Pergamon

PII: S0016-0032(96)00126-3

J. Franklin Inst. Vol. 335B, No. 3, pp. 401-439, 1998

Published by Elsevier Science Ltd
Printed in Great Britain
0016-0032/98 $19.00+0.00

A Theory of Discontinuities in Physical System Models

by PIETER J. MOSTERMAN

Department of Electrical and Computer Engineering, Box 1824, Sta B, Vanderbilt University, Nashville, TN 37235, U.S.A.

and GAUTAM BISWAS

Department of Computer Science, Box 1679, Sta B, Vanderbilt University, Nashville, TN 37235, U.S.A.

(Received 22 August 1996)

ABSTRACT: *Physical systems are by nature continuous, but often display nonlinear behaviors that make them hard to analyze. Typically, these nonlinearities occur at a time scale that is much smaller than the time scale at which gross system behavior needs to be described. In other situations, nonlinear effects are small and of a parasitic nature. To achieve efficiency and clarity in building complex system models, and to reduce computational complexity in the analysis of system behavior, modelers often abstract away any parasitic component parameter effects, and analyze the system at more abstract time scales. However, these abstractions often introduce abrupt, instantaneous changes in system behavior. To accommodate mixed continuous and discrete behavior, this paper develops a* hybrid *modeling formalism that dynamically constructs bond graph model fragments that govern system behavior during continuous operation. When threshold values are crossed, a meta-level control model invokes discontinuous state and model configuration changes. Discontinuities violate physical principles of* conservation of energy *and* continuity of power, *but the principle of* invariance of state *governs model behavior when the control module is active. Conservation of energy and continuity of power again govern behavior generation as soon as a new model configuration is established. This allows for maximally constrained continuous model fragments. The two primary contributions of this paper are an algorithm for inferring the correct new mode and state variable values in the hybrid modeling framework, and a verification scheme that ensures hybrid models conform to physical system principles based on the principles of* divergence of time *and* temporal evolution in behavior transitions. *These principles are employed in energy phase space analysis to verify physical consistency of models.* © 1997 The Franklin Institute. Published by Elsevier Science Ltd

1. Introduction

Physical systems are by nature continuous, but their behaviors often exhibit complex nonlinearities which obscure system and model characteristics. Nonlinear behaviors are hard to analyze. Behavior generation algorithms may have to deal with steep gradients that introduce numerical stiffness and hamper meaningful analysis and interpretation. Furthermore, since system linearity cannot be assumed, a number of well studied powerful linear analysis methods in the quantitative and qualitative domains (**1**, **2**) cannot be applied to analyze system behavior.

Typically nonlinearities in system behavior are simplified by abstraction. A complex nonlinearity may be represented as a set of piecewise linear behaviors, or the time scale may be modified (granularity is decreased), and, as a result, certain behaviors seem to undergo discontinuous changes. To systematically model abstractions and the ensuing discontinuities in behaviors, the notion of mixed continuous/discrete *hybrid* models of physical systems is developed in this paper. The continuous behavior regions are modeled using energetic bond graph models (**3**, **4**) from which the traditional differential state equations can be derived. At points of discontinuity the structural topology of the bond graph model may change, causing a switch from one system configuration to another, which results in a corresponding change in the state equations that govern continuous behavior. The discontinuous transition from one model (system configuration) to another provides interesting challenges:

1. How can we guarantee that the model topology in the new state corresponds to the correct system configuration?
2. How do we guarantee that the initial state derived for the new configuration is correct, i.e. it produces behavior that corresponds to what the real physical system would generate?

The model transition process is best represented as a discrete switching process, and this process can be quite elaborate (**5**, **6**). This is particularly true if the system under consideration traverses several intermediate discontinuous states before converging on a new continuous mode of operation.

It has been established that the bond graph modeling scheme (**3**, **4**) provides a general methodology for describing and analyzing the continuous characteristics of physical systems. It provides an elegant domain-independent formalism for building lumped parameter physical system models from a small set of primitive components, and embodies the principle that dynamic behavior of a system is governed by the energy exchange patterns among the components of the system. Overall system behavior generation is governed by the principle of conservation of energy, and more specifically by the first and second law of thermodynamics (**7**, **8**).

To allow for graceful introduction of discontinuities through model configuration changes in the bond graph framework, new paradigms have to be introduced. In recent years, this has been an active area of research (**9**–**14**). Promising approaches that have emerged include the use of

- modulated transformers with boolean modulus (**13**),
- the introduction of a new bond graph primitive to model an ideal switch (**12**), and
- controlled junctions that turn *on* and *off* to switch model configurations (**14**).

A primary goal of this research is to formulate a sound theory of hybrid models from a physical perspective. Logical approaches to hybrid modeling have led to the use of *dense time* finite state automata (**15**–**17**), but they lack an underlying physical theory, which results in this methodology producing descriptive modeling languages that result in the generation of physically underconstrained models. Some attempts to prove the physical correctness of these models have been successful under limited conditions, such as the restriction that state variables can only change at a constant rate. This severely limits the applicability of such models to real physical systems. In other work, Iwasaki *et al.* **(18)** focus on the behavior generation process for hybrid models and make no attempt to verify their physical consistency. They introduce the notion of *hypertime*, an infinitesimal interval during behavior generation when actual time does not advance. This facilitates changes in a model state or configuration through a series of discontinuous changes happening in infinitely short time intervals, and produces executable hybrid models, but it may result in incorrect transfer of system state between model configurations. Moreover, behavior generation is still possible for physically inconsistent models which renders these models valid only by merit of their execution semantics.

An important contribution of this work is the establishment of a theory of discontinuities in physical system models. This theory has its basis in dynamic physical system theory and the laws of thermodynamics. The bond graph framework extended by controlled junctions defines the hybrid system modeling language. The hybrid modeling language allows for the expression and analysis of different kinds of nonlinearities as piecewise linear models. Also, the hybrid scheme defines an algorithm for model switching and deriving system state that can be verified† for physical consistency.

The hybrid modeling formalism introduces controlled junction elements into bond graph models. These junctions, controlled by finite state automata, can be switched *on* and *off*. When *on*, their active state of operation, they behave like normal junctions. When *off*, their inactive state, they inhibit energy transfer. This hybrid formalism embodies a uniform approach to analyzing mixed continuous/discrete system behavior without violating fundamental physical principles, such as conservation of energy and momentum. Establishing a sound theory of hybrid models requires the development of a theory of discontinuities, and the effects of interleaving discontinuities with continuous behavior changes. Our basic premise is that system state cannot change during a sequence of discrete switches because they do not have a real manifestation. Energy redistribution among system components occurs only in the continuous modes of system operation.

The rest of the paper is organized as follows. Section II reviews the basic characteristics of physical systems and their interpretation in the bond graph framework. Section III establishes the nature of discontinuities, their links to abstraction of physical system models, and their effects on overall behavior generation and analysis. Section IV focuses on the hybrid modeling scheme, and the Mythical Mode Algorithm (MMA) for establishing discontinuous changes in behavior generation. Section V introduces the notions of physical consistency and the verification of hybrid system models based on physical theory. The verification methodology, energy phase space analysis, is applied to a number of systems to demonstrate its effectiveness. Section VI discusses

† At this point, it may be useful to emphasize the difference between model verification and model validation. To verify a physical system model involves checking whether it violates any physical laws and constraints, i.e. it conforms with theory. Validation establishes how well the model generates behaviors that correspond to the exact, real situation of interest.

the implications of the hybrid modeling theory, and makes suggestions for future work in this area.

II. Characteristics of Physical Systems

In this section, we discuss primary characteristics of physical systems, and the bond graph modeling formalism, which, based on lumped parameter energetic modeling, conforms to these characteristics.

2.1. *Energetic modeling*

Physical systems theory is based on the concept of *reticulation* which assumes that certain properties of a system can be isolated and lumped into processes with well-defined parameter values, and the system can be defined as a network of interacting processes (7). For the *lumped parameter* assumption to hold the resulting models have to be *persistent*, i.e. the model corresponding to the physical system cannot change abruptly. Therefore, abrupt dynamic effects, such as the introduction of turbulence due to sudden large pressure differences in a tank, cannot be analyzed with such models. This is especially important when instantaneous configuration changes occur, and the lumped parameter assumption requires immediate redistribution of energy to establish homogeneity within each buffer element in the new mode of operation. In a mode of operation where the model evolves continuously, homogeneity within individual buffers is easily maintained. However, instantaneous configuration changes violate this condition, and only after a new continuous mode of operation is attained, system stability is restored and redistribution of energy resumes based on principles that govern lumped parameter modeling.

Bond graphs, based on energetic modeling of physical systems, adopt the lumped parameter approach to modeling and describe a physical system at any given time as a distribution of energy over connected physical elements. This energy distribution reflects the history of the system, and, therefore, defines its *state*. Future behavior is determined by its current state description, and subsequent *input* to the system. Changes in state of a physical system are attributed to energy exchange among its components which can be expressed in terms of the time derivative or flow of energy, *power*. Irrespective of domain (i.e. mechanical, fluid, pneumatic, electric, etc.) power is the product of two conjugate variables: the *intensive* variable or effort, e, and the *extensive* variable or flow, f. Therefore, effort and flow are called *power* or *signal variables*. Intensive variables are specified at points in a system, and may vary from point to point (e.g. pressure, temperature). Extensive variables on the other hand, are defined over an extent (e.g. volume, charge), and are typically additive in nature.† For example, if one considers two blocks of the same material at the same temperature, and brings them together to form one system, the volume of the overall system is the sum of the individual volumes. On the other hand, the temperature of the combined system remains the same.

† Though variables of an additive nature are extensive, not all extensive variables are necessarily additive in nature. This is particularly true in the case of fields.

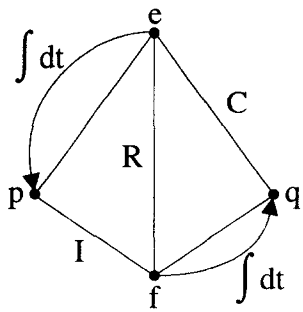


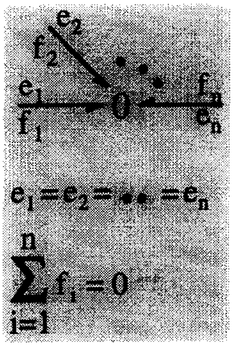
FIG. 1. The tetrahedron of state.

2.2. *Primitive elements in bond graph models*

Energy can be represented as stored effort and stored flow. The energy corresponding to stored effort is called *generalized momentum*, p, and the energy corresponding to the stored flow is *generalized displacement*, q. Consequently, p and q are called *energy variables*, and constituent elements that store generalized momentum and displacement in the bond graph framework are called inductors, I, and capacitors, C, respectively. These ideal energy storage element relations are shown by the *tetrahedron of state* (**19**) in Fig. 1, and represent the reversible processes in nature. Because of their *integrating* nature, the actual energy stored in these elements is a function of the initial value of the energy they contain. Each initial value, therefore, introduces a *degree of freedom* in the system and an additional dimension to the state vector. If the integral form cannot be used, the stored energy value in the buffer is completely determined by the other components in the system model, and the element does not introduce an additional degree of freedom to the system.† Irreversible processes are represented by the dissipative element, R, which generates entropy. In an isothermal environment, this flow of energy to the thermal domain is not shown explicitly.

The R, C, and I elements exchange energy with other elements via ports. To connect more than two basic elements together, a *junction* structure is required. Junctions typically allow an arbitrary number of components to be connected together. They preserve continuity of power by adhering to the generalized forms of Kirchoff's current and voltage laws, which define the two forms of junctions, 0- and 1-junctions, respectively. Figure 2 illustrates the constituent equations of the two types of junctions. Junction relations are instantaneous, i.e. they do not introduce temporal effects. Two special types of junctions, or signal transformers: the transformer, TF, and the gyrator, GY, complete the basic elements in the bond graph language. The transformer establishes a ratio between input and output efforts and its reciprocal between flows. It can be used as an *impedance transformer* in one physical domain, and as a *class transformer* between domains (**20**). The gyrator operates similarly, by establishing a relation between input effort and output flow, and its reciprocal between input flow and output effort.

† In bond graph terminology, this buffer is then said to operate in *derivative causality*.



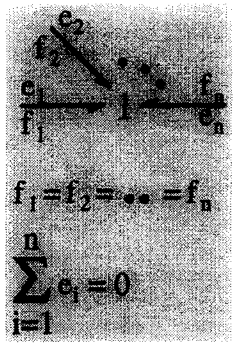


FIG. 2. Continuity of power across junctions is ensured by their constituent equations.

2.3. *The model context*

The first law of thermodynamics states that

"internal energy is conserved in processes taking place in an isolated system" (**8**).

