Hybrid Modeling Specifications for Dynamic Physical Systems

Pieter J. Mosterman and **Gautam Biswas**
Center for Intelligent Systems
Box 1679, Sta B
Vanderbilt University
Nashville, Tennessee 37235
`pjm, biswas@vuse.vanderbilt.edu`

Abstract

Non linear behaviors of real-world physical systems are often abstracted into piecewise linear models by simplifying component parameters or coarsening the time scale of behavior analysis. The net result is often abrupt, discontinuous changes in system behavior, which typically violate the principles of continuity of power and conservation of energy. We have developed a hybrid modeling paradigm that combines energetic bond graph models with finite state automata for discrete meta-level control of model configuration changes. This provides a systematic framework for behavior generation based on the principle of *invariance of state*. The principle of *divergence of time* verifies consistency of models using phase space analysis.

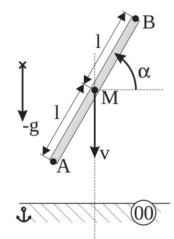

Figure 1: **A collision between a body and a floor.**

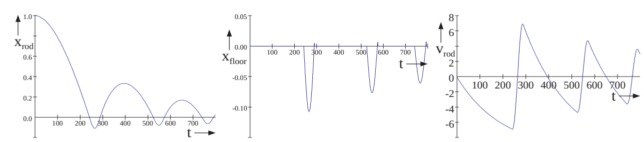

Figure 2: **An ideal elastic collision between a rigid body and a stiff floor.**

1 Introduction

All physical system behavior is governed by the principle of conservation of energy and continuity of power [3]. Simulators for generating dynamic physical system behavior strictly adhere to these principles. However, nonlinear continuous behaviors are often simplified in simulation for computational purposes to piecewise linear behaviors, and the resulting modeling *abstractions* make the system behavior appear discontinuous. These abstractions are typically categorized as *time scale* abstractions and phenomena or *parameter* abstractions [10].

Consider a thin rod falling freely under gravity at a prespecified angle, Fig. 1. Assuming an *ideal elastic collision* with the floor the rod tip velocity reverses at the point of impact. In reality, this process takes a small amount of time during which the kinetic energy of the rod is transformed into internal elastic energy in the rod and the floor (they are both compressed), and this builds up a reaction force that reverses the rod-tip velocity. This behavior is illustrated in Fig. 2 for an ideal rod–floor collision (i.e., all kinetic energy is returned), with the floor assigned a very high stiffness value. The process of force and velocity reversal occurs at a much smaller time scale than the time scale at which the overall behavior of interest evolves, therefore, it makes sense to abstract the ideal elastic collision in time and the reversal of velocity phenomena appears to be instantaneous.

Parameter abstraction occurs if the collision is considered to be ideal *non-elastic*; when the rod collides with the floor, its tip remains in contact with the floor. This can only happen if the rod and floor are both considered ideal rigid bodies and their coefficients of elasticity parameters are abstracted away.

Time scale and parameter abstractions result in continuous models that are interspersed with time points where discontinuities occur. These discontinuities are best described by model configuration changes, and the system is said to operate in *multiple modes* during which behavior evolves continuously. Model configuration changes are determined by a meta-level control model which determines the active mode of operation at the end of a sequence of discrete changes. Previous research has applied pre-enumeration of realizable model configurations to implement a *global* control model using transition functions [7] and Petri net structures [4].

In general, pre-enumeration of all realizable model configurations is a daunting task which can be overcome by introducing local *switching* mechanisms [1]. These mechanisms dynamically determine which model fragments are active and compose them together to define the global mode of operation. However, the introduction of compositionality requires more complex mode transition semantics because local activation of model fragments may cause a number of consecutive discrete changes before a new mode of continuous operation is arrived at. These mixed continuous and discrete systems define a *hybrid* modeling paradigm. During discontinuous change, physical principles of conservation of energy and continuity of power may be violated and the interaction between the continuous formalism and localized discrete effects is governed by the principle of *invariance of state* [10], which requires the state vector to be invariant across discrete changes.

