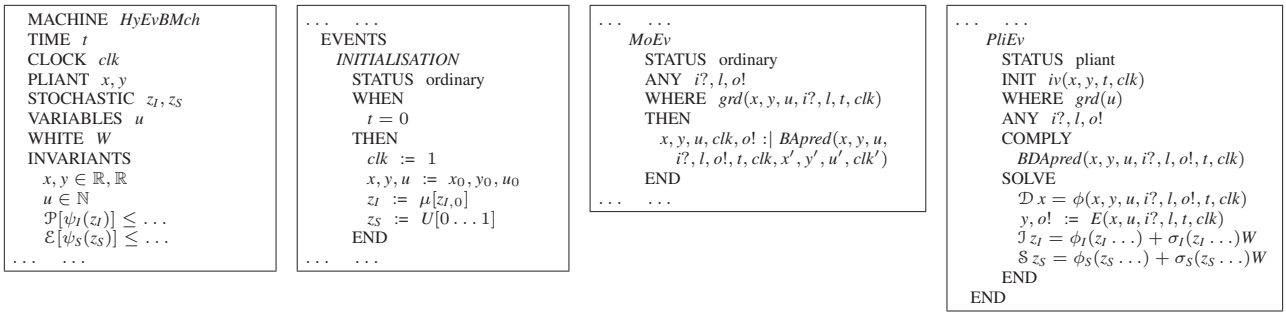

Fig. 1. A schematic Hybrid Event-B machine, including stochastic elements.

dimension. However, in this section, we just skip over them, and return to discuss them more thoroughly in Section III-B.

HyEvBMch starts with declarations of time and of a clock. In Hybrid Event-B time is a first class citizen in that all variables are functions of time, whether explicitly or implicitly. However time is special, being read-only, never being assigned, since time cannot be controlled by any human-designed engineering process. Clocks allow a bit more flexibility, since they are assumed to increase their value at the same rate that time does by default, but may be (re)set during mode events (see below).

(Non-stochastic) variables are of two kinds. There are mode variables (like u, declared as usual) which take their values in discrete sets and change their values via discontinuous assignment in mode events. There are also pliant variables (such as x, y), declared in the PLIANT clause, which usually take their values in topologically dense sets (normally $\mathbb{R}$) and which are allowed to change continuously, such change being specified via pliant events (see below).

Next are the invariants. These resemble invariants in discrete Event-B, in that the types of the variables are asserted to be the sets from which the variables' values *at any given moment of time* are drawn. More complex invariants are similarly predicates that are required to hold *at all moments of time* during a run.

Then we get to the events. The *INITIALISATION* has a guard that synchronises time with the start of any run, while all other variables are assigned their initial values in the usual way. As hinted above, in Hybrid Event-B, there are two kinds of event: mode events and pliant events.

Mode events are direct analogues of events in discrete Event-B. They can assign all machine variables (except time itself). In the schematic *MoEv* of Fig. 1, we see three parameters $i?, l, o!$, (an input, a local parameter, and an output respectively), and a guard *grd* which can depend on all the machine variables. We also see the generic after-value assignment specified by the before-after predicate *BApred*, which can specify how the after-values of all variables (except time, inputs and locals) are to be determined. (Stochastic variables can also figure in mode events, in the same way as non-stochastic variables, though we have not mentioned them directly in *MoEv*.)

Pliant events specify the continuous evolution of the (non-stochastic, for the discussion of this section) pliant variables over an interval of time. The schematic pliant event *PliEv* of Fig. 1 shows the structure. There are two guards: there is *iv*, for specifying enabling conditions on the pliant variables, clocks, and time; and there is *grd*, for specifying enabling conditions on the mode variables. The separation between the two is motivated by considerations connected with refinement.

The body of a pliant event contains three parameters $i?, l, o!$, (once more an input, a local parameter, and an output respectively) which are functions of time, defined over the duration of the pliant event. The behaviour of the event is defined by the COMPLY and SOLVE clauses. The SOLVE clause specifies behaviour fairly directly. For example the behaviour of pliant variable y and output $o!$ is given by a direct assignment to the (time dependent) value of the expression $E(\ldots)$. Alternatively, the behaviour of pliant variable x is given by the solution of the first order ordinary differential equation (ODE) $\mathcal{D}\, x = \phi(\ldots)$, where $\mathcal{D}$ indicates differentiation with respect to time. (In fact the sematics of the $y, o! = E$ case is given in terms of the ODE $\mathcal{D}\, y, \mathcal{D}\, o! = \mathcal{D}\, E$, so that x, y and $o!$ satisfy the same regularity properties.) The remaining lines of the SOLVE clause refer to stochastic variables. We postpone discussion of these to Section III.

The COMPLY clause can be used to express any additional constraints that are required to hold during the pliant event via its before-during-and-after predicate *BDApred*. Typically, constraints on the permitted range of values for the pliant variables, and similar restrictions, can be placed here.