For a system to adhere to the first law it has to be isolated. Because a completely isolated system is of little practical use, conservation of energy is achieved by explicitly specifying energetic interaction with the environment. This represents the system *context* and is modeled by *sources* and *sinks* of effort and flow, S_e and S_f, respectively. When all interaction is specified the net change of energy in the system is governed by the principle of conservation of energy. The change of energy in a system can be attributed to losses by dissipation through resistive elements. This energy loss needs to be modeled explicitly as a source of entropy in the case where it constitutes free energy. Note that although conservation of energy is the most fundamental law of physics, it is the hardest to enforce (**21**, **22**), since only the significant interactions are captured by the system model. An additional assumption in macrophysics is the constraint of *power continuity*, which is based on the assumption of conservation of energy. It is observed that energy cannot be annihilated at one point in a system and produced at the same rate at another point. It has to traverse the intermediate space (**7**). Therefore, any physical system not only conserves energy, but by nature is continuous in its signal (power) variables, effort and flow.

2.4. *Temporal character*

The constituent equations of the basic elements can be categorized temporally as either *instantaneous* or *integrating*. Algebraic relations are instantaneous; they are embodied by resistive elements and junctions. Because of the nature of the time integral relation, integrating effects evolve over time; they are associated with the buffer elements, capacitances and inductances, that are independent, i.e. they impose additional degrees of freedom on the system. Buffers that cannot be assigned an initial value independent of the rest of the system, are dependent. Typically, this occurs when

- a source or sink is modeled to enforce a specific amount of stored energy on a buffer (source-buffer dependency), or
- buffers are directly connected to each other without intervening dissipators between them (buffer–buffer dependency).

These two situations can be directly attributed to choices made when designing the system model. In the first case, component mechanisms that are assumed to have insignificant effects with respect to the modeling task or scope are neglected and buffer dependency only introduces additional loading effects. In the second case, the dependent buffers most likely represent the same buffer effect. A lumped parameter assumption can be imposed to replace the dependent buffers by an equivalent combined buffer.

2.5. *Summary*

Bond graphs provide a powerful modeling formalism (**4**, **19**, **23**) for physical systems which can be defined in terms of a small set of domain-independent physical mechanisms: the ideal sources (S_f, S_e), an energy dissipation element (R), two energy storage elements C and I, and pure lossless energy distribution elements, the 0- and 1-junctions and transformers TF and gyrators GY. These mechanisms provide an elegant basis for analyzing the continuous behavior of physical systems by inherently enforcing energy conservation and power continuity constraints on system models. Moreover, they provide an intuitive interpretation of energy flow and storage in a system, and can be used to generate state space equations with state variables that are directly linked to the system's energy storage elements.

III. The Nature and Effects of Discontinuities

This section introduces the background for modeling abstractions that lead to the generation of discontinuous behavior. To systematically analyze the resultant mixed continuous/discrete behavior of the system, a hybrid representation that includes bond graphs and finite state automata is established. Discontinuous changes are modeled to manifest as sequences of instantaneous switches that may produce changes in system configuration, and, therefore, the relations among state variables. We establish the *invariance of state* principle as governing the analysis of energetic behavior of the system during discontinuous switchings, and the derivation of the initial state in a new continuous mode of operation after the switchings.

3.1. *The nature of discontinuities*

Physical systems by nature are continuous, and discontinuities are artifacts of simplifications and assumptions introduced into the system model. In general, discontinuities in behavior generation can be attributed to two abstraction phenomena:

- *time scale abstraction*, and
- phenomena or *parameter abstraction*.

The time scale for the actual nonlinear behavior of the system may be much faster than the time scale at which system behavior needs to be analyzed. If system behavior were explicitly modeled at this small time scale, appropriately positioned small energy storage and dissipative effects have to be included in the system model. The ensuing time constants may obscure or complicate the generation of the more gross (or abstract) phenomena that are of interest, therefore, time scale abstraction techniques may need to be introduced to focus on the more useful behaviors. Furthermore, small time

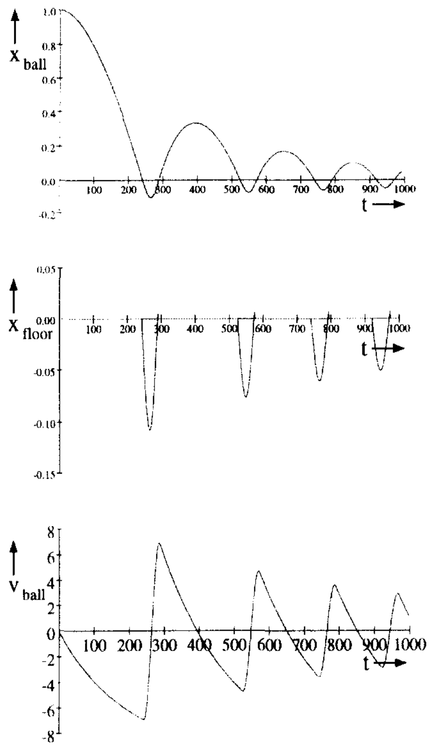


FIG. 3. An ideal elastic collision between a rigid body and a very stiff floor.

constants cause steep gradients in system behavior, which results in numerical stiffness problems when conventional simulation methods are used, or lead to the use of less accurate implicit integration methods to avoid stiffness problems. For example, consider a ball bouncing on a floor. When the ball on its downward trajectory hits the floor, elastic compression of the floor, or the ball, or both, enables storage of the ball's kinetic energy as internal compression energy which eventually builds up a negative force that imparts a negative velocity to the ball causing it to fly back upward. Therefore, the velocity of the ball changes continuously, but starting from the point of impact, for a short time interval, the velocity changes with a very steep slope. This is shown in Fig. 3 for a system where the ball is considered to be an ideal rigid body and the floor is modeled to have a relatively large stiffness value. The effects of compression are indicated by the ball and floor displacement becoming negative for a very short period of time. During these periods the ball velocity undergoes a steep change from a negative value (downward motion) to an equal and opposite positive value (upward motion). When modeled in less detail, the velocity of the ball can be instantaneously negated. If the interaction between the ball and floor is modeled as an ideal elastic collision the overall behavior implies that the ball continues to bounce but with decreasing amplitudes because of air resistance. To achieve simpler, but correct behavior generation, the stiffness effects are abstracted into an instantaneous change in velocity and the bouncing ball model combines continuous behavior components with *abrupt* or discontinuous behavior changes. Since the elasticity of the ball or floor is not abstracted

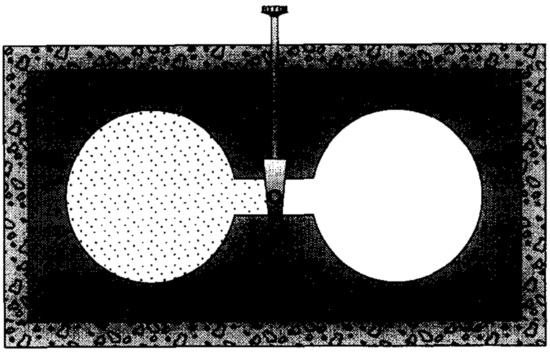


FIG. 4. Free expansion of a gas by diffusion after an instantaneous change.

away, but condensed into an instantaneous effect, this is an example of time scale abstraction. A more detailed discussion of this phenomenon is presented in Section V.

A second cause for discontinuities in models can be attributed to component parameter abstractions. The effects of particular component characteristics, often parasitic terms, are simplified or ignored. However, this may reduce the degrees of behavioral freedom in the system by making buffers dependent and make it hard to ensure that system behavior remains continuous. For example, if the ball and floor elasticity in the bouncing ball example are abstracted away, the system shows an ideal *nonelastic* collision and the ball comes to an immediate stop at the point of impact. Because of the rigid body assumption, the kinetic energy of the ball dissipates instantaneously the moment its velocity is forced to be 0. In reality, small elasticity or deformation effects allow the ball to maintain a velocity that is not directly coupled to the floor. Again, the behavior exhibits a steep gradient, but remains continuous. This implies that the elasticity and deformation effects introduce additional degrees of freedom in the system model that result in continuous system behavior. On the other hand, if these effects are very small, and they have a negligible effect at the macroscopic level, they can be abstracted away in the system model. In that case, at the point of impact, the ball with nonzero velocity is directly coupled to the floor which has 0 velocity, and a discontinuity occurs.

3.2. *Effects of the lumped parameter assumption*

To study the lumped parameter assumption in detail, consider the free expansion experiment conducted by Gay-Lussac and Joule, shown in Fig. 4 (**21**). A chamber is made up of two connected bulbs with an idealized open-closed valve between them. Initially, only the left bulb contains a gas and the valve is closed. When the valve connecting the two volumes is opened, the gas in the left bulb expands freely and starts diffusing into the right bulb. Even if the connecting orifice is non-resistive, this diffusion introduces nonhomogeneous turbulence effects that are active for a period of time. This is not an issue from a thermodynamics perspective, where the goal is to establish energy balance after the distribution has become homogeneous (e.g. determining the temperature of the water in the compartment surrounding the two volumes). Therefore, the lumped parameter assumption holds. Based on the underlying modeling assumption, the transitional effects are negligible to the time scale of interest during continuous evolution. However, the lumped parameter assumption does not hold during discontinuous changes. In case of the free expansion experiment, if the valve were closed quickly enough after opening to operate as a sequence of two instantaneous changes, the gas could not have diffused yet. So, the homogeneous distribution of gas over the two volumes is never actually established. In fact, an immediate closing of the opening leads to no redistribution of energy which conforms with the observation that in real time there never was a connection.

3.3. *The effects of discontinuities*

Discontinuities in physical system models have a number of effects that do not occur in the analysis of continuous system behavior:

- buffers may become dependent, thus changing the dimension of the state vector which may result in an apparent violation of conservation of energy.
- a discontinuous change may trigger a chain of discontinuous changes.

These situations may cause continuity of power to be violated, a phenomena that cannot occur when physical systems are modeled with traditional bond graph elements. They are described in detail next.

3.3.1. *Buffer dependency and conservation of energy.* A consequence of the modeling assumption that small energy storage or dissipative effects can be abstracted away is that this may result in a direct coupling between other previously independent buffers in a system model. Because of idealized connections, these coupled buffers then become dependent, operate as one, and thus the number of degrees of freedom in the system is reduced. In modeling these connections as ideal, the microscopic effects manifest as instantaneous changes at the macroscopic level. Based on the lumped parameter assumption, dependent buffers have to be treated as one which may require a discontinuous change of stored energy. A configuration change resulting in buffer dependency may cause a Dirac pulse† generated at the instant the change occurs. The pulse represents an amount of energy that dissipates discontinuously as heat, much like the loss of energy due to resistive dissipation which is modeled as a source of entropy. If the thermal domain were explicitly modeled, the loss of energy by the Dirac pulse may be modeled as a source of entropy. In other words, this implies that the system transfers energy to the environment which is modeled as a source of entropy (actually a sink). If the environment is assumed to be *isothermal*, this source is not modeled explicitly. Note that the instantaneous loss of energy would occur during a short time interval if small dissipative elements capturing the energy redistribution effects were introduced into the model in the connection between the buffers.

3.3.2. *Sequences of changes.* Discontinuous changes occur when signal values cross a threshold level during continuous evolution of the system (**5**, **24**, **25**). The effect of this is that energetic connections are activated or deactivated. This may cause adjoining signals to change discontinuously, and cross switching thresholds themselves. The result is the activation or deactivation of one or more energetic connections, leading to a sequence of discontinuous changes.

Because of the modeling assumption that discontinuous changes are instantaneous,

† This is a pulse of finite area, but infinitesimal width that occurs at a point in time.

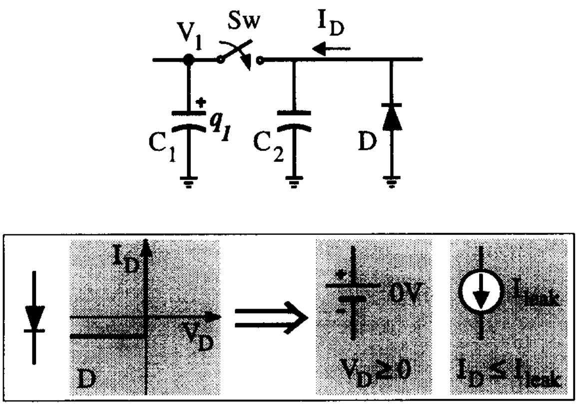


FIG. 5. Two capacitors that become dependent and switch modes of operation of the diode.

such a sequence of discontinuities occurs instantaneously and real-time does not *progress* until a model configuration is reached where no more switches occur. These sequences of instantaneous changes make it difficult to infer the new mode of continuous operation. Additionally, the system state needs to be advanced across a series of model configurations to correctly map the state vector of the last continuous mode onto the new one.

3.3.3. *Illustration.* Consider an electrical circuit with two capacitors in parallel connected by an ideal switch. When the switch is open, the two buffers can charge and discharge independently, but they become dependent when the connecting switch is closed requiring the two buffers to achieve a common potential. The total charge on the two capacitors before the switch was closed has to be preserved so that the physical principle, conservation of charge, is not violated. Assuming that the initial charge on C_1 is q_1, and there is no initial charge on C_2, the common potential after the switch is closed is $V' = q_1/(C_1+C_2)$. The amount of energy before closing the switch is $q_1^2/2C_1$. After closing the switch the charge on C_1 is $q_1' = [C_1/(C_1+C_2)]\,q_1$ and the charge on C_2 is $q_2' = [C_2/(C_1+C_2)]\,q_1$, therefore, the amount of energy in the system, $q_1'^2/2C_1 + q_2'^2/2C_2$, is $q_1^2/[2(C_1+C_2)]$. This implies that closing the switch causes a loss of energy equal to

$$\frac{q_1^2}{2C_1} - \frac{q_1^2}{2(C_1+C_2)} = q_1^2 \frac{C_2}{2C_1(C_1+C_2)}. \tag{1}$$

Imposition of the conservation of charge principle appears to result in an instantaneous loss of energy in the system, i.e. the conservation of energy principle is violated. As described, this loss has to be explicitly modeled in case the environment is not isothermal.