If a number of consecutive discrete changes triggers a previous change in the sequence, a loop of discrete, instantaneous, changes may ensue. Because of the assumption that discontinuities are instantaneous, real time would halt, which is in conflict with the observation that physical system behavior always progresses or diverges in time. The physical principle of *divergence of time* is used in energy phase space analysis to *verify* physical consistency of hybrid models [10].

This paper presents a systematic study of hybrid models for dynamic physical systems. The modeling paradigm combines the use of bond graphs to capture the continuous evolution of system behavior with local finite state automata to model discrete changes in system model configuration.

2 Model Requirements

The concept of *reticulation* applies to physical system models, and allows a system to be defined as a configuration of ideal processes. When this concept is extended to allow discrete changes, an ideal element, the switch, that implements configuration changes without directly participating in energy interactions within the system is required.

Consider the falling rod example discussed earlier, and assume that the rod and floor are ideal rigid bodies, therefore, their collision is ideal non-elastic. Initially the rod falls freely, therefore, it has linear momentum in the vertical direction only. Upon collision, its vertical momentum is distributed into three components: horizontal, vertical, and angular momentum, as a function of the length of the rod and angle of collision. Coulomb friction acts when the rod and floor are in contact: $if\ v \neq 0\ then\ |F_{friction}| = \mu F_{normal}$, where μ is the coefficient of friction (this example is adapted from [8]). If the length and angle are such that the magnitude of the horizontal force, $|F_{x,collision}|$, exceeds the thresh-

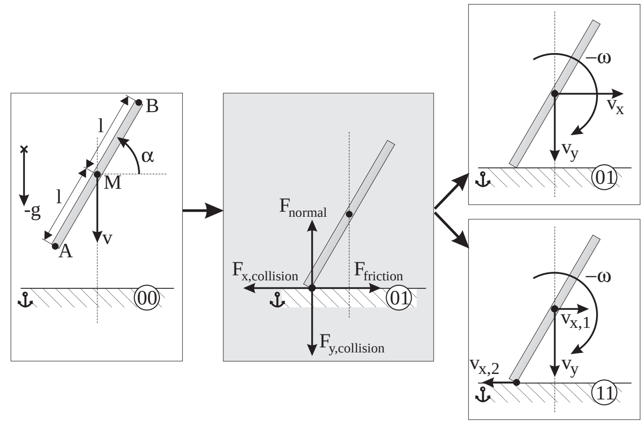

Figure 3: **Possible modes of operation of a thin rigid rod falling onto a rigid floor.**

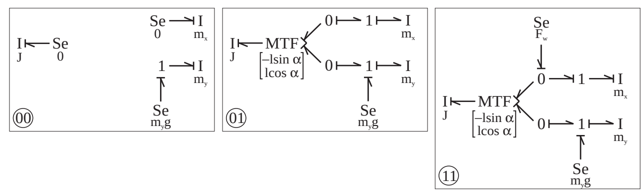

Figure 4: **Model configurations of the falling rod.**

old value μF_{normal}, then the rod starts to slide (Fig. 3). Otherwise, it sticks at the point of collision and rotates about this point. For ease of analysis, the overall behavior can be split into three different modes: (1) rod falls freely, (2) rod sticks at the point of collision and rotates, and (3) rod slides after collision and rotates.

Bond graph models of rigid bodies (van Dijk [13]) describe individual models for each mode of operation. When the rod is falling freely no forces act on the rotational and horizontal inertias, and gravity acts on the vertical inertia. This is mode 00 in Fig. 3, and the corresponding bond graph appears in Fig. 4. In the second model configuration (01), the rod rotates around the point of contact, A, and its angular velocity determines the horizontal and vertical velocity of the center of mass, M, based on the angle of the rod, α, and distance l between A and M. This model (Fig. 4) also shows the effects of gravity on the three inertias, creating accelerated rotation. Note that there is no linear motion along the floor, because the horizontal force applied by the rod does not exceed the threshold required to overcome the force of friction. In the third model configuration, mode 11, the force of friction is exceeded, therefore, the rod has linear motion along the floor. Rotation completely determines the vertical velocity of the center of mass because the point of rotation is fixed in this direction. A friction force, F_w, works against the horizontal velocity at the point of contact which adds to the velocity of the center of mass.