The COMPLY clause has another purpose. When specifying at an abstract level, we do not necessarily want to be concerned with all the details of the dynamics — it is often sufficient to require some global constraints to hold which express the needed safety properties of the system. The COMPLY clauses of the machine's pliant events can house such constraints directly, leaving it to lower level refinements to add the necessary details of the dynamics.

The semantics of (the non-stochastic part of) a Hybrid Event-B machine is as follows. It consists of a set of *system traces*, each of which is a collection of functions of time, expressing the value of each machine variable over the duration of a system run. (In the case of *HyEvBMch*, in a given system trace, there would be functions for non-stochastic variables clk, x, y, u, each defined over the duration of the run.)

Time is modelled as an interval $\mathcal{T}$ of the reals. A run starts at some initial moment of time, t_0 say, and lasts either for a finite time, or indefinitely. The duration of the run $\mathcal{T}$, breaks up into a succession of left-closed right-open subintervals: $\mathcal{T} = [t_0 \ldots t_1), [t_1 \ldots t_2), [t_2 \ldots t_3), \ldots$. The idea is that mode events (with their discontinuous updates) take place at the isolated times corresponding to the common endpoints of these subintervals t_i, and in between, the mode variables are constant and the pliant events stipulate continuous change in the pliant variables.

Although pliant variables change continuously (except perhaps at the t_i), continuity alone still allows for a wide range of mathematically pathological behaviours. To eliminate these, we make the following restrictions and recommendations which apply individually to every subinterval $[t_i \ldots t_{i+1})$:

I Zeno: there is a constant δ_{Zeno}, such that for all i needed, $t_{i+1} - t_i \geq \delta_{\mathsf{Zeno}}$. N. B. Since Zeno behaviour is a global property, its prohibition cannot be enforced statically without knowing the global reachability relation.

II Limits: for every variable x, and for every time $t \in \mathcal{T}$, the left limit $\lim_{\delta \to 0} x(t-\delta)$ written $\overrightarrow{x(t)}$ and right limit $\lim_{\delta \to 0} x(t+\delta)$, written $\overleftarrow{x(t)}$ (with $\delta > 0$) exist, and for every t, $x(t) = \overleftarrow{x(t)}$. [N. B. At the endpoint(s) of $\mathcal{T}$, any missing limit is defined to equal its counterpart.]

III Differentiability: The behaviour of every pliant variable x in the interval $[t_i \ldots t_{i+1})$ is given by the solution of a well posed initial value problem $\mathcal{D}\, xs = \phi(xs \ldots)$ (where xs is a relevant tuple of pliant variables and $\mathcal{D}$ is the time derivative). "Well posed" means that $\phi(xs \ldots)$ has Lipschitz constants which are uniformly bounded over $[t_i \ldots t_{i+1})$ bounding its variation with respect to xs, and that $\phi(xs \ldots)$ is measurable in t.

Regarding the above, the Zeno condition is certainly a sensible restriction to demand of any acceptable system, but in general, its truth or falsehood can depend on the system's full reachability relation, and is thus very frequently undecidable, so I is a recommendation rather than a condition that can be imposed statically.

The stipulation on limits, with the left limit value at a time t_i being not necessarily the same as the right limit at t_i, makes for an easy interpretation of mode events that happen at t_i. For such mode events, the before-values are interpreted as the left limit values, and the after-values are interpreted as the right limit values. In fact, the right continuity we stipulate in II places our behaviours in the space *càdlàg* (French: "continue à droite, limite à gauche"). This space is of prime interest for stochastic calculus, of which more in Section III, but for now this is just a happy coincidence, and the limits issue could quite easily be handled in different ways within the confines of the non-stochastic formalism.

The differentiability condition guarantees that from a specific starting point, t_i say, there is a maximal right open interval, specified by t_{MAX} say, such that a solution to the ODE system exists in $[t_i \ldots t_{\mathrm{MAX}})$, and such that the solution is absolutely continuous [11], [12]. Within this interval, we seek the earliest time t_{i+1} at which a mode event becomes enabled, and this time becomes the preemption point beyond which the solution to the ODE system is abandoned, and the next solution is sought after the completion of the mode event.

In this manner, assuming that the *INITIALISATION* event has achieved a suitable intial assignment to variables, a system run is *well formed*, and thus belongs to the semantics of the machine, provided that at runtime:

- Every enabled mode transition is feasible, i.e. has an after-state, and on its completion enables a pliant transition (but does not enable any mode transition).
- Every enabled pliant transition is feasible, i.e. has a time-indexed family of after-states, and EITHER:
  - (i) During the run of the pliant transition a mode transition becomes enabled. It preempts the pliant transition, defining its end. ORELSE
  - (ii) During the run of the pliant transition it becomes infeasible: finite termination. ORELSE
  - (iii) The pliant transition continues indefinitely: nontermination.