To illustrate a sequence of discontinuous changes, consider the effect of the diode (Fig. 5) that operates in one of two possible modes:

- an effort source; it enforces 0 V, independent of the current.
- a flow source; it imposes a negative leakage current, independent of the voltage.

Initially, the voltage drop across the diode is 0 and it operates in its effort source

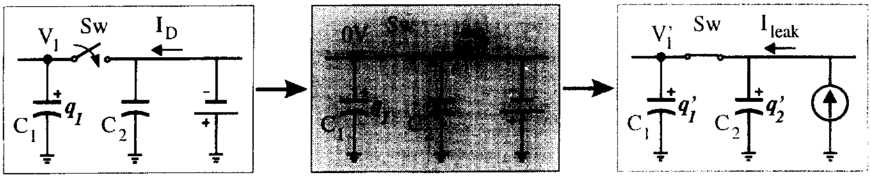


FIG. 6. A series of discontinuous changes may contain mythical modes.

mode (see Fig. 6). When the switch is closed, this effort source enforces 0 V on both of the capacitors which requires C_1 to discharge instantaneously. This results in a current flow that approaches negative infinity and based on the switching specification of the diode ($I_D \leqslant I_{leak}$) the model configuration changes immediately to one where the diode operates as a current source. Since no more discontinuous changes occur, the capacitors become dependent and redistribution of charge occurs as in the two capacitor system described above. Imposing the modeling assumption that discontinuous changes occur instantaneously, the model configuration where the diode operates as an effort source and the switch is closed is departed instantaneously and never achieved in reality. Therefore, the infinite current is never actually established, it is only used to infer the new mode of continuous operation. If it were considered a real mode of operation, C_1 would discharge instantaneously during the intermediate mode of operation. Therefore, when the diode switches to its current source mode of operation no energy would be left to maintain the leakage current.

3.3.4. *Limitations*. The occurrence of a number of discontinuous changes may make buffers alternate between mutual dependence and independence during a sequence of changes. From a physical perspective this situation has to be analyzed carefully.

To illustrate a situation where this may lead to a conflicting system model, consider the electrical circuit in Fig. 7, where the relay turns off when the voltage drop across C_1 equals the voltage drop across C_2. If the relay was on initially, the moment the switch is closed this condition holds and the relay opens. Because of the instantaneous nature of discontinuities, the model configuration where there is a connection between both the capacitors is departed immediately. As discussed in terms of the free expansion experiment, even though dissipative effects of the connection are not modeled, the energy redistribution still takes time. In real time, the basic model configuration does not change, i.e. the capacitors are disconnected all the time. However, the relay is open and its switching condition implies that the charge on both of the capacitors has been redistributed to reflect their equal voltage drop, which it has not. This indicates that

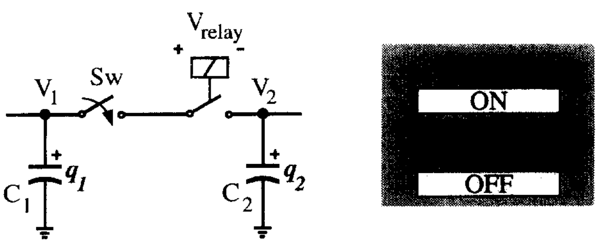


FIG. 7. Instantaneous buffer dependency changes cause problems.

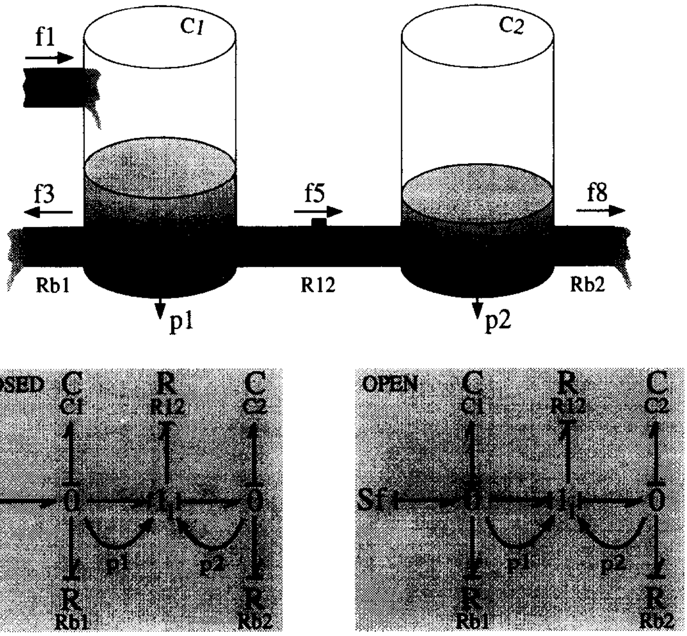


FIG. 8. Models of a bi-tank system with a closed and open latch.

bond graph theory in conjunction with instantaneous junction switching does not generate consistent behavior in this scenario, and either resistive or inductive effects of the connection have to be included, or another modeling approach has to be applied.

To summarize, switching conditions based on energy stored in buffers that are alternating dependent/independent within a sequence of discontinuous change are inconsistent with bond graph assumptions, and, therefore, prohibited. In this case, either model refinement or another modeling approach has to be chosen. Notice that gradients based on these energy variables *can* be used as demonstrated by the capacitor–diode example. Though there may not be an actual flow of current, a gradient exists the moment the switch closes, and this causes the diode to change to its mode of operation where it enforces a leakage current.

3.4. *A basic representation*

A scheme to accommodate discontinuous power changes needs to be introduced into the bond graph framework. An idealized discrete switching element can impose a discontinuous binary, *on/off*, relation on energy transfer paths in the system to model configuration changes. In a bond graph this ideal *switch* can be incorporated into energetic connections or junctions, that turn *on* (energetic connection is present) and *off* (energetic connection is absent). The expanded form of the junctions are referred to as *controlled junctions*.

Figure 8 shows two connected tanks that get disconnected when a latch closes. The bond graph fragments show that if the latch is closed, the 1_1 junction is deactivated, and the flow through this junction becomes 0. Therefore, there is no energy transfer across this junction, which implies that the bonds incident on the junction can be

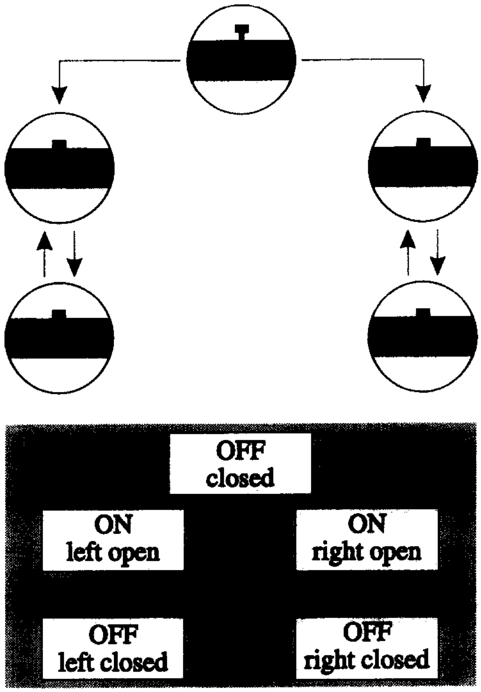


FIG. 9. A physical system may contain discontinuities that have memory.

eliminated in this system mode. The net result is that the two tanks become independent subsystems thus producing a seamless implementation of the mode-switching process.

The physical *on/off* state for each controlled junction is affected by continuous variables which cause the switching when variable values cross prespecified thresholds (e.g. $p_1 > p_2$ in the above examples). The control logic that governs the *on/off* relations of controlled junctions are defined by combinational or sequential automata and called the control specification, CSPEC. The input to a CSPEC are signal values (effort and flow) and the output is an *on/off* signal that turns the controlled junctions *on* (renders it active) and *off* (renders it inactive). The state transition diagram determining the CSPEC may have several internal states that map onto the *on* and *off* states of the controlled junction.

The need for sequential control logic for a controlled junction is demonstrated for the latched bi-tank system in Fig. 9. Initially the latch in the connecting pipe is in its upright position. When a threshold pressure difference is exceeded, the latch opens to the left or to the right. Once it has opened in one direction, it cannot open in the opposite direction anymore, therefore, future *on/off* states of the controlled junction are a function of the past state of the latch.

An important observation applies to the *state vector* of a system, which contains a necessary and sufficient number of variables to completely describe the system state. In continuous physical systems the set of variables that describe the energy distribution in the system capture its entire history, and, therefore, serve as the state vector for the system. However, when models contain discontinuities that are controlled by sequential logic, the continuous state vector cannot completely specify system history. Future behavior becomes dependent on the internal states of the sequential automata, and a hybrid state vector is required to capture the necessary sequential automata states as well. As an illustration, consider the latched bi-tank system. If the internal state of the latch in Fig. 9 is not known, future behavior cannot be determined uniquely based on energy variable values alone. The same energy distribution can result in two different states for the latch, depending on its past history, and system behavior can evolve along two different trajectories. To disambiguate this situation, the system state vector needs an additional discrete component, which specifies the model configuration at different points in time.

3.5. *Evolution of system behavior: invariance of state*

The capacitor–diode example in Section 3.3.3 illustrates that a discontinuous change may generate additional discontinuous changes. In general, one change may trigger a chain of discontinuous changes. Within our modeling framework complex computations due to nonlinearities and fast time constants are simplified by representing them as discontinuous changes that occur instantaneously. Therefore, modeled time during such a series of changes does not progress. Any system configuration that occurs during a sequence of switches without an intervening continuous state, has no real existence. These system configurations are transitional, or *mythical*,† because the system does not have separate physical manifestations in the intermediate mythical modes. A further consequence of this is that the system cannot exchange energy with its environment during this period, in other words, it is *isolated*. This is compatible with the fact that there is no redistribution of stored energy within the system during this sequence.

If switching specifications are such that at any point in a discontinuous sequence of switches the system comes back to an already generated discontinuous configuration, a loop of discontinuous changes ensues. Since discontinuous changes are assumed to occur instantaneously, this implies that system behavior stops progressing or diverging in real time. This is obviously in conflict with physical reality, and *divergence of time* constitutes an important condition for verifying a system model for physical consistency. Discussion on the divergence of time principle also appears in other work (**16**, **25**). The principle that energy is not redistributed during discontinuous, instantaneous changes, but only after a new mode is reached is termed the principle of *invariance of state*. Our algorithm for deriving discontinuous changes, described in Section IV, and the corresponding model verification technique, discussed in Section V, are based on this principle.

Note that correct inference of model configuration changes requires all buffers that become dependent to be analyzed in derivative causality, possibly invoking Dirac pulses. Though integral causality may be chosen for one of a set of dependent buffers, during continuous evolution of the system, discontinuous changes may cause steep changes in the energy stored in them, which may lead to a sequence of switches.

† Our notion of mythical configurations differs from the de Kleer and Brown (**26**) notion of mythical causality and time, which was used to describe and explain nonequilibrium behavior of dynamic systems.

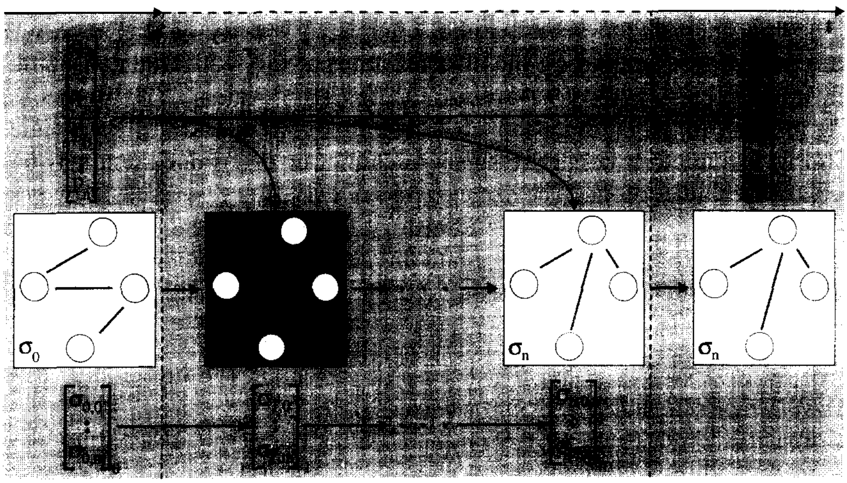


FIG. 10. During mythical changes the discrete state vector changes to reflect configuration changes. The continuous state vector is only changed once a real mode is reached.