2.1 Transferring the State

When the system changes to a new mode of operation, the system state has to be derived from the last continuous mode of operation. This is referred to as the *initial value problem.* If no buffer or energy storage elements go into derivative causality, the state vector is unchanged. Otherwise, (1) one or more buffers may become dependent on a source, or (2) two or more buffer elements become dependent on each other, and the state vector between the two configurations is different [10, 11]. In the first case, the energy stored in the dependent buffers is determined by the value of the source. In the second case, conservation of state is applied so the total amount of charge and momentum in both modes remains the same. For n dependent buffers, buffer 0 is chosen in integral causality. The new value of its stored energy, q_0^+, is determined by

$$q_0^+(1 - \sum_{i=1}^{n-1} g_{i,0} g_{0,i} \frac{C_i}{C_0}) = q_0 + \sum_{i=1}^{n-1} g_{0,i} q_i \qquad (1)$$

where $g_{i,0}$ is the gain of the path from buffer i to buffer 0 and C_i the buffer value of buffer i. Note that this may result in loss of energy to the environment [11].

As an example of buffer-buffer dependency, consider the situation where the falling rod hits the floor and starts to rotate around the point of contact. The bond graph model of this mode, 01 (Fig 4), shows that all three inertias are dependent, their linear velocities are completely determined by the angular velocity. To continue behavior generation, the state vector for the new mode 01 has to be computed from the previous mode 00, where the rod falls freely. This requires the initial momentum to be redistributed among the three inertias to establish consistent linear and angular velocities. Following this procedure, ensures the conservation of momentum across the configuration changes.

2.2 Sequences of Discrete Changes

Local switches can trigger other switching conditions, resulting in sequences of mode transitions. This is illustrated by the scenario where the rod starts to slide after hitting the floor. Upon collision, the first model configuration that is generated implies all three inertias are dependent. This results in new signal values (forces) which may cause the sliding mode of operation to become active if $|F_x| > \mu F_{normal}$. Whether this model configuration becomes active can only be inferred by traversing the intermediate configuration of three dependent inertias. Note, however, that this mode has no real representation but is only used for inferring the new mode of continuous operation. These modes are called *mythical.*

2.3 Invariance of State

In previous work [10], we have shown that the state vector adheres to the principle of *invariance of state*, which states that the state vector in the last mode of continuous operations is preserved until a new continuous mode is arrived at. This is illustrated for the case where the falling rod slides upon hitting the floor given $\mu = 0$. At the point of contact, the system switches from mode 00 to mode 01 (Fig. 4) first and since the rod starts to slide, $|F_x| > \mu F_{normal}$, mode 11 next. This immediate transition implies that the intermediate mode, 01, is mythical, which is crucial in solving the initial value problem. The implication is that the state vector in mode 11 should be directly computed from the state vector in mode 00. In this case, the linear momentum p_x has an initial value of 0 as its buffer element, m_x, is independent of the other two inertias in mode 11. If mode 01 were real, the state vector would first be transferred to mode 01, the dependency of the three inertias would result in p_x to change to an abrupt non zero value, and, therefore, the rod would have a non zero horizontal velocity at the point it started sliding. This behavior is incorrect, because there is no horizontal force acting ($\mu = 0$), the horizontal velocity should remain 0.

3 Evaluation of Existing Methodologies

Bond graphs form an elegant formalism to model the continuous dynamic behavior of physical systems. Initial work in introducing local switching mechanisms focused on nonlinear resistances and modulated transformers (MTF) [5]. However, the use of a nonlinear resistance introduces dissipative effects and violates the requirements for ideal switching. Other disadvantages of dissipative elements are the increase in model complexity, and because of their typically small resistive values they may introduce numerical stiffness in simulators. Finally, the occurrence of algebraic loops [13] increases when additional resistances are introduced. Modulated transformers with a boolean modulus avoid these problems but always propagate causality which is not correct.