Thus in a well formed run mode events alternate with pliant events. The last event (if there is one) is a pliant event (whose duration may be finite or infinite).

N. B. Many formalisms for hybrid systems permit a succession of mode events to execute before the next pliant event runs (to use our terminology). We avoid this for a number of reasons. Firstly, it spoils the simple picture that at each time, each variable has a unique value, and the runtime semantics of a variable is a straightforward function of time. Secondly, it avoids having to define the final value of a succession of mode events via a fixpoint calculation. Thirdly, and perhaps most importantly, it maintains the discrete Event-B picture in which events are (implicilty) seen as taking place at isolated points of real time, insofar as Event-B behaviours are seen as relating to the real world. We regard the overturning of such unstated assumptions as particularly dangerous in an engineering context — c.f. the Mars Lander incident, in which the U.S. and European teams interpreted measurements according to different units, without anyone ever thinking to check which units were actually intended.

III. Adding Stochastic Behaviour to Hybrid Event-B

Adding a stochastic dimension to the conventional discrete transition systems used for the majority of computing formalisms and applications has been intensively studied over many years. From the large literature we just cite the popular PRISM tool [13], and monograph [14]. Increasingly, other formalisms are also adding probabilistic aspects to their functionality, e.g. [15].

As noted in the Introduction, in this paper, we do not examine this aspect in our extension of Hybrid Event-B, since our primary focus is on applications where the mode

transitions happen to be non-stochastic.[1] Rather, we restrict our investigation to the inclusion of noise in the continuous behaviours interleaving between individual discrete transitions. This is the realm of stochastic calculus.

Stochastic calculus also has a long pedigree and a considerable literature. Once a subject that was confined to mathematical probabilists and theoretical physicists, the discipline underwent an explosion of interest when a stochastic calculus approach led to the discovery of the Black-Scholes equation. This equation, which won its discoverers the Nobel Prize for Economics in 1997, proved to be the key to a principled way to price financial derivatives.[2] The huge importance of the financial industry did the rest.

The enormous interest in stochastic calculus continues to this day. Now, there is a bewildering variety of books on stochastic calculus, and many earlier treatments of probability have been expanded to include the topic in order to tap into the huge popularity that stochastic calculus currently enjoys, especially when it is slanted towards financial applications. From the voluminous literature, we cite only [16], [17], [18], [19], [20].

A. Some Stochastic Calculus Essentials

We start on familiar territory. Using conventional mathematical notation, a first order ordinary differential equation can be written in the following form:

$$\frac{dX}{dt} = \phi(t, X) \tag{1}$$

where $\frac{dX}{dt}$ is the conventional derivative of X (with respect to, as it happens, time), and ϕ is well enough behaved. Then, the usual global results follow (e.g. uniqueness of the solution, and its continuity (also with respect to parameter variation)). Now, we want to amplify this picture to include a noise term that can affect the solution:

$$\frac{dX}{dt} = \phi(t, X) + \text{“}noise\text{”} \tag{2}$$

The question now arises, how to interpret the noise term. To cut a long story short, an interpretation of noise as mathematically formal white noise is possible. Unfortunately, mathematical white noise is discontinuous and of unbounded variation everywhere. The prospects of doing relatively normal calculations with such an object disappear. If you sweat the mathematics hard enough, the white noise idea can be made to work, but the white noise process has to be constructed as a probability measure on the space of tempered distributions. This is conceptually very exotic, even by the extremely generous standards of conventional stochastic calculus.

Instead, the usual approach is to proceed as follows. We observe that, in general, when we integrate some ‘function’, it becomes less irregular — typically, a discontinuous function becomes continuous, a continuous function becomes differentiable, etc. If we thus consider the ‘integral’ of the white noise process, we are able to put aside the distributional machinery, and confine attention to more conventional probabilistic techniques. The formulation (2) is transmuted to:

$$dX = \phi(t, X)dt + \sigma(t, X)dW \tag{3}$$

In (3), which is really a shorthand for an easier to define integral equation relating a finite increment of X to a finite increment of ϕ combined with a finite increment of noise,

$$X(t_H) - X(t_L) = \int_{t_L}^{t_H} \phi(t, X)dt + \int_{t_L}^{t_H} \sigma(t, X)dW \tag{4}$$

W is a standard Wiener process (also known as Brownian motion), a mathematically tractable ‘integral’ of the formal white noise mentioned above.

The Wiener process is of unbounded variation and nondifferentiable (almost) everywhere, but may, with some care, be constructed to be continuous. This encourages the possibility of integrating it, as (4) would suggest.