3.6. *Transferring the state*

After a new mode of continuous evolution is inferred, the new system state has to be derived. This is referred to as the *initial value problem*. The discussion above implies that the continuous part of the state vector cannot change during a sequence of discontinuous change (see Fig. 10). This implies that the energy distribution in the initial state of a new continuous configuration or mode is derived from the energy distribution in the last state of the previous continuous mode of operation. When energy redistribution occurs, depending on possible buffer dependencies, three situations characterize the initial value problem that results from discontinuities in a model:

- switching modes causes no buffer–buffer or source–buffer dependencies. In this case, the transition causes no change in the continuous state vector.
- transitions between real modes cause two or more buffers to become dependent, and this causes the size of the state vector to change. As discussed, individual energy variable values change, but their total remains the same so conservation of state holds. The new energy distribution is determined by the total value of stored energy in buffers that are involved and the ratios of their parameters (**27**). Mode transitions may result in loss of energy to the environment as a result of changes in the system configuration.
- mode transitions cause source–buffer dependency. In this case, the switching causes a source (i.e. the environment) to instantaneously transfer energy into or out of the system. The new values of stored energy in the buffers involved is set to the source enforced values after the new real mode is established.

As a result, the system transfers from a real mode to another, but the initial energy distribution in the new mode may be different from the distribution in the last real mode. This introduces a final step in switching model configurations. When the new real mode has been established, the initial value problem is solved taking into account the buffer dependencies that have occurred as a result of configuration changes in the

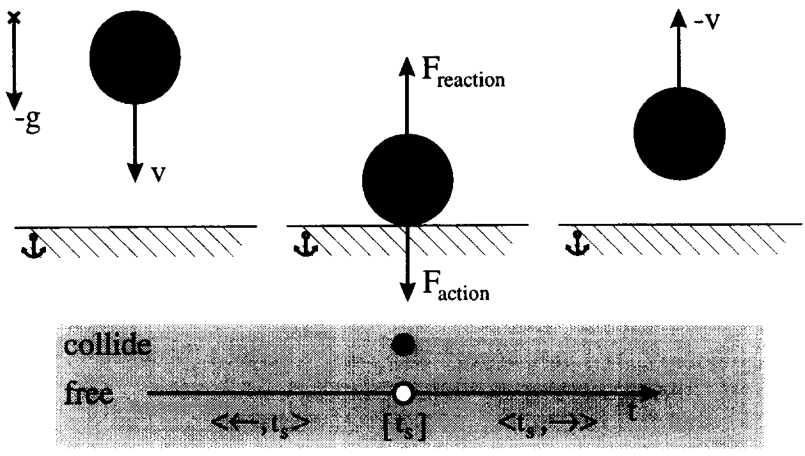


FIG. 11. Real modes can have a point or interval presence in time and have to ensure time continuity.

model. Next, simulation continues from the point in real time where the sequence of discontinuous changes was initiated.

3.7. *Temporal evolution in behavior transitions*

The semantics of behavior generation from the hybrid models needs to combine real modes with continuous evolution of system behavior with instantaneous discrete behavior changes where real time does not advance. Note that a real mode of system behavior can encompass an *interval* or *point* in real time, whereas all discrete changes have to occur at well defined points in time. Consider the example of the ideal elastic collision of the bouncing ball shown in Fig. 11. Model configurations where the ball is moving freely (up or down) represent continuous modes of operation where system behavior evolves over time. The system model is abstracted so that the collision process is ideal and elastic, and holds only at a point in time at which the ball momentum is reversed. If the ball and floor were in contact for any period longer than a point in time, the ball's momentum would transfer to the floor, and it would come to rest. On the other hand, if this real mode did not exist, i.e. the ball and floor never touched, the ball could not exchange momentum with the floor, which implies its velocity would never reverse. The configuration where the ball is in contact with the floor is abstracted to a point, which is then followed by an interval of time where the ball travels upwards exhibiting continuous behavior.

In summary, this method for representing the bouncing ball behavior as a discontinuous change from a real mode (moving downward) to a second real mode (point of contact with the floor modeling the collision) and then a discontinuous change to a third real mode (moving upward) is much cleaner than a discontinuous model which represents the reversal in ball velocity as an initial value problem [e.g. (**28**)]. In the latter situation, the point in time at which the collision occurs is considered to be the start point of the second time interval with the ball moving freely upward, and the model specifies the initial velocity of the ball at the start point. This model implies the ball is always moving freely, up or down, reversing its velocity at a particular position with no explicit physical phenomenon, such as a collision, to account for the change. The net result is a model that violates the principle of invariance of state, because stored momentum changes abruptly without explicit interaction with the environment.

IV. The Hybrid Modeling Scheme

The hybrid modeling scheme for physical systems has to ensure the generation of physically consistent behavior that in a gross sense approximates the true behavior of the system. The correctness of the model can only be determined by comparing true system behavior with model behavior. However, exhaustive validation of this type is impractical, therefore, as a first step, models can be verified by ensuring that none of their behaviors violate general laws of physics.

Section III outlined the issues and principles that characterize and govern behavior generation:

- the principles of conservation of state and energy for continuous system behavior; this is adequately ensured by the bond graph formalism.
- the principles of invariance of state, divergence of time, and temporal evolution in behavior transitions for discontinuous behavior. They are summarized below.
  1. Principle of *invariance of state*. During discontinuous changes, continuity of power cannot be guaranteed, however, conservation of energy embodied in the invariance of state principle has to hold. This means that ideal discontinuous modeling elements, switches, do not dissipate energy. If dissipative effects have to be modeled, appropriate resistance elements or heat sinks have to be introduced into the model.
  2. Principle of *divergence of time*. A sequence of discontinuous changes must always terminate in a continuous mode of operation with a well-defined system state derived from the last continuous system state. In other words, the system model has to adhere to the principle of divergence of time and conservation of state for it to be physically correct. The modeling formalism must provide means for ensuring that divergence of time is not violated and include support to help identify subsystems that violate this principle. Further analysis could then result in the model being refined to a correct one.
  3. Principle of *temporal evolution in behavior transitions*. A continuous mode of operation may persist at a point in time or during an interval. In case of an interval that is left-open, changes in model configuration cannot cause discontinuous changes in the energy distribution in the system. Discontinuous changes are caused by buffer-dependency, therefore, this forms an additional verification mechanism to ensure physical correctness of the hybrid system model.
- interaction between the discontinuous and continuous parts must adhere to the principle of invariance of state. Furthermore, when discontinuous changes occur, causal relations among variables in the system may change. The modeling formalism must have a consistent, algorithmic, scheme for causality assignment as a function of system configuration.

This section presents in detail the formal specification of the hybrid modeling scheme:

- the modeling language, and
- the behavior generation algorithms.

We begin by establishing the nature of discontinuities in bond graphs, and compare this with previous approaches to incorporate switching elements in the bond graph framework.

4.1. *Discontinuities in bond graphs*

Bond graphs are based on principles of conservation of energy and continuity of power. Discontinuous changes represent modeling abstractions which may violate these principles. In the bond graph framework, discontinuities have to be dealt with at a *meta-model* level, where the energy model embodied in the bond graph scheme is suspended in time, and discontinuous model configuration changes happen instantaneously. Therefore, the meta-model describes a control structure that causes changes in bond graph topology using idealized switches that do not violate the principles of energy distribution in the system imposed by the bond graph. After a new model configuration is derived, the model state is transferred from the previous configuration to the new one. Further switches may occur, and the meta-model is active until a bond graph configuration where no more switches occur is derived. At this point, the principles of conservation of energy and continuity of power govern the evolution of continuous system behavior.

To keep the overall behavior generations consistent, the meta-model control mechanism and the energy-related bond graph models are kept distinct. The configuration changes are implemented as local structural changes, where model components get connected or disconnected at junctions controlled by the meta-model mechanism. Note that the actual switch mechanism is not modeled as a bond graph element. However, changes in configuration result in reassignment of causality in the bond graph. This will become clear in examples presented later.

4.2. *Previous work*

Initial attempts to include discontinuous *switching* behavior in bond graphs used nonlinear resistances (**29**) and modulated transformers (MTF) (**19**, **29**, **30**). The main disadvantage of a nonlinear resistance is its dissipative nature, and this violates the principle that configuration switching be ideal and lossless. The MTF approach switches its transformation factor between 1 and 0 to establish and break connections. However, a modulated transformer with transformation factor 0 allows propagation of causality in the bond graph model, which interferes with the ideal configuration switching and resulting re-assignment of causality.

To circumvent this problem, a combination of a modulated transformer and resistor was used (**31**). This allows for configuration changes in a bond graph framework without change in causality. As an example, consider the two basic configuration changes in Fig. 12. To disconnect the buffer element and resistor, either their current (series connection) or voltage (parallel connection) can be made 0. In both cases, power transferred across the junction becomes 0 and the elements are disconnected. To model a disconnecting series connection (1-junction) the MTF–R combination always has flow causality. Its dual for disconnecting parallel connections (0-junction) is also shown

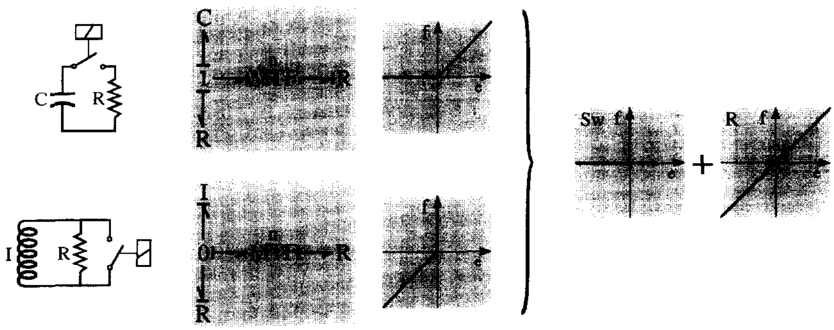


FIG. 12. The use of a modulated transformer for switching.

in Fig. 12. The presence of the resistance, R, in the model, again implies that switching involves dissipation, therefore, there can be no ideal switching. Switching is a combination of an R and Sw element, where the switch Sw would constitute a primitive relation. To model an ideal switch requires a combination of two MTF–R constructs, performing complementary switching actions. Furthermore, the resistance again introduces additional complexity and numerical stiffness in behavior generation. Moreover, the resistances are parasitic, and their small values are hard to estimate accurately. Finally, in case of a number of switches, the large number of parasitic resistances can lead to algebraic loops.

Recently, Strömberg, Top and Söderman (**12**, **32**, **33**) introduced an ideal *switch* to model ideal switching behavior, and Broenink and Wijbrans (**11**) introduced *switching bonds* to handle discontinuities in models. The ideal switch, Sw, is a new bond graph element that either enforces 0 effort or 0 flow on a junction to turn it on or off (see Fig. 13). Note that this represents a degenerate form of the regular source elements. Because of the 0 effort or 0 flow, power ($effort * flow$) associated with a switch is always 0, which conforms with the concept that switching be ideal and introduced to control configuration changes in the system models. However, Strömberg *et al.* define the ideal switch as part of the energetic model, rather than a component that handles control actions. As noted earlier, the ideal switch should be used only as a *transitional* element

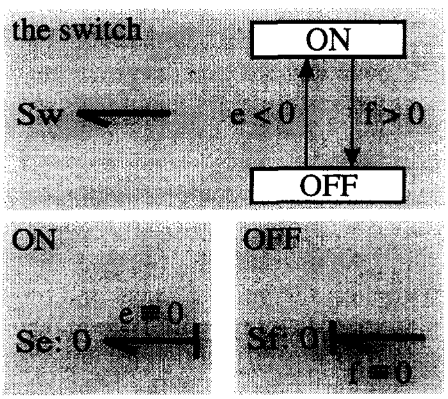


FIG. 13. Operation of the switching element.

that is used to infer a new mode of operation from an existing one. Furthermore, it is an abstraction of physical phenomena so it should have no direct link to energetic computations like the other bond graph elements. Though alleviated by the introduction of *switching fields* (**34**), the use of these switches obscures hierarchical structure descriptions. Typically system mode changes occur when switches turn on and off. Relations among switches are hard to identify and their use as bond graph elements clutters the model. Finally, as demonstrated in this paper, localized discrete effects require a global configuration change algorithm, which is not part of the switched bond graph modeling paradigm.

Switching bonds fit the conceptual framework that separates energetic model aspects from control mechanisms. They conform to a structural description of real-time systems, which consists of a *data flow* and a *control flow* part (**35–37**). Lent conjectures that this holds for systems in general (**38**). For switching bonds, the data flow part is represented by the bond graph formalism and the control flow part is represented by a *control box* that contains switching logic modeled as global finite state automata or Petri nets. System configuration changes occur through these switching bonds, which either establish or break energy connections between components and subsystems. Though the concept of an integrated data and control model has a lot of merit, interaction between the two through switching bonds causes problems in the form of hanging junctions. Even more important, changing boundary conditions due to switching are often handled incorrectly (**33**).