The causality problem of $MTFs$ was solved by using a combination of an MTF and resistor with fixed causality [2]. This construct allows for configuration changes without re-assigning causality but again introduces dissipation into the switching process. The other disadvantages listed for nonlinear resistances also appear.

To ensure ideal switching processes and to eliminate meaningless causal assignments when changes in model configurations occur, Strömberg, Top, and Söderman [12] introduced an ideal *switch* as a bond graph element, and Broenink and Wijbrans [4] introduced the notion of *switching bonds* to handle discontinuities in models. The switch in *switched bond graphs* relies on a localized bond graph element to model discontinuities, whereas the switching bonds approach recognizes the global implications of local switching effects and incorporates a higher level control mechanism to establish configuration changes. However, deactivating power bonds may cause problems in the form of hanging junctions and incor-

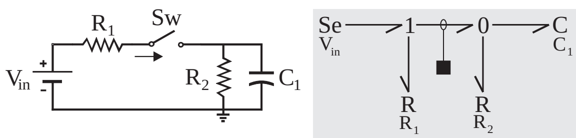

Figure 5: **Switching bonds.**

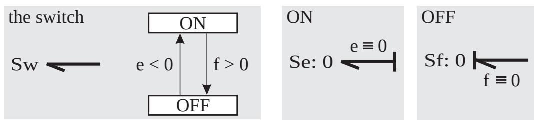

Figure 6: **Operation of the switching element.**

rect loading of diconnected model fragments. Further, changing boundary conditions due to switching are incorrectly handled. For example, in Fig. 5 if the switch opens, the corresponding bond in the bond graph disappears. Though this disconnects the parallel part, the bond graph incorrectly still shows a series connection.

Switched bonds introduce discontinuities as a bond graph element, Sw, that enforces 0 effort or 0 flow on a junction (Fig. 6), thus inhibiting energy transfer in one state and allowing completely unobstructed flow of energy in the other. Whereas these switches satisfy the requirement of an ideal transitional element for meta-level control, they clutter bond graph models and obscure hierarchical structure descriptions.

Mosterman and Biswas address the above problem by using *controlled junctions* in a *hybrid bond graph* framework [9]. This modeling formalism is based on a physical theory of discontinuities in models of dynamic physical systems [10] and achieves (1) ideal switching by a controlled junction, and (2) global configuration changes from local switches by a control algorithm.

4 Hybrid Bond Graph Modeling and Simulation

The hybrid modeling formalism relies on bond graphs to model continuous system behavior and *finite state automata* [6] to model discrete configuration changes.

4.1 Controlled Junctions

The finite state automata take as input signal values from the bond graph model, and generate control signal values that cause specific controlled junctions to be turned *on* and *off*. These controlled junctions act like normal junctions when they are in their *on* state. When turned *off*, they inhibit transfer of energy between model fragments that are connected through the junction. To correctly handle boundary conditions, it is critical to load disconnected model fragments appropriately. To model ideal switch behavior [12], when a 0-junction is turned *off*, it is replaced by 0 value effort sources on each

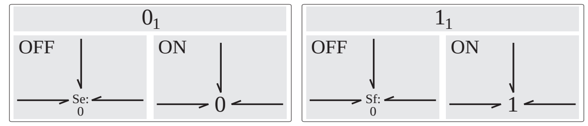

Figure 7: **Operation of the controlled junction.**

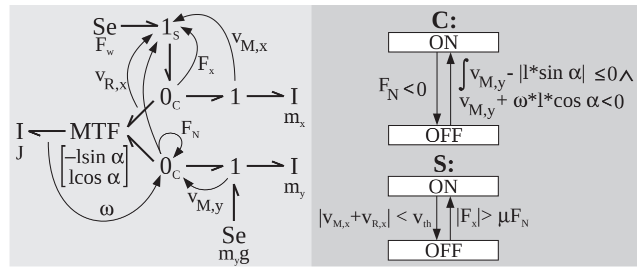

Figure 8: **Hybrid bond graph model of a thin rigid rod colliding with a rigid floor.**

of the adjoining bonds. When a 1-junction is turned *off*, it is replaced by 0 value flow sources (see Fig. 7).