Unfortunately, defining the integral of W (and especially of σW) brings surprises. Although W is continuous, when σ depends nontrivially on W (as (4) suggests is permissible), W varies so violently that the choice of sample point t_{θ_j} in a Riemann-Stiltjes approximation to the integral, $\sum_{t_L}^{t_H} \sigma(t_{\theta_j}, X(t_{\theta_j}))W(t_{j+1} - W(t_j)$, where $t_j \leq t_{\theta_j} \leq t_{j+1}$, makes a macroscopic difference to the answer as the limit $(t_{j+1} - t_j) \to 0$ is taken. The leftmost choice, $t_{\theta_j} = t_j$, yields the Itô integral, while the midpoint choice, $t_{\theta_j} = (t_j + t_{j+1})/2$, yields the Stratonovich integral, these being the two most commonly used, out of an infinity of possible (and different) definitions.

For the stochastic differential equations that thus arise, with mild constraints on ϕ and σ, existence and uniqueness of solutions can be proved in a suitable sense. In fact the same conditions we imposed in Section II will do, i.e. measurability in time and the Lipschitz property in state variables. The only difference is that the solutions that are thus guaranteed are merely continuous in time (instead of being absolutely continuous). This is due to the violence of the fluctuations in W, and corresponds to the fact that despite being ‘integrals of (3)’ these solutions are not, in any conventional sense, differentiable anywhere. In fact the approximation process by which the solutions are constructed, converges only in the L^2 sense, rather different to the pointwise convergence for normal differential equations.

In the context of normal differential equations, the general existence and uniqueness theorems that apply, promise far more solutions than can ever be calculated by analytical techniques. In the same manner, the general existence and uniqueness theorems for stochastic differential equations also promise immeasurably more than calculation can deliver. Ironically, this simplifies the task for automated approaches based on calculation, since there is simply a lot less that can be calculated. (Correspondingly, there is greater emphasis

[1] This is just for convenience in this study. There would be no reason to insist on it in principle.

[2] Actually, only Scholes and Merton (who also worked on the problem) got the Nobel Prize. Black died in 1995, so was ineligible to receive it.

on simulation based approaches to stochastic calculus, the techniques for which have been highly developed.)

Mostly, the stochastic differential equations that yield to analytic calculation are similar to restricted cases of analytically solvable ordinary differential equations. For instance, a class of scalar equations in linear form can be solved by a variation of parameters technique. The possibilities become even more restricted when scalars become vectors, and as multiple sources of noise are allowed.

The range of possibilities for closed form solutions is enlarged somewhat by permitting changes of variables. If we know a solution $X(t)$ to (3), and we have a function:

$$Y(t) = g(t, X(t)) \tag{5}$$

then Y is a solution of:

$$dY = \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dX + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX)^2 \tag{6}$$

where $(dX)^2$ is calculated by squaring (3), and then simplifying using the rules:

$$(dt)^2 = (dt)(dW) = (dW)(dt) = 0 \tag{7}$$

$$(dW)^2 = dt \tag{8}$$

provided the various derivatives of g exist in the usual sense. The system (6)-(8) constitutes what is known as the Itô formula, and is applicable when stochastic integrals are understood in the Itô sense. When stochastic integrals are understood in the Stratonovich sense (i.e. a Stratonovich interpretation is imposed on the integral in (4)), the rather disquietening second order term in (6) is removed and we get a more normal chain rule. We will see a small example of these things in our case study below.

It might appear improbable that a framework constructed via mathematical contortions as extreme as those just described should have any relevance to applications in the real world. The fact is, however, that it does. The principal reason for this is that the idealised stochastic processes that the mathematics creates have independent increments on disjoint time intervals. And whereas such exotic properties as nondifferentiability everywhere have little relevance to physical applications, in practice such details make little difference on macroscopic timescales. By contrast, the independence on disjoint time intervals is a good reflection of the properties of physical noise, dramatically enhancing the utility of these models.

B. Stochastic Hybrid Event-B

The discussion of Section III-A can be restated in the following manner. While we remain in the realm of normal differential equations, all of whose ingredients are continuous in all variables, the terms appearing in such a differential equation stand for themselves. In other words, they represent directly (i.e. are merely notations for) elements of the mathematical semantic domain; and the equality between the left and right hand sides of such a differential equation denotes the literal identity of the same real function represented by the notations appearing on either side of it.

However, when we contemplate extending the idea of a differential equation to increasingly irregular behaviours, then the procedures by which we give rigorous meaning to what we intuitively wish to understand by the notational elements of a differential equation become increasingly complex and unituitive. In other words, the differential equation itself becomes a *syntax*, for which the procedures referred to give the semantics.[3]

As with almost all nontrivial constructions in real analysis, the procedures that we are referring to, define the desired semantics in the sense of Platonic mathematics — i.e. nonconstructive arguments are fully utilised when needed in the construction of the desired limits. And as we noted, the proportion of instances of such arguments that can be instantiated in symbolic calculations is rather small.