To address the problems of idealized switches as bond graph elements and switching bonds, Mosterman and Biswas (**14**) proposed the use of *controlled junctions* in the bond graph framework to handle idealized configuration changes. The controlled junction recognizes the presence of energy and control flow structures, and the need for two distinct, but interacting formalisms that operate within an integrated framework to model physical systems. The interaction between the two formalisms; the control structure and the bond graph occurs via controlled junctions (rather than bonds) that can switch *on* and *off*, thus effectively providing the ability to switch configurations within the bond graph framework. The use of *finite state automata* (**39**, **40**) to model the control mechanisms creates a systematic and powerful *hybrid* modeling framework. After formally specifying the modeling language, we demonstrate that the interaction between the two formalisms can be proven to be consistent, and, therefore, results in verifiable physically correct models.

4.3. *The modeling language*

A system with n components each with k behavior modes, can assume n^k overall configurations, however, in practical situations, only a small fraction of the configurations are physically realized. In previous work, researchers have defined a number of approaches, such as transition functions [e.g. (**28**)], finite state automata and switching bonds (**11**), to handle discrete changes in physical system configuration and behavior. Most of these methods assume that the range of system behaviors are pre-enumerated, but, in general that can be a very difficult task. Recent compositional modeling approaches (**41**, **42**) overcome this problem and build system models *dynamically* by composing model fragments. Our hybrid modeling scheme adopts this meth-

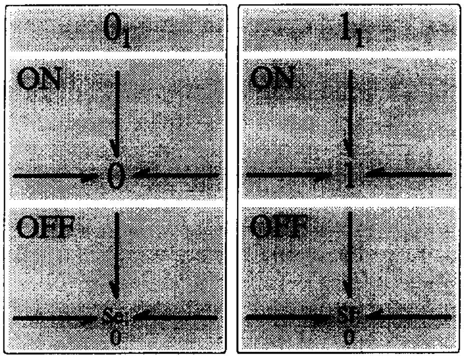


FIG. 14. Operation of the controlled junction.

odology, and implements a dynamic model switching methodology in the bond graph modeling framework.

Instead of identifying a global control structure and pre-enumerating bond graph models for each of the modes, the overall physical model is developed as one bond graph model that covers the energy flow relations within the system. Discontinuous mechanisms and components in the system are then identified, and each mechanism is modeled locally as a *controlled junction* which can assume one of two states—*on* and *off*. The local control mechanism for a junction is implemented as a *finite state automaton* and represented as a *state transition graph* or *table*.

4.4. *Controlled junctions*

Controlled junctions operate in one of two modes: *on* and *off*. When *on*, controlled junctions behave like traditional 0- and 1-junctions. To invoke configuration changes, controlled junctions can be turned *off*, or deactivated, in which state they inhibit transfer of energy between model fragments. Each power bond connected to a deactivated controlled junction can be loaded by a source element of 0 value to correctly handle boundary conditions of the disconnected model fragments. When a 0-junction is deactivated, an effort source with effort value 0 is used as load. Similarly, for a deactivated 1-junction, a flow source with flow value 0 is used as load. In case of configuration changes it becomes critical to establish correct loading on all adjoining bonds. In most bond graph models of physical systems, 0- and 1-junctions appear in alternating sequence, therefore, when a junction is deactivated, its adjoining power bonds can be removed from the bond graph to establish the new model configuration. Also, the effort causality as enforced by a deactivated 0-junction typically does not propagate across neighboring 1-junctions. Neither does flow causality of a deactivated 1-junction propagate across adjoining 0-junctions. By enforcing 0 effort or flow, the transfer of energy across these junctions becomes 0, therefore, controlled junctions exhibit *ideal switch* behavior. Modeling discontinuous behavior in this way is consistent with bond graph theory.

In our hybrid modeling representation, controlled junctions are marked with subscripts (e.g. 0_1 in Fig. 14). Each controlled junction has an associated finite state automaton that determines whether the controlled junction is in the *on* or *off* state. In other words, the finite state automaton defines a junction's control specification (CSPEC). The input of each CSPEC consists of

- power variables from the bond graph, and
- external control signals.

This input is depicted in the hybrid bond graph as arrows into the controlled junction. Mathematical operations may be applied to the power signals before they are input to the CSPEC. These operations can be modeled by a block diagram [as is used in the bond graph modeling tool CAMAS (**43**)] to manipulate signal values on active bonds. The output of the CSPEC sets the associated controlled junction to *on* or *off*. Internally the CSPEC can have any number of states, its control logic can be combinational or sequential, but each transition is a manifestation of a physical effect, therefore, its internal states have to map alternately to physical *on* and *off* states of the controlled junction. Conditions for a valid CSPEC are:

- each internal state must map onto an *on* or *off* state of the controlled junction.
- transition conditions on the edges have to evaluate to boolean values in each mode of operation.
- the CSPEC conditions have to result in at least one continuous mode of operation for all reachable energy distributions.

The set of local control mechanisms associated with controlled junctions constitute the *control model* of the system. The control model performs no energy transfer, therefore, it is distinct from the bond graph model that deals with the dynamic behavior of the physical system variables. Control models describe the *transitional*, i.e. mode-switching behavior of the system. A *mode* of a system is determined by the combination of the *on/off states* of all the controlled junctions in its hybrid bond graph model. Note that the system modes and transitions are dynamically generated, and do not have to be pre-enumerated.

The use of controlled junctions for an ideal nonelastic collision between a ball and floor is illustrated in Fig. 15. The ball inertia is modeled as m, the air resistance as R_1, and gravity as an effort source, mg. In the *off* state, there is no connection between the ball and the floor and the bonds connecting the ball to the resulting effort source with 0 value can be disregarded. The flow source with 0 value is disconnected as well. The CSPEC part of the model shows that at the point the ball touches the floor ($\int v_{\text{ball}} \leqslant 0$), the controlled junction turns *on* and the flow source with 0 value becomes connected to the mass. This implies that the ball velocity $= 0$, but the connection between the ball and floor source also causes a change of causality, forcing the inertial element to operate in derivative causality. The ball stays connected to the floor as long as the force it exerts is >0. When this force becomes negative, the junction turns *off* again, and the ball inertia becomes an independent buffer element again. Its momentum is reversed, and the ball flies up. This example illustrates a seamless integration of multi-mode behaviors in one model based on a local switching mechanism. Other examples of hybrid bond graph models are discussed in (**5**, **24**, **25**).

4.5. *Mode switching in the hybrid model*

The invariance of state principle discussed in Section 3.5 establishes that only the *discrete state* vector changes during discontinuous mode changes and this is determined

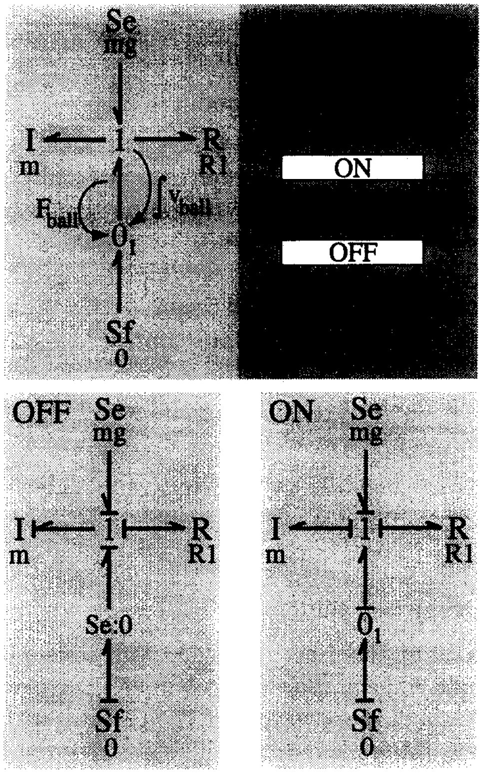


FIG. 15. Hybrid bond graph model of an ideal nonelastic collision.

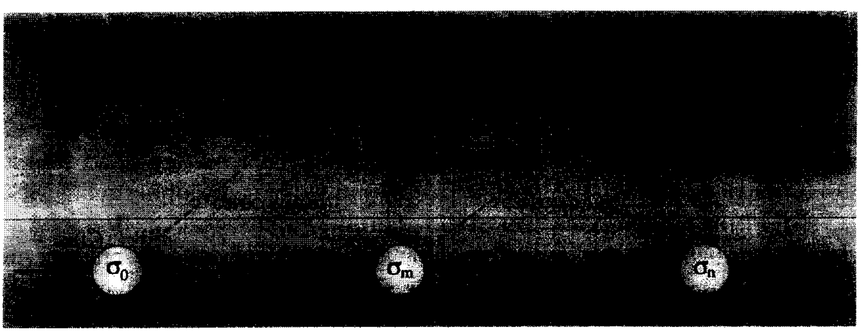


FIG. 16. A sequence of mythical mode switches.

by the finite state automata associated with the *controlled junctions*. After these changes result in a real mode of operation (where system behavior evolves as a function of time), the discrete switching ends and the continuous state vector for this new mode has to be derived. This is illustrated in Fig. 16. Mythical modes are depicted as dark circles and real modes as open circles. In real mode σ_0, a signal value crosses a threshold at time t_s, which causes a discontinuous change of model configuration to σ_1. Based on the original energy distribution (P_0, Q_0) values for the set of power variables (E_1, F_1) in this new configuration are calculated. The energy distribution (P_0, Q_0) is shown as the state vector **x**. The new values cause another instantaneous mode change and a new model configuration, σ_2, is reached. Again, the set of new power variables values (E_2, F_2), are calculated based on the original energy distribution (P_0, Q_0). Further mythical mode changes may occur till a real mode, σ_m, is reached. Now, the energy distribution, or continuous state variable values, of the departed real mode is mapped to the newly attained real mode. This is trivial when no buffer dependencies have occurred. Otherwise, conservation of state across dependent buffers yields the new individual state variable values based on the ratios of their buffer values. Finally, a new time interval has to be derived where the system behavior begins to evolve continuously. This is determined by finding the left limit of the system at the switching time, t_s. This limit may again result in a sequence of configuration changes $(\sigma_{m+1}, \sigma_{m+2}$, etc.) based on (P_m, Q_m) that terminates in a model configuration that persists over an interval of time, σ_n. Real time continuous simulation resumes at t_s^- so system behavior in real time implies that modes σ_0, σ_m, and σ_n follow each other instantaneously. The formal *Mythical Mode Algorithm* (MMA) is outlined below.

1. Calculate the energy values (Q_0, P_0) and signal values (E_0, F_0) for bond graph model σ_0 at time t_s.
2. Use CSPEC to infer a possible new mode given (E_0, F_0).
3. If one or more controlled junctions switch state then:
   (a) derive the bond graph for this configuration.
   (b) propagate causality using the SCAP algorithm **(19)** to establish buffer-dependency.
   (c) calculate the signals (E_i, F_i) for the new mode, σ_i, based on the initial values (Q_0, P_0).
   (d) use CSPEC again to infer a possible new mode based on (E_i, F_i) for the new mode, σ_i.
   (e) repeat Step 3 until no more mode changes occur.
4. Establish the mode, σ_m, as the new real system configuration at the point of discontinuity.
5. Map (Q_0, P_0) to the energy distribution for σ_m, (Q_m, P_m).
6. Take the left limit at the point of discontinuity of mode σ_m and calculate (E_m, F_m) based on (Q_m, P_m).
7. Use CSPEC to infer a possible new mode given (E_m, F_m).
8. Perform Step 3 with (Q_m, P_m) as initial values.
9. Establish the final mode, σ_n, as the new model configuration.
10. Map (Q_m, P_m) to the energy distribution for σ_n, (Q_n, P_n).

It is important to note that system behavior evolves in two ways:

- continuously in real time, where traditional bond graph models govern the generation process.
- discontinuously as instantaneous model configuration changes as defined by the MMA above.

The discontinuous changes represent meta-level control phenomena and are not part of the real time line. However, every discontinuous change sequence must be preceded and followed by a real mode of operation so that system behavior continues to evolve in time. This is discussed in detail in the next section. In addition, two different real model configurations cannot exist for the same energy distribution at any one time.†

Based on the MMA, a hybrid system model can be defined by a function $\mathbf{f}(t)$ describing the evolution of the continuous component of the state vector $\mathbf{x}(t)$ in a particular mode of operation σ_k with the input vector $\mathbf{u}(t)$

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \sigma_k, \mathbf{u}(t), t). \tag{2}$$

The vector σ_k includes the *on/off* states of all controlled junctions. When switching is initiated, a sequence of mythical states may ensue, with $\sigma_{k+1} = \sigma_k$ as the initial mode at time t_s.

$$\begin{cases} \sigma_{k+1} = \sigma_k \\ t_s = t \end{cases}. \tag{3}$$

The CSPEC conditions for switching between modes of operation are captured by the recursive function $\mathbf{g}$. The signal values at every new mode, $\mathbf{x}^+(t_s)$, are computed from the original vector $\mathbf{x}(t_s)$:

$$\begin{cases} \sigma_{k+1} = \mathbf{g}(\mathbf{x}^+(t_s), \mathbf{x}(t_s), \sigma_{k+1}, \mathbf{u}(t_s)) \\ \mathbf{x}^+(t_s) = \mathbf{f}(\mathbf{x}(t_s), \sigma_{k+1}, \mathbf{u}(t_s), t_s) \end{cases}. \tag{4}$$

The function $\mathbf{f}$ is calculated at the time of switching t_s. This continues till $\mathbf{f}$ produces no new modes. The last mode defines the new real state vector $\mathbf{x}^+(t_s)$ and σ_m. Advancing the state vector to $\mathbf{x}(t_s^+)$ may cause another sequence of configuration changes to occur. To infer these new configuration changes,

$$\begin{cases} \sigma_{k+1} = \mathbf{g}(\mathbf{x}^+(t_s^+), \mathbf{x}(t_s^+), \sigma_{k+1}, \mathbf{u}(t_s^+)) \\ \mathbf{x}^+(t_s^+) = \mathbf{f}(\mathbf{x}(t_s^+), \sigma_{k+1}, \mathbf{u}(t_s^+), t_s^+) \end{cases} \tag{5}$$

is invoked until no more configuration changes occur. Then, the model has reached a new time interval in which continuous evolution occurs. Note that the configuration changes that result from a transition at a point in time to an interval cannot cause a discontinuous change in the state vector $\mathbf{x}$, i.e. its left limit exists and has value $\mathbf{x}^+(t_s)$. This implies that energy re-distribution can only occur at points in time. In case of a left open interval, energy re-distribution would occur at time $t_s + \varepsilon$ but there is always a $\delta < \varepsilon$ that defines time $t_s + \delta$ at which energy re-distribution is required instead.