The input to the controlling finite state automata are shown as signals into a controlled junction (Fig. 8) which is marked with a subscript that links the junction to a finite state automaton. The finite state automata represent a junction's *control specification* (CSPEC) which may contain sequential logic with any number of internal states. However, each one of these has to map onto either a junction's *on* or *off* state and, since each state reflects a physical manifestation, *on* and *off* states have to alternate in each transition sequence. Logically all conditions on state transitions have to evaluate to a boolean value. Finally, CSPEC conditions have to be such that every sequence of discrete transitions ends in a new model configuration that has a real manifestation.

The controlled junctions combined with their control logic constitute part of the control model of the hybrid modeling paradigm which is distinct from the energy model since it performs no energy transfer. Note that each combination of *on/off* states of all controlled junctions constitutes a global *mode* of operation. These modes can be dynamically generated.

As an example consider the hybrid bond graph model, shown in Fig. 8, of the idealized rigid thin rod colliding with the idealized rigid floor (Fig. 1). Initially, the rod is moving freely and controlled junctions 0_C and 1_S are *off*. Replacing the junctions with their 0 value sources results in the bond graph (mode 00) shown in Fig. 9. The position of the rod-tip closest to the floor, y_A, is determined by the position of the center-point, $\int v_{M,y}$, and the distance of the rod-tip from the center point, $|l * sin\alpha|$. If y_A becomes 0, the rod collides with the floor and 0_C comes *on*, and the model transitions into mode 01. If the rod-length and angle of collision are such that 1_S comes *on*,

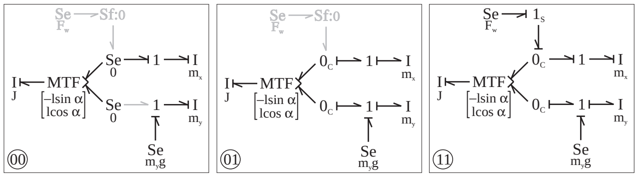

Figure 9: **Dynamically generated configurations.**

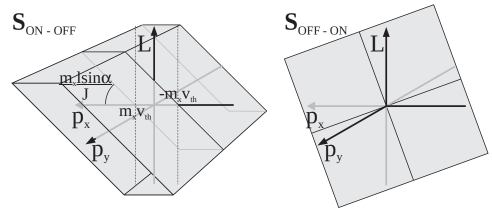

Figure 10: **Transition conditions for S.**

the model transitions into mode 11 and the rod begins to slide. Note that removing the 0 value sources that have no effect results in the model configurations as determined based on a global pre-enumeration approach, Fig. 4. This shows how the hybrid bond graph approach provides a seamless integration of configuration changes based on local switches. Other examples of hybrid bond graph models are discussed in [9, 10, 11].

4.2 The Switching Algorithm

A local switch may trigger a sequence of configuration changes, which ends when a new configuration where the system evolves continuously is arrived at. Initially the system operates in mode σ_0. At time t_s a signal value crosses a threshold value which invokes a local switch, establishing mode σ_1 as the new model configuration. Now, the system state, $x(t_s)$, is transferred to this configuration and re-calculation of signal values invokes one or more new local switches and the model moves to mode σ_2 instantaneously. The initial state vector, $x(t_s)$, is used again to infer possible other configuration changes until a mode, σ_m is arrived at where no more instantaneous changes occur, based on $x(t_s)$. This mode is established as the new real mode at t_s with state $x^+(t_s)$. As a result, the system transfers from a real mode, σ_0, to another, σ_m, which follow each other immediately in real time. Details of the implementation of this simulator system are presented elsewhere [9].

5 Verifying Model Behaviors

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 [11].

To verify divergence of time for the falling rod in Fig. 3, the switching conditions for all CSPECs have to be expressed in terms of the stored energy in the system. The resulting energy phase space is 6-dimensional: $(\alpha, \mu, y, p_x, p_y, L)$ across all modes: 00, 01, 10, and 11. To establish consistency of this model, the intersection of the transition conditions for all modes would be solved for simultaneously to determine if there was a non empty transitional area. For sake of clarity and to visualize the energy phase space, we limit the study to CSPEC **S** in modes 01 and 11 in the three dimensional energy phase space with dimensions (p_x, p_y, L).