Again referring to conventional mathematical discourse, much is inferred from the context, and the notations for differential equations are modified as much or as little as is necessary to make the discourse as a whole unabiguous. For instance, when both Itô and Stratonovich integrals are mentioned in the discourse, the notation $\int_{t_L}^{t_H} \sigma(t, X) \circ dW$ is often used for the latter to distinguish them.

In the case of a formal notation like the extension of Hybrid Event-B that we are contemplating, we need to be precise enough about such notational matters that we become sure that the underlying formal system implementing the semantics (in a tool, for example) does the right thing. Moreover, when we have a situation in which there is a possible choice in the matter of semantics, and that that choice is of relevance to users of the formal notation (such as in the case just mentioned of the choice between Itô and Stratonovich semantics for SDEs), then the notation must be rich enough that the desired semantics can be indicated unabiguously through the syntax.

We return to the schematic syntax of Fig. 1 to fill in the discussion of the elements passed over in Section II. The first item omitted from Section II is the 'STOCHASTIC z_I, z_S' declaration. This declares z_I and z_S to be stochastic variables, with sample space constructed from $\mathbb{R}$. The latter point precludes the need to declare their type in the INVARIANTS clauses.

The next item of omitted syntax occurs in the declaration 'WHITE W'. This is a reference to Wiener noise, the distribution (essentially a time parameterised Gaussian) that independent increments of standardised Wiener noise are assumed to obey. The fact that we say 'WHITE' instead of 'WIENER' will be explained shortly. While we are discussing this, it is as well to mention that Wiener noise is not the only kind that is considered in stochastic calculus. Although we confine our treatment here to the Wiener case, there are also other distributions that are used in appropriate circumstances, such as the Cauchy distributions and the Lévy distributions.

[3] As an example, when we loosen the requirement on the right hand side of an ordinary differential equation to mere *measurability* in time (rather than actual continuity) –as we do in Section II– the solutions generated are guaranteed to have a derivative only *almost everywhere*; i.e. the equality between the left and right hand sides of the differential equation holds only up to a set of measure zero, and on this set of measure zero the left hand side of the differential equation is meaningless.

(These share, along with the Gaussian distribution, needed self-similarity properties.) For extensions of Hybrid Event-B beyond what we consider here, we could introduce additional declarations 'CAUCHY *C*' or 'LEVY *L*' to name them.

The next items of omitted syntax occur in the INVARIANTS clauses. The declaration $\mathcal{P}[\psi_I(z_I)] \leq \ldots$ states that the probability of $\psi_I(z_I)$ is always less than some unspecified expression. Likewise $\mathcal{E}[\psi_S(z_S)] \leq \ldots$ states that the expectation value of $\psi_S(z_S)$ is always less than some similarly unspecified expression. The symbols $\mathcal{P}$ and $\mathcal{E}$ economise on the need to write out the definitions that specify these quantities in detail. Further symbols denoting other probabilistic quantities could be introduced in a similar fashion if needed. Note that it is unlikely that a stochastic variable like z_I or z_S will occur 'raw' (i.e. outside the context of some statistic such as an expectation or a probability estimate), in an invariant. If the sample space of a stochastic variable were bounded, then a property like $z \leq BOUND$ would be provable (rather trivially, by definition). However the Wiener noise we focus on is unbounded, so useful finite properties only arise through the use of summative statistics such as expectations.

The next omitted syntactic items occur in the *INITIALISATION* event, after the initialisations of the non-stochastic variables. The line '$z_I := \mu[z_{I,0}]$' indicates that the initial distribution of the stochastic variable z_I is a single point mass distribution centred at $z_{I,0}$. The line '$z_S := U[0 \ldots 1]$' indicates that the stochastic variable z_S is initially distributed uniformly over the interval $[0 \ldots 1]$. names for other common distributions may be introduced in a similar way. Also, more complex expressions for intial distributions can be written using lambda notation, for instance: $(\lambda\, s \bullet Prob(s))$ where s ranges over the sample space (which is always $\mathbb{R}$).

The next pieces of omitted syntax occur in the SOLVE clause of the *PliEv* event. After the lines that define the behaviour of the deterministic variables, $x, y, o!$, there are two lines that stipulate how the stochastic variables z_I and z_S evolve. To save clutter, we eliminate the differentials from a representation like (3), giving a syntax closer to (2), although the content is always treated as though it was like (3).