A complete simulation system that incorporates the MMA algorithm has been developed under Microsoft Windows using Visual Basic 3.0 Professional Edition (**44**), and tested on a number of physical system examples. The behavior generation algorithm has three key modules:

1. ESPEC, the continuous system simulator of the energy specifications as modeled by a bond graph, represented as function $\mathbf{f}$.

† Unless the system cannot be modeled in sufficient detail economically. In such situations, it is custom to allow additional abstraction that results in nondeterministic models. Nondeterministic models are not considered here.

2. CSPEC, that determines switching of controlled junctions, represented as function **g**.
3. MMA, the mythical mode algorithm which combines the discrete and continuous elements and coordinates the calculation of posteriori values $\mathbf{x}^+(t_s)$ and $\mathbf{x}^+(t_s^+)$ with recursive mode switching; when discontinuous switching is complete, it maps the continuous state vector from the previous real mode of operation to the new one.

Details of the implementation of this simulator system are presented elsewhere (5).

V. Physical Consistency

The principle of divergence of time implies that in a hybrid simulation of physical system behavior, a sequence of discontinuous changes should always terminate in a continuous mode so that system behavior continues to evolve in time. To verify that this condition is satisfied for a given system model, it has to be shown that every individual CSPEC and combinations of interacting CSPECs are consistent with the divergence of time principle.

To prove consistency, the key observation is that energy is not redistributed during a sequence of discontinuous changes. Any redistribution of energy takes place only at the end of the sequence of switches after a new mode of continuous operation is established. Thus the invariance of the energy distribution in the last real mode to switching can be used to verify physical consistency.

5.1. *Energy phase space analysis*

The CSPEC switching conditions are generally based on power variables, and they have to be mapped onto the energy variables in the previous real mode that was departed to exploit the invariance of state principle. The next step maps the functions onto an energy phase space (multidimensional space, where the energy variables define the orthogonal axes of the space). Note that the CSPEC functions define the boundaries in the energy space for the different modes of operation of the system. However, configuration changes may change the relations among power and energy variables in the different modes or model configurations, so the relevant CSPEC functions are plotted individually for each mode of the system. Transitions from one mode to another occur along boundaries (defined by the CSPEC functions), but as discussed earlier, the energy distribution is invariant under any sequence of transitions. Therefore, when switching conditions specify a transition from one mode of operation to another, there has to be a region of overlap in the energy phase spaces of these two modes. In other words, a nonnull intersection in the energy space between two modes, defines the reachability between the modes.

Based on this multiple energy phase space analysis, a necessary condition for divergence of time can now be established. Because of the invariance of the dimensions of the energy phase spaces across discontinuous changes, once a sequence of discontinuous changes is initiated from a real mode, the energy phase space must intersect at least one other mode or configuration where no further switching occurs, i.e. a second real mode is reachable from the first real mode. Otherwise, a sequence of discontinuous changes will never terminate and the model never returns to a configuration where time can evolve continuously again. To summarize, divergence of time in the hybrid model is verified as follows:

1. Determine all model configurations.
2. Map CSPEC switching conditions onto the energy distribution in the previous real mode.
3. Construct an energy phase space for each configuration and identify the transition regions or boundary conditions for each of them.
4. Form the conjunction of all energy phase spaces.
5. If the conjunction shows transition areas with no real mode involved, divergence of time is violated for the related energy distribution in the model.

This multiple energy phase space analysis provides a powerful qualitative methodology for building verifiably correct hybrid models of physical systems based on the principle of divergence of time. Its application is illustrated next.

5.2. *Examples*

Consider an ideal elastic collision illustrated in Fig. 17. When the ball hits the floor with a velocity v, for an ideal elastic collision model, it instantaneously reverses its velocity and starts traveling upwards. In the corresponding hybrid bond graph model, the change of velocity is shown as a modulated flow source which delivers its power at

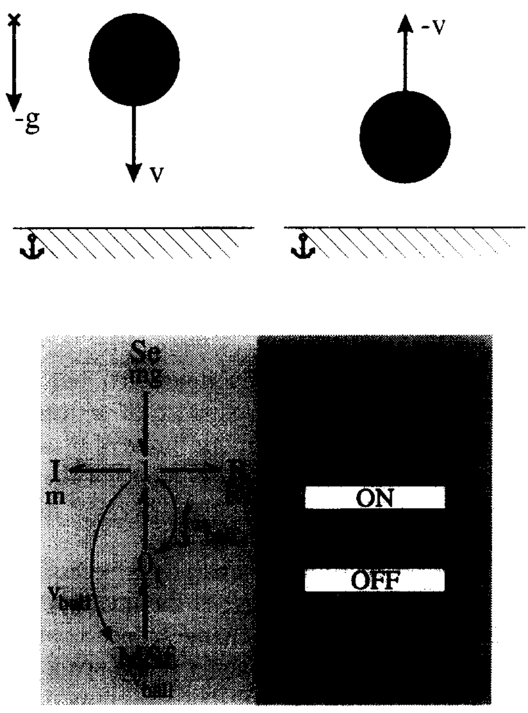


FIG. 17. An ideal elastic collision.

the point on time when the ball hits the floor. The moment the ball hits the floor, the controlled junction 1 (0_1) comes on, as specified by its CSPEC. The moment the ball starts to move upwards, the controlled junction turns off, and the ball is modeled by an independent bond graph fragment again. Note that the momentum returned by the floor to the ball could have been set to any value ($\leqslant 100\%$) by the modeler. This is because the floor is not part of the system and any exchange of energy with the environment has to happen through effort and flow sources. Also, the modulated flow source is only active at a *point* in time. In the MMA, this point causes an energy re-distribution and then a mode switch. In the new mode of operation, at $t+\varepsilon$, the controlled junction 0_1 is *off* again.

Next, to demonstrate a bouncing ball slowly coming to rest on the floor, we consider a combination of an ideal elastic and an ideal nonelastic collision. Here the ball bounces with a decreasing amplitude because of air resistance. At a certain time, it is observed to come to rest, and this is modeled as an ideal nonelastic collision. This nonelastic mode of operation is activated when the ball hits the floor with momentum less than a pre-set threshold value. A complete model of this system that includes the combination of the effects of an elastic and nonelastic collision is shown in Fig. 18.

To analyze the physical consistency of this model, the switching conditions in the CSPECs have to be expressed in terms of stored energy in the last real mode before switching. The stored energy space is two dimensional, one dimension for the momentum of the ball, p_{ball}, and the second dimension for the position of the ball, x_{ball}. All CSPEC transitions are specified in terms of energy variables, except for the force exerted by the ball on the floor, F_{ball}. This is computed from the bond graph as $F_{\text{ball}} = -F_g + F_m + F_{R1}$. When the value of F_{ball} is negative, the ball disconnects from the floor. To derive the conditions under which the controlled junction 1 transitions from *on* to *off* during a sequence of switches, F_{ball} has to be expressed in terms of stored energy variables x_{ball} and p_{ball} when junction 1 switches *on*. The buffer dependency that arises when the flow source becomes active forces F_m into a derivative relation

$$F_m = \frac{\mathrm{d}p}{\mathrm{d}t}. \tag{6}$$

Note that when the controlled junction 1 is *off*, $p_{\text{ball}} = mv_{\text{ball}}$, but when the junction

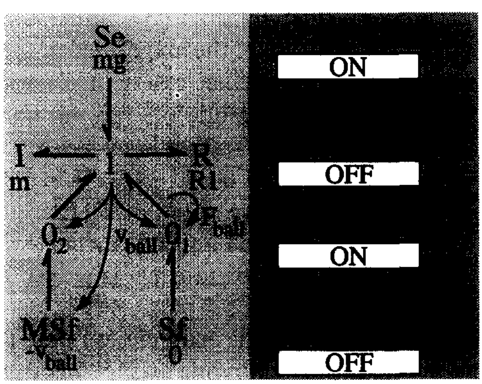


FIG. 18. Hybrid bond graph model of a combined ideal elastic and nonelastic collision.

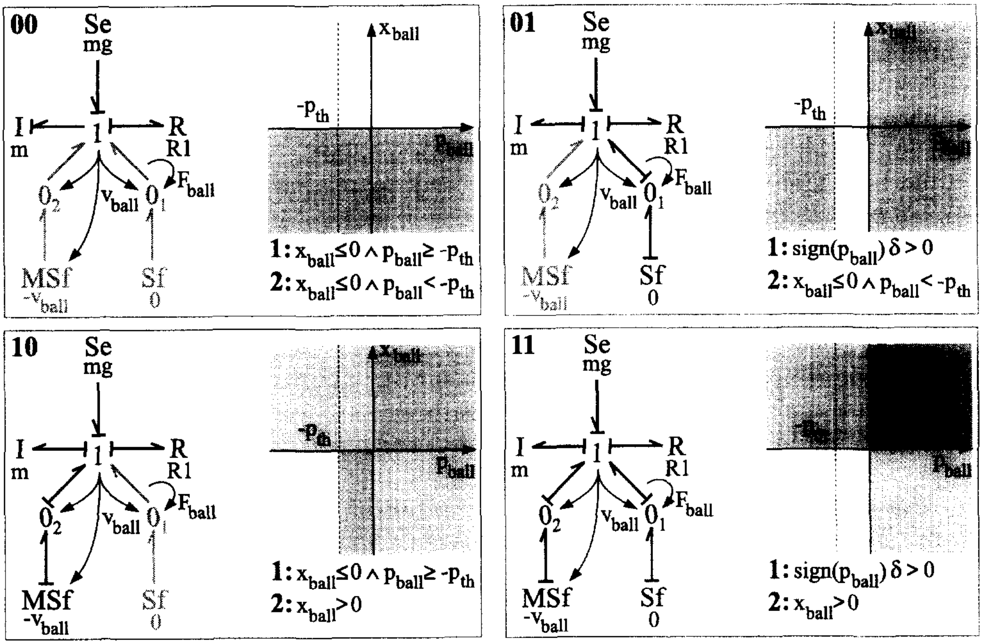


FIG. 19. Energy phase spaces for each of the modes of operation of a combined ideal elastic and nonelastic collision.

switches *on*, the velocity and momentum of the ball become 0 instantaneously. The constituent equation (6) implies that the value of F_m at this point is a Dirac pulse, δ, which approaches either positive or negative infinity, depending on whether the stored momentum was negative or positive, respectively. If the momentum was 0, F_m equals 0. Let the function *sign* be defined as

$$sign(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0. \\ 1 & \text{if } x > 0 \end{cases} \tag{7}$$

The condition for switching *off* immediately after switching *on* becomes $F_{\text{ball}} = -F_g - sign(p)\delta + F_{R1} \leqslant 0$. Because of the magnitude of the Dirac pulse, the effect of the gravitational force and the air resistance can be neglected at switching if $p \neq 0$. With the minus sign compensated, the condition for the controlled junction 1 to transition immediately from the *on* to *off* state is $sign(p)\delta \geqslant 0$. This inequality holds for all values of $p > 0$. If $p = 0$ then $F_m = 0$ and $F_{R1} = 0$, so the switching condition becomes $-F_g \leqslant 0$. Because of the negative value of F_g (the gravitational force acts downward), this condition is never satisfied and no further switching occurs. Consequently, the area for which $p > 0$ (modes 01 and 11) is grayed out in the phase space in Fig. 19.