The condition under which the CSPEC moves to its *off* state is recomputed in terms of energy variables, and the condition $|v_{M,x} + v_{R,x}| < v_{th}$ is rewritten in terms of the constituent variables in mode 11. With $v_{M,x}$ the horizontal velocity of the center of mass, $v_{R,x}$ the horizontal velocity as a result from the angular velocity, and v_{th} a preset threshold velocity, this yields $\frac{m_x l sin\alpha}{J} L - m_x v_{th} < p_x < \frac{m_x l sin\alpha}{J} L + m_x v_{th}$ (Fig. 10). To translate the condition under which CSPEC **S** switches *on*, $|F_x| > \mu F_N$, into its constituent energy variables, it is observed that in mode 01, F_x and F_y have derivative relations, and the switching conditions becomes $|\frac{dp_x}{dt}| > \mu(\frac{dp_y}{dt} - mg)$. At the time of collision the linear momenta change discontinuously, generating Dirac pulses, thus the effect of mg can be neglected. A comparison of the two Dirac pulse areas results in the switching condition $|p_{x,1} - p_{x,0}| > \mu(p_{y,1} - p_{y,0})$ with $p_{x,0}$ the momentum before and $p_{x,1}$ the momentum after collision. Consider the situation where $p_{x,0} = 0$ and $p_{x,1} > 0$, mode 01 in Fig 3. Then switching occurs if $p_{x,1} - p_{x,0} > \mu(p_{y,1} - p_{y,0})$, where $p_{x,1}$ and $p_{y,1}$ can be expressed in terms of $p_{x,0}$, $p_{y,0}$, and L_0 which constitute the total amount of momentum present in the system before collision. Substitution establishes an inequality of the form $c_1(\mu, \alpha, l)p_{x,0} + c_2(\mu, \alpha, l)p_{y,0} + c_3(\alpha, \mu, l)L_0 > 0$.

These switching conditions are shown in Fig. 10 as a plane with normal vector $[c_1(\mu, \alpha, l)\ c_2(\mu, \alpha, l)\ c_3(\alpha, \mu, l)]^T$, which establishes the boundary between the transitional and non-transitional space. Because the normal vector is a function of α, depending on its direction any point in the energy space may satisfy the transition condition. However, the plane always contains the origin which is also true for the transition space of $\mathbf{S}_{ON-OFF}$. Therefore, if both spaces are intersected for $v_{th} > 0$, there is a space where the energy distribution along (p_x, p_y, L) causes **S** to switch *on* as well as *off*. In this space the model cannot reach a new mode of continuous evolution and divergence of time is violated. In this area, when the rod makes contact with the floor, transition conditions imply that it begins to slide. However, when it does, the velocity at point A is below v_{th} so the rod stops sliding and sticks. But, this is the initial model configuration when the rod makes contact with the floor, for which the forces were determined to make the rod slide, therefore, the model goes into a loop of discontinuous changes.