Thus the line '$\mathcal{I}\, z_I = \phi_I(z_I \ldots) + \sigma_I(z_I \ldots)W$' defines the behaviour of z_I via an Itô stochastic differential equation. So: '$\mathcal{I}$' indicates Itô; 'ϕ_I' (i.e. the terms not involving W multiplicatively) are the deterministic (or drift) components of the equation; and '$\sigma_I W$' (i.e. the terms involving W multiplicatively) denote the noise terms. The lack of differentials in the notation make the equation resemble (2), so it is possibly less confusing syntactically to refer to W as white noise via the earlier declaration 'WHITE'. However, since W is only a name (i.e. it is just an individual symbol labelling the noise terms), there is no barrier to *interpreting* the information in the equation '$\mathcal{I}\, z_I = \ldots$' as though it were in the standard differential form (3), in which case reading W as 'Wiener' would make sense. Having the same initial letter for both names is convenient.

An important point to note is that deterministic equations, like $\mathcal{D}\, x = \ldots$ must not feature raw stochastic variables. Any expression featuring a stochastic variable is itself a random variable, with a distribution over the same sample space. However, statistics like expectations remove this variability (even if expectations are viewed as (constant) random variables in probability theory), so the inclusion of terms like $\mathcal{E}[z_I]$ in the right hand side of $\mathcal{D}\, x = \ldots$ would be permissible.

A related observation is that the *BDAPred* predicate that can occur in a pliant event like *PliEv* would be subject to the same restrictions regarding stochastic variables as arose in the context of invariants, so that raw stochastic variables would not be expected to occur. Unlike invariants though, the contents of *BDAPred* could contain time-varying properties.

Turning to the line '$\mathcal{S}\, z_S = \phi_S(z_S \ldots) + \sigma_S(z_S \ldots)W$', it is analogous to the Itô case. It simply stipulates, via the symbol '$\mathcal{S}$', that the stochastic differential equation is to be interpreted according to Stratonovich semantics. Otherwise, things are as previously.

As noted before, there are also other cases of stochastic differential equations which are not covered in the preceding account. However, they will require syntactic devices very similar to those we have described, so their omission here is not serious from the point of view of designing the overall shape of the extended Hybrid Event-B formalism.

IV. A Simple Case Study

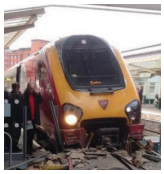

Fig. 2. A recent buffer collision in the UK.

In this section we present a simple case study to illustrate the preceding techniques. Our case study is based on a toy train stopping application. Adhesion of the train to the track is a serious problem for the management of railways. Automated systems for calculating the train's position from onboard instrumentation can be thrown awry by slippage between the wheels and the track. Also, trains approach station platforms extremely slowly, in order that uncertainties in the behaviour of the braking system do not cause the train to overshooot, especially if there is a risk of hitting a buffer. (See Fig. 2 for an extreme case of overrun, contributed to by brake failure.)

Here, we model the uncertainty of braking by adding a noise term to the desired ideal deceleration equation of motion. The ideal, deterministic braking law is thus:

$$dV = -a\, dt \tag{9}$$

in which V is the current velocity, and a is the (constant) deceleration. The braking law has been written in differential form, ready for what is to follow.

We add a noise term to get:

$$dV = -a\, dt + \sigma V dW \tag{10}$$

In (10) the deterministic part is still $-a$, but there is also the noise term $\sigma V dW$. The noise term is multiplicative (i.e the noise term dW is multiplied by the dependent variable V),

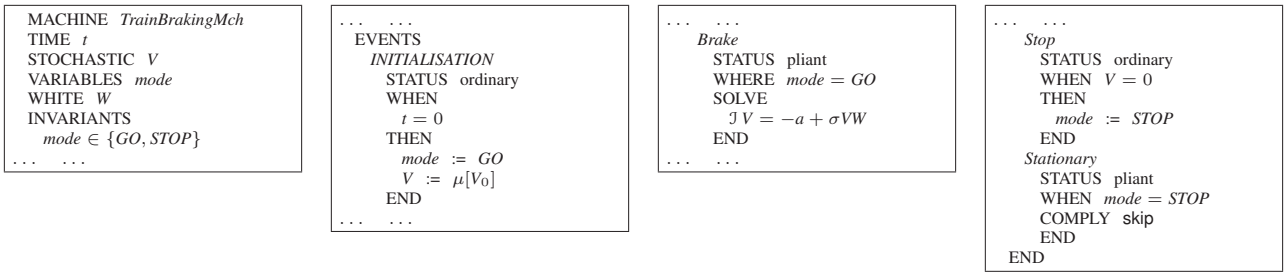

Fig. 3. A Stochastic Hybrid Event-B machine for simple train braking.

since we would not expect the train velocity to be subject to braking uncertainty once it had stopped.

Fig. 3 shows a complete Stochastic Hybrid Event-B machine for depicting this situation. The modelling starts at the moment the braking is about to start, with time synchronised to 0, and the velocity V presumed to be known exactly, thus initialised to a point mass distribution centred on the starting velocity V_0. The *INITIALISATION* event enables the braking pliant event *Brake*, via the setting of the *mode* variable to *GO*. During *Brake*, the velocity is subject to the stochastic law (10), written in the formal Stochastic Hybrid Event-B notation as $\Im V = -a + \sigma VW$, indicating moreover, that we intend to interpret the noise in the Itô sense.