To generate the complete analysis for this example, we note that the two controlled junctions are dependent, because a switch in one instantaneously affects variables that are used in transition conditions of the CSPEC of the other. Therefore, the two switches

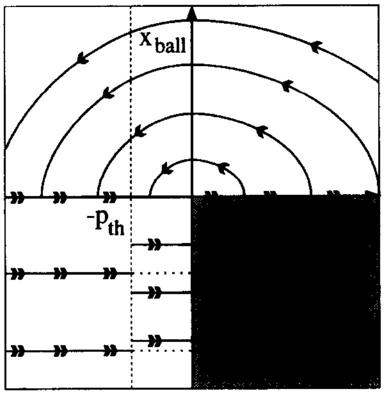


FIG. 20. Conjunction of the multiple energy phase spaces for each mode of operation of a combined ideal elastic and nonelastic collision. Double arrow heads depict discontinuous changes that occur after a new real mode is inferred.

have to be analyzed together for consistency of their combined effect. Phase spaces are established for each of the four modes of the combined elastic and nonelastic collision and labeled 00, 01, 10, and 11, where the first digit indicates whether the controlled junction 2 is *on* (1) or *off* (0), and the second digit indicates the same for controlled junction 1. The energy phase spaces for each of these modes of operation are shown in Fig. 19. In the phase spaces, the areas that are instantaneously departed are grayed out and the conjunction of the four energy phase spaces is shown in Fig. 20. This conjunction shows that there is an area in the energy distribution space for which the system model is in a transitional or switching state in every mode, and does not include a real mode of operation. Since the dimensions of the energy phase space are invariant across switches, this energy distribution cannot reach a real mode of operation during a sequence of switches in this region, thus violating the divergence of time condition.

In this part of the phase space, when the ball hits the floor, it has positive momentum. For the bouncing ball, this mode is unreachable, and, therefore, the model is physically consistent. The system always moves towards negative momentum and it instantaneously reverses (depicted by double arrows in Fig. 20) when the displacement becomes zero. So, analytically, the displacement never becomes negative. However, due to numerical disturbances, or initial conditions, the model may arrive in the physically inconsistent area of operation, especially, in case the floor is another moving body. In such situations, when simulating the system, the CSPEC conditions model the desired physical scenario inadequately.

To establish a physically correct system, the CSPEC switching conditions have to be modified. From the physical system it is clear that when the ball is moving upward but its position is below the floor, no collision should occur. So, additional constraints can be imposed based on the momentum of the ball, and the conditions $p_{\text{ball}} < 0$ and $p_{\text{ball}} \geqslant 0$ can be added to the *off/on* and *on/off* transitions of controlled junction 1, respectively. This results in the energy phase spaces shown in Fig. 21, and the conjunction of the energy phase spaces has a real mode of operation for each energy distribution. The combinational switching logic implies that this real mode of operation

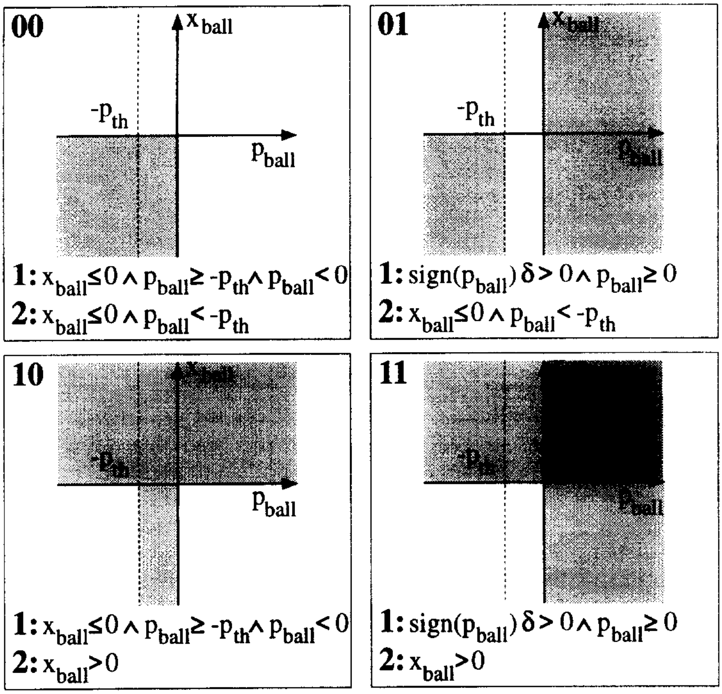


FIG. 21. Multiple energy phase space analysis of the modified model shows there is a real mode of operation for each energy distribution.

is reachable in one switching step. A simulation of the physically consistent system is shown in Fig. 22. The air resistance (R_l), causes the amplitude of the ball to decrease until the momentum of the ball upon collision falls below the threshold value p_{th} and the ball comes to rest on the floor.

To demonstrate the usefulness of the method for more complex systems, consider a valve mechanism used in automobiles (**45**, **46**), shown in Fig. 23. The valve is opened by a cam which pushes a rod over a specific angle of rotation, as illustrated in Fig. 24. As the rod is pushed upwards, the rocker arm serves as a lever to push the valve down. Though the valve spring exerts a force to maintain contact between the cam and rod tappet, it may lift off the cam and bounce back. This can be represented as a combined elastic/nonelastic collision. The hybrid bond graph model for the system is shown in Fig. 25. Notice that this model is very similar to the bouncing ball. However, in this situation, the value of the displacement, x_{th}, that causes contact between the cam and the rod tappet is variable and the momentum of the collision becomes $p_{col} = m(v_{cam} - v_{rod})$.

The energy phase space analysis of the hybrid bond graph model in Fig. 25 is shown in Fig. 26. Superimposition of the four spaces like before, shows the model to be physically consistent. A simulation of a revolution of the cam system is shown in Fig. 27. Initially, the cam and rod–tappet are continuously in contact and rotating the cam results in a vertical displacement of the tappet. When the valve mechanism reaches its maximal vertical displacement, the inertia of the valve-mechanisms causes the tappet

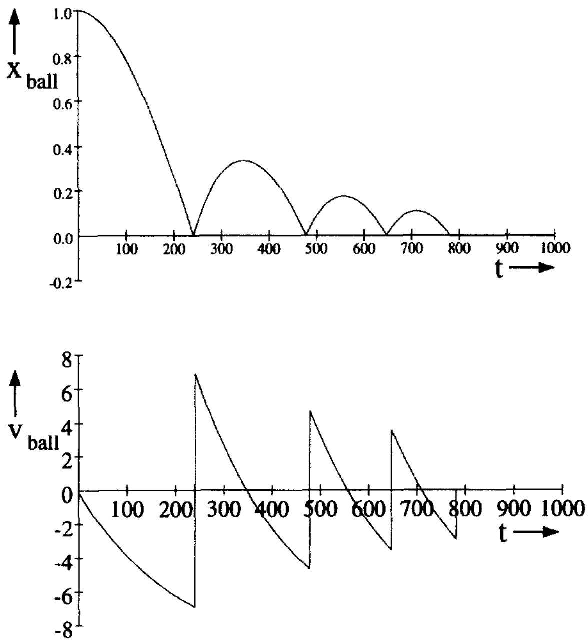


FIG. 22. A bouncing ball which comes to rest when its momentum falls below a specific threshold value ($m = 0.2$, $R_1 = 1$, $g = -9.8$).

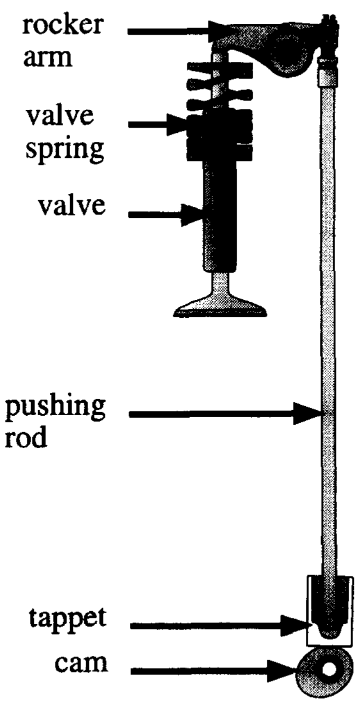


FIG. 23. Valve mechanism as used in automobiles.

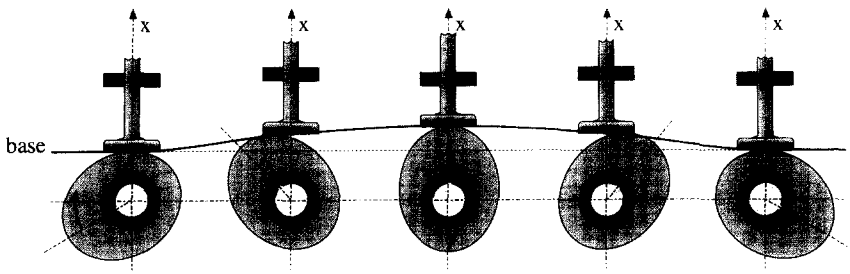


FIG. 24. Cam mechanism which opens a valve during one revolution.

to lift off the cam. Due to the valve spring, the tappet moves down and eventually collides with the cam. This collision is modeled as ideal elastic, until the tappet collides with momentum less than a certain threshold level, at which time the collision is modeled as ideal nonelastic and the tappet comes to rest on the cam again. Note that one of the phase space dimensions represents a source variable. Because sequences of switches occur instantaneously, external variables cannot change during such a sequence, therefore, external variables are invariant across a sequence of switches as well, and consequently they can be used in the energy phase space analysis. The dimension of the displacement variable x_{rod} is omitted for clarity reasons.

VI. Discussion and Conclusions

The inherent continuity of physical systems has been a key characteristic that has been exploited in effective modeling and simulation of these systems. However, it has often been found desirable to represent nonlinear behavior as discontinuities and piecewise linear behaviors in order to: (1) avoid model complexity, (2) prevent numerical stiffness, and (3) allow for linear system analysis methods. This results in mixed continuous/discrete system models that introduce a number of unique phenomena that are specific to their hybrid nature. In this paper, we develop a theory for hybrid physical

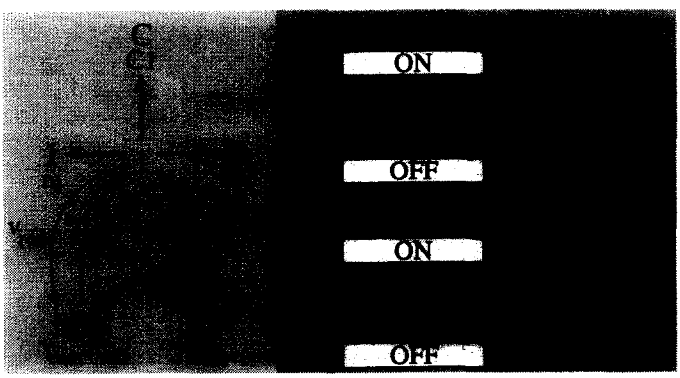


FIG. 25. Hybrid bond graph model of the cam–rod system.

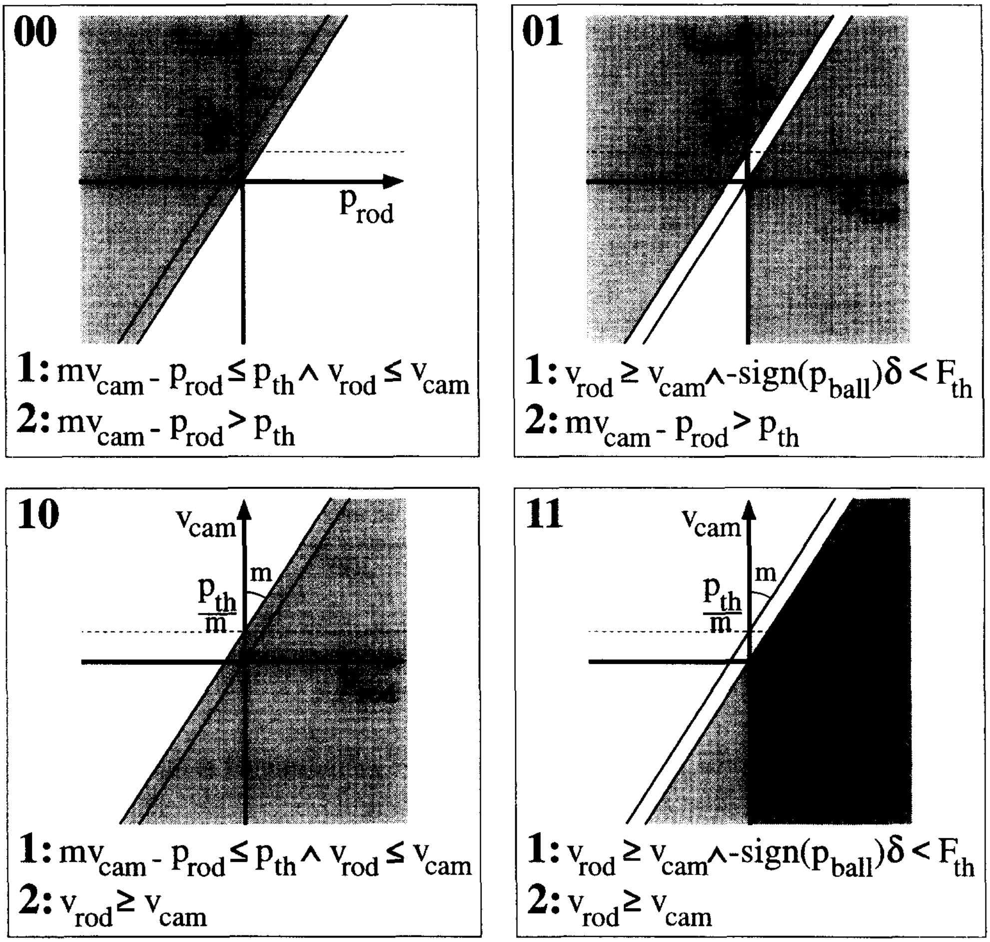
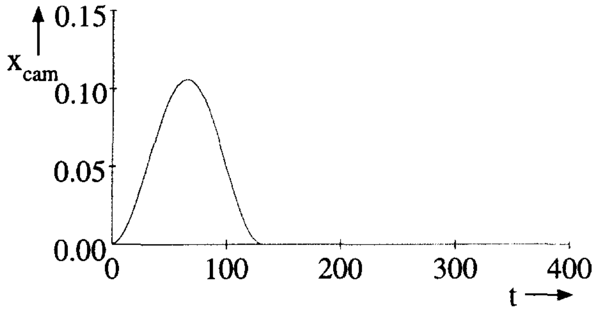


FIG. 26. Energy phase space analysis of cam–rod system.

system models that combine the use of the energetic bond graph modeling language with discrete finite state automata. Interesting observations are made about the lumped parameter assumption in the bond graph methodology, and its limitations in handling discrete changes in model configurations.