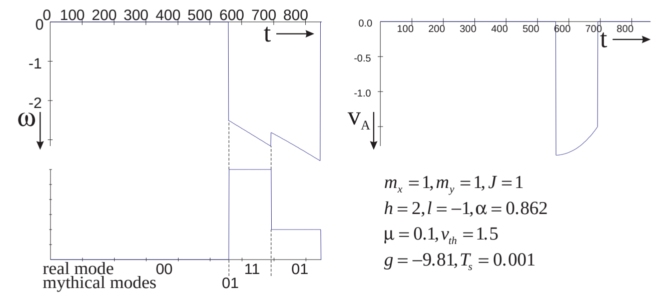

Figure 11: **Physically consistent simulation.**

To eliminate this modeling inconsistency, either (1) the threshold velocity can be set to 0, (2) additional constraints can be included for CSPEC **S**, or (3) the system can be modeled in greater detail adding parasitic phenomena so that the model exhibits continuous behavior. In case additional constraints are included for CSPEC **S**, either the *on→off* condition or the *off→on* condition can be modified. To make the model physically consistent, the *off→on* transition condition can include $|v_{M,x} + v_{R,x}| \geq v_{th}$ Simulating this model for one scenario after adding this constraint is shown in Fig. 11. After an initial stage where the rod falls down at a specific angle, it hits the floor and moves into configuration 01. Given the state vector of the system, it is inferred that an immediate configuration change to mode 11 follows. In this configuration the rod slides with a velocity that is decreasing in magnitude due to the friction force acting on the rod-tip. At one time, this velocity falls below a preset threshold value and if the state vector is such that the rod can get stuck without immediately satisfying the condition to slide, the system moves into mode 01. In this mode it is stuck and rotates around the point of contact until it falls flat on the floor.

6 Conclusions

Physical systems are by nature continuous and perceived discontinuities are caused by abstracting nonlinear continuous behaviors, to reduce model complexity and avoid numerical stiffness in simulation. From a compositional modeling viewpoint, modeling discontinuous configuration changes locally makes the task of model generation feasible and efficient. The resultant hybrid modeling scheme (1) incorporates ideal switching mechanisms that do not interfere with the energetic properties of the system, (2) transfers the system state correctly between configurations, and (3) handles sequences of configuration changes, to generate global system behavior.

The notion of a meta-level control model is applied by Mosterman and Biswas [9] in combination with the controlled junction to provide for consistent interaction between control model and energy model. The application of the global mythical mode algorithm in conjunction with continuous simulation to derive system behavior is demonstrated for the falling rod system. Divergence of time analysis verifies model consistency.

References

[1] G. Biswas and X. Yu. A formal modeling scheme for continuous systems: Focus on diagnosis. In *Proc. IJCAI-93*, pp. 1474–1479, Chambery, France, Aug. 1993.

[2] W. Borutzky, G. Dauphin-Tanguy, and J.U. Thoma. Advances in bond graph modelling: theory, software, applications. *Mathematics and Computers in Simulation*, 39:465–475, 1995.

[3] P.C. Breedveld. Multibond graph elements in physical systems theory. *J. of the Franklin Inst.*, 319(1/2):1–36, Jan./Feb. 1985.

[4] J.F. Broenink and K.C.J. Wijbrans. Describing discontinuities in bond graphs. In *ICBGM93*, pp. 120–125, San Diego, CA, 1993.

[5] D. Karnopp and R.C. Rosenberg. *Analysis and Simulation of Multiport Systems.* John Wiley, NY, 1975.

[6] Z. Kohavi. *Switching and Finite Automata Theory.* McGraw-Hill, Inc., New York, 1978.

[7] B. Kuipers. Qualitative simulation. *Artificial Intelligence*, 29:289–338, 1986.

[8] P. Lötstedt. Coulomb friction in two-dimensional rigid body systems. *Z. angew. Math. u. Mech.*, 61:605–615, 1981.

[9] P.J. Mosterman and G. Biswas. Modeling discontinuous behavior with hybrid bond graphs. In *1995 Intl. Conf. on Qualitative Reasoning*, pp. 139–147, Amsterdam, May 1995.

[10] P.J. Mosterman and G. Biswas. Analyzing discontinuities in physical system models. In *Qualitative Reasoning: The* 10^{th} *Intl. Workshop*, pp. 164–173, Fallen Leaf Lake, CA, May 1996.

[11] P.J. Mosterman and G. Biswas. A formal hybrid modeling scheme for handling discontinuities in physical system models. In *AAAI-96*, pages 985–990, Portland, OR, Aug. 1996.

[12] U. Söderman and J.E. Strömberg. Switched bond graphs: Towards systematic composition of computational models. In *ICBGM95*, pp. 73–79, Las Vegas, Jan. 1995.

[13] J. van Dijk. *On the role of bond graph causality in modelling mechatronic systems.* PhD dissertation, University of Twente, CIP-Gegevens Koninklijke Bibliotheek, Den Haag, The Netherlands, 1994.