Once the noisy braking process has reduced the velocity to zero, the *Stop* event fires, sets the *mode* variable to *STOP*, and thus disables *Brake*. The new value of *mode* enables the *Stationary* pliant event, which, in simply specifying COMPLY skip, affirms that the dynamics has terminated and that the system has reached its final state.

We turn to the solution of this system. One way of approaching this is to use variation of parameters, mentioned earlier. For this, we can regard (10) as a homogeneous linear equation, $dV = \sigma V dW$, to which an inhomogeneous term, $-a\,dt$ has been added. The homogeneous equation has the solution:

$$V^0(t) = e^{-\frac{1}{2}\sigma^2 t + \sigma W(t)} \tag{11}$$

This may be verified as follows. First, we notice that the differential of the exponent is $-\frac{1}{2}\sigma^2 dt + \sigma dW(t)$. Second, we notice that when we apply Itô's formula (6)-(8) to $g(t,x) \equiv e^x$, replacing its argument x with the exponent in (11), then the derivative comes out as:

$$\begin{aligned} dV^0 &= \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dV^0 + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dV^0)^2 \\ &= 0 + e^{[\ldots]}[-\frac{1}{2}\sigma^2 dt + \sigma dW(t)] + \\ &\quad \frac{1}{2}e^{[\ldots]}[-\frac{1}{2}\sigma^2 dt + \sigma dW(t)]^2 \\ &= e^{[\ldots]}[-\frac{1}{2}\sigma^2 dt + \sigma dW(t) + \frac{1}{2}(0 + 0 + \sigma^2 dt)] \\ &= \sigma V^0(t) dW(t) \end{aligned} \tag{12}$$

where, in the penultimate line of (12), we have used Itô's formula to eliminate negligible higher order differentials, and replace $(dW)^2$ with dt, to get the result desired.

To solve (10) fully, we now set $V(t) \equiv V^0(t)Z(t)$ and follow the variation of parameters recipe. In general, we would need to consider that $Z(t)$ is itself a stochastic process, and act accordingly, but it turns out that in this case all the stochastic behaviour is already catered for in $V^0(t)$, so $Z(t)$ turns out to be deterministic. Thus:

$$\begin{aligned} dV &= V^0 dZ + Z dV^0 \\ &= e^{[\ldots]} dZ + Z\sigma V^0(t) dW \\ &= -a dt + \sigma V^0(t) Z dW \end{aligned} \tag{13}$$

so that

$$e^{[\ldots]} dZ = -a\,dt \tag{14}$$

At this point, separation of variables can be applied to yield the complete solution:

$$V(t) = e^{-\frac{1}{2}\sigma^2 t + \sigma W(t)} \left\{ V_0 - a \int_0^t e^{(\frac{1}{2}\sigma^2 s - \sigma W(s))} ds \right\} \tag{15}$$

The preceding illustrates some of the black magic of stochastic calculus. Usually, the trickiest aspect is knowing when one can safely apply 'normal calculus rules', and when it is unsafe to do so. For instance, (15) cannot be reduced further, since it contains a sample path $W(t)$ of Wiener noise, in a way that cannot be simplified any more.

In many physical applications, the Stratonovich integral provides a more convincing model of the underlying process, since it can be derived as a limit of smooth approximations to Wiener noise, a tactic which seems eminently plausible in the context of modelling physical processes. Taking this approach above would eliminate all the $\frac{1}{2}\sigma^2$ terms.

Since (15) depends on an explicit Wiener sample path, it is not easy to deduce anything specific from it. For instance, one cannot display how its value varies with the sample path since the space of sample paths is infinite dimensional. Reducing the detail considered offers more promise of progress. Thus, attempting to calculate statistics like the mean or variance of the distribution gives a better idea of what (15) actually means.

Working out statistics such as the mean and variance of (15) over Wiener paths $W(t)$ is technically nontrivial. Of even more interest though, are random variables that quantify the time taken for the velocity to drop to zero, i.e. for the train to stop, and beyond that, the time integral of $V(t)$ over this period. The former (without intending any irony) is a 'stopping time'

in the terminology of stochastic processes. Calculating such quantities is technically even more demanding than merely calculating means, and often, a safe estimate based on (say) twice the variance will be adequate in practice. For lack of space, we defer the pursuit of these technicalities to another place.

Nevertheless, we point out that once some route to estimating quantities like the stopping time has been determined, then the relevant quantities can be reformulated as invariants of the system model. In this way, the calculational excursions involved in pursuing them can be brought back to connect with familiar notions, namely invariants in Hybrid Event-B. In the case of our toy model, such an invariant might state that 'with probability X, the distance travelled during the braking episode is less than Y'.