The concept of reticulation is employed in defining a basic switching element that is used in conjunction with the traditional 0- and 1-junctions in bond graph models. The switching element basically controls junctions in that it can activate and deactivate them. This introduces the basic notion of power discontinuity and configuration changes into models. Local switching logic is implemented by a CSPEC that is associated with each controlled junction. CSPECs take as input power variables from the bond graph, possibly operated on by block diagrams, and form a separate control model on top of the energy flow model of the bond graph. An important characteristic of our hybrid bond graph modeling scheme allows for global behavior generation based on



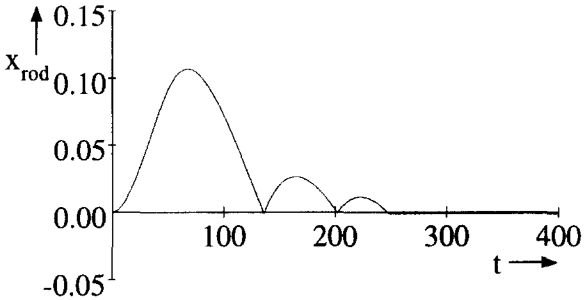


FIG. 27. Simulation of cam boundary position and rod position during one cam revolution [$m = 1.5$, $C_1 = 0.01$, $R_1 = 10$, $g = -9.8$, $v_{rod}(0) = 0$, $F_{C1}(0) = 0.1$].

local switching conditions, a necessary condition for compositional modeling of complex systems.

Conditions of physical consistency are exploited to develop a behavior generation algorithm which correctly derives a new model configuration where system behavior evolves as a continuous function of time based on differential equations. The algorithm correctly transfers the continuous state vector between modes of operation, i.e. discontinuous model changes, in the hybrid bond graph framework. A multiple energy phase space analysis technique provides a systematic methodology for model verification in terms of the principle of divergence of time. This analysis also reveals that, though there may be energy distributions where a mode of continuous evolution does not exist, these modes may not be reachable, and the system is physically consistent for all other energy distributions.

Unlike previous work [e.g. Alur, Henzinger and Nicollin (**15**, **16**)] which required linearity assumptions, i.e. state variables had to have a constant rate of change for divergence of time to be provable, the multiple energy phase space model is more general. Furthermore, this technique provides a lot of additional useful information, in the sense that it can be used to identify possible physical inconsistencies, which suggests modifications to switching conditions to eradicate these. This aids in developing physically correct models, which is different from the approach of Iwasaki *et al.* (**18**), where the concept of hypertime is introduced to make a system model executable.

Because of the exponential time complexity of the multiple energy phase space analysis (all possible combinations of switches have to be analyzed), our current research focuses on partitioning a physical system model into areas of instantaneous interrelated switching effects. These represent topological areas in the model, where switches are instantaneously affected by one another through related signals. Preliminary work in this area is presented in (**47**).

References

(**1**) D. G. Luenberger "Introduction to Dynamic Systems: Theory, Models, & Applications", Wiley, New York, 1979.

(**2**) D. Weld and J. de Kleer, "Readings in Qualitative Reasoning about Physical Systems", Morgan Kaufmann, San Mateo, CA, 1990.

(**3**) D. C. Karnopp, D. L. Margolis and R. C. Rosenberg, "Systems Dynamics: A Unified Approach", 2nd Edn, Wiley, New York, 1990.

(**4**) H. M. Paynter, "Analysis and Design of Engineering Systems", The M.I.T. Press, Cambridge, MA, 1961.

(**5**) P. J. Mosterman and G. Biswas, "Modeling discontinuous behavior with hybrid bond graphs", in "1995 International Conference on Qualitative Reasoning", pp. 139–147, Amsterdam, May 1995.

(**6**) T. Nishida and S. Doshita, "Reasoning about discontinuous change", in "Proc. AAAI-87", pp. 643–648, Seattle, WA, 1987.

(**7**) P. C. Breedveld, "Multibond graph elements in physical systems theory", *J. Franklin Inst.*, Vol. 319(1/2), pp. 1–36, Jan./Feb. 1985.

(**8**) K. Denbigh, "The Principles of Chemical Equilibrium: With Applications in Chemistry and Chemical Engineering", 3rd Edn, Cambridge University Press, London, 1971.

(**9**) G. M. Asher, "The robust modelling of variable topology circuits using bond graphs", in "Proceedings of the International Conference on Bond Graph Modeling", pp. 126–131, San Diego, CA, 1993.

(**10**) W. Borutzky, "Representing discontinuities by sinks of fixed causality", in "1995 International Conference on Bond Graph Modeling and Simulation (ICBGM '95)" (Edited by F. E. Cellier and J. J. Granada), pp. 65–72, Las Vegas, Jan. 1995.

(**11**) J. F. Broenink and K. C. J. Wijbrans, "Describing discontinuities in bond graphs", in "Proceedings of the International Conference on Bond Graph Modeling", pp. 120–125, San Diego, CA, 1993.

(**12**) J. E. Strömberg, J. Top and U. Söderman, "Variable causality in bond graphs caused by discrete effects", in "Proceedings of the International Conference on Bond Graph Modeling", pp. 115–119, San Diego, CA, 1993.

(**13**) W. Borutzky, "Discontinuities in a bond graph framework", *J. Franklin Inst.*, Vol. 332B, pp. 141–154, 1995.

(**14**) P. J. Mosterman and G. Biswas, "Behavior generation using model switching a hybrid bond graph modeling technique", in "1995 International Conference on Bond Graph Modeling and Simulation (ICBGM '95)", (Edited by F. A. Cellier and J. J. Granada), pp. 177–182, Las Vegas, Jan, 1995.

(**15**) R. Alur, C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P.-H. Ho, X. Nicollin, A. Olivero, J. Sifakis and S. Yovine, "The algorithmic analysis of hybrid systems", in "Proceedings of the 11th International Conference on Analysis and Optimization of Discrete Event Systems" (Edited by J. W. Bakkers, C. Huizing, W. P. de Roeres, and G. Rozenberg), pp. 331–351, Springer, 1994.

(**16**) T. A. Henzinger, X. Nicollin, J. Sifakis and S. Yovine, "Symbolic model checking for real-time systems", *Information Comput.*, Vol. 111, pp. 193–244, 1994.

(**17**) X. Nicollin, J. Sifakis and S. Yovine, "From atp to timed graphs and hybrid systems", in

"Lecture Notes in Computer Science" (Edited by J. W. Bakkers, C. Huizing, W. P. de Roeres and G. Rozenberg), Vol. 600, pp. 549–571, Mook, The Netherlands, June 1991.

(18) Y. Iwasaki, A. Farquhar, V. Saraswat, D. Bobrow and V. Gupta, "Modeling time in hybrid systems: how fast is 'instantaneous'?", in "1995 International Conference on Qualitative Reasoning", pp. 94–103, Amsterdam, May 1995.

(19) R. C. Rosenberg and D. Karnopp, "Introduction to Physical System Dynamics", McGraw-Hill, New York, 1983.

(20) J. U. Thoma, "Bond graph for thermal energy transport and entropy flow", *J. Franklin Inst.*, Vol. 292(2), pp. 109–120, 1971.

(21) G. Falk and W. Ruppel, "Energie und Entropie: Eine Einführung in die Thermodynamik", Springer, Berlin, 1976.

(22) R. Penrose, "The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics", Oxford University Press, Oxford, 1989.

(23) D. Karnopp and R. C. Rosenberg, "System Dynamics: A Unified Approach", The Massachusetts Institute of Technology, Cambridge, MA, 1968.

(24) P. J. Mosterman and G. Biswas, "Analyzing discontinuities in physical system models", in "Qualitative Reasoning: The Tenth International Workshop" (Edited by Y. Iwasaki and A. Farquhar), pp. 164–173, Stanford Sierra Camp, Fallen Leaf Lake, CA, May 1996; AAAI Technical Report WS-96-01, AAAI Press, 1996.

(25) P. J. Mosterman and G. Biswas, "A formal hybrid modeling scheme for handling discontinuities in physical system models", in "Proc. AAAI-96", pp. 985–990, Portland, OR, 1996.

(26) J. de Kleer and J. S. Brown, "A qualitative physics based on confluences", *Artificial Intelligence*, Vol. 24, pp. 7–83, 1984.

(27) P. J. Mosterman and G. Biswas, "Hybrid modeling specifications for dynamic physical systems", in "Int. Conf. on Bond Graph Modeling and Simulation", pp. 162–167, Phoenix, AZ, Jan. 1997.

(28) B. Kuipers, "Qualitative simulation", *Artificial Intelligence*, Vol. 29, pp. 289–338, 1986.

(29) D. Karnopp and R. C. Rosenberg, "Analysis and Simulation of Multiport Systems", Wiley, New York, 1975.

(30) J. Garcia, G. Dauphin-Tanguy and C. Rombaut, "Bond graph modeling of thermal effects in switching devices", in "1995 International Conference on Bond Graph Modeling and Simulation (ICBGM '95)" (Edited by F. E. Cellier and J. J. Granada), pp. 145–150, Las Vegas, Jan. 1995.

(31) W. Borutzky, G. Dauphin-Tanguy and J. U. Thoma, "Advances in bond graph modelling: theory, software, applications", *Math. Comput. Simulation*, Vol. 39, pp. 465–475, 1995.

(32) J. E. Strömberg, U. Söderman and J. L. Top, "Bond graph supported estimation of discretely changing parameters", "Proceedings of the International Conference on Bond Graph Modeling", pp. 47–52, San Diego, CA, 1993.

(33) J. Top, "Conceptual modeling of physical systems", Ph.D. dissertation, University of Twente, CIP-DATA Koninklijke Bibliotheek, Den Haag, The Netherlands, 1993.

(34) U. Söderman and J. E. Strömberg, "Switched bond graphs: towards systematic composition of computational models", in "1995 International Conference on Bond Graph Modeling and Simulation (ICBGM '95)" (Edited by F. E. Cellier and J. J. Granada), pp. 73–79, Las Vegas, Jan. 1995.

(35) T. DeMarco, "Structured Analysis and System Specification", Prentice-Hall, Englewood Cliffs, NJ, 1979.

(36) D. J. Hatley and I. Pirbhai, "Strategies for Real-Time Systems Specification", Dorset House, New York, 1988.

(37) P. T. Ward and S. J. Mellor, "Structured Development for Real-Time Systems", Prentice-Hall, Englewood Cliffs, NJ, 1985.

(38) F. W. Umbach, "Organisatie van Instrumentatie", H. E. Stenfert Kroese B.V., Leiden/Antwerpen, 1981.

(39) A. V. Aho, J. E. Hopcroft and J. D. Ullman, "The Design and Analysis of Computer Algorithms", Addison-Wesley, Reading, MA, 1974.

(40) Z. Kohavi, "Switching and Finite Automata Theory", McGraw-Hill, New York, 1978.

(41) G. Biswas and X. Yu, "A formal modeling scheme for continuous systems: focus on diagnosis", in "Proceedings IJCAI-93", pp. 1474–1479, Chambery, France, Aug. 1993.

(42) B. Falkenhainer and K. Forbus, "Compositional modeling: finding the right model for the job", *Artificial Intelligence*, Vol. 51, pp. 95–143, 1991.

(43) J. F. Broenink, "Computer aided physical modeling and simulation: a bond graph approach", Ph.D. dissertation, University of Twente, Febodruk, Enschede, The Netherlands, 1990.

(44) Microsoft Corporation, "Visual Basic, Version 3.0 Professional Edition", One Microsoft Way, Redmond, WA 98052-6399, 1993.

(45) C. W. Ham, E. J. Crane and W. L. Rogers, "Mechanics of Machinery", 4th Edn, McGraw-Hill Book Company, New York, 1958.

(46) E. F. Obert, "Internal Combustion Engines: Analysis and Practice", 2nd Edn, International Textbook Company, Scranton, PA, 1950.

(47) P. J. Mosterman and G. Biswas, "Verification of hybrid physical system models", in "Proc. ASME-96", pp. 707–714, Atlanta, GA, Nov. 1996.