V. Conclusions

In the preceding sections we have reviewed conventional Hybrid Event-B, and we then embarked on a brief excursion into some of the salient features of stochastic calculus, in order to inform the design of a suitable extension of Hybrid Event-B that would include modelling capabilities based on it. In the event, the modification to Hybrid Event-B proved to be fairly mild, due in part to the prior design of the semantics of Hybrid Event-B having been done in a way that made the extension straightforward. Still, relatively straightforward though the extension was, it nevertheless necessitated a significant review of stochastic calculus issues in order to justify that the proposed extension should indeed be as it was described.

Related to the impact on the Hybrid Event-B formalism, is the impact on any tool that supports it. Here, we can again argue that the impact will be mild. The reason for this is that, as we argued above, those cases of stochastic differential equations that can be addressed using symbolic means, can be solved by deploying the same family of techniques that support the continuous portion of conventional Hybrid Event-B. It merely needs to be signaled to an implementation that a particular equation in the SOLVE clause is a stochastic differential equation, with the noise terms it contains suitably declared beforehand, and the implementation will be able to organise the appropriate calculations. In this manner we hope to be able to include support for straightforward stochastic calculus early in the development of tool support for Hybrid Event-B. We illustrated our formalism with a simple case study based on the uncertainties of a toy train stopping application.

One thing we deliberately omitted from the treatment of this paper is the more familiar style of probabilistic reasoning associated with discrete events. Although we were concerned with deterministic external control, via the mode events associated with predictable behaviour, the more stochastic aspect certainly comes to the fore when we consider failure modes and the overall dependability of systems composed of not wholly reliable components. Such an integrated treatment requires a deeper excursion into stochastic calculus, engaging with the more complex Lévy processes. This is another aspect that we defer for investigation elsewhere.

References

[1] J. Sztipanovits, "Model Integration and Cyber Physical Systems: A Semantics Perspective," in *Proc. FM-11*, Butler and Schulte, Eds. Springer, LNCS 6664, p.1, http://sites.lero.ie/download.aspx?f=Sztipanovits-Keynote.pdf, 2011, Invited talk, FM 2011, Limerick, Ireland.

[2] J. Willems, "Open Dynamical Systems: Their Aims and their Origins. Ruberti Lecture, Rome," 2007, http://homes.esat.kuleuven.be/~jwillems/Lectures/2007/Rubertilecture.pdf.

[3] "Summit Report: Cyber-Physical Systems," 2008, http://iccps2012.cse.wustl.edu/_doc/CPS_Summit_Report.pdf.

[4] L. Barolli, M. Takizawa, and F. Hussain, "Special Issue on Emerging Trends in Cyber-Physical Systems," *J. Amb. Intel. Hum. Comp.*, vol. 2, pp. 249–250, 2011.

[5] J.-R. Abrial, *The B-Book: Assigning Programs to Meanings*. Cambridge University Press, 1996.

[6] ——, *Modeling in Event-B: System and Software Engineering*. Cambridge University Press, 2010.

[7] R. Banach, M. Butler, S. Qin, N. Verma, and H. Zhu, "Core Hybrid Event-B: Adding Continuous Behaviour to Event-B," 2012, submitted.

[8] R. Banach and M. Butler, "Cruise Control in Hybrid Event-B," in *Proc. ICTAC-13*, ser. LNCS, Liu, Woodcock, and Zhu, Eds., vol. 8049. Springer, 2013, pp. 76–93.

[9] ——, "A Hybrid Event-B Study of Lane Centering," in *Proc. CSDM-13*, Aiguier, Boulanger, Krob, and Marchal, Eds. Springer, 2013, pp. 97–111.

[10] R. Banach, "Pliant Modalities in Hybrid Event-B," in *Proc. Jifeng He Festschrift 2013*, ser. LNCS, Liu, Woodcock, and Zhu, Eds., vol. 8051. Springer, 2013, pp. 37–53.

[11] Wikipedia, "Absolute continuity."

[12] H. Royden and P. Fitzpatrick, *Real Analysis*. Pearson, 2010.

[13] PRISM Probabilistic Model Checker, http://www.prismmodelchecker.org.

[14] A. McIver and C. Morgan, *Abstraction, Refinement and Proof for Probabilistic Systems*. Springer, 2004.

[15] UPPAAL, http://www.uppaal.org.

[16] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2010, 5th ed.

[17] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford University Press, 2001, 3rd ed.

[18] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Springer, 1992, 4th ed.

[19] M. Jeanblanc, M. Yor, and M. Chesney, *Mathematical Methods for Financial Markets*. Springer, 2009.

[20] A. Kyprianou, *Fluctuations of Lévy Processes with Applications*. Springer, 2014.