Journal of Simulation (2012), 1–12

www.palgrave-journals.com/jos/

An analysis of the cost of validating semantic composability

C Szabo[1]* and YM Teo[2]

[1]The University of Adelaide, Adelaide, Australia; [2]National University of Singapore, Singapore

Validation of semantic composability is a non-trivial problem and a key step in component-based modelling and simulation. Recent work in semantic composability validation promises to reduce verification, validation, and accreditation efforts. However, the underlying cost of current validation approaches can undermine the promised benefits, and the trade-off between validation accuracy and validation cost is not well understood. In this paper we present, to the best of our knowledge, the first quantitative study on the cost of validating semantic composability. Our study covers four representative validation approaches, including two DEVS-based methods, Petty and Weisel formal validation, and deny-validity, and for simplicity, we use computation time as a measure of validation cost. For a queueing model with 1000 components, there is significant trade-off between validation accuracy and cost, with the time-based deny-validity costing seven times that of timeless Petty and Weisel formalism.

Journal of Simulation advance online publication, 29 June 2012; doi:10.1057/jos.2012.11

Keywords: validation; cost; composability

1. Introduction

In component-based simulation model development, composed and validated simulation model components are combined and re-combined to suit specific user requirements (Kasputis and Ng, 2000; Davis and Anderson, 2003). In the development process, the user ideally specifies his requirements to an integrated component-based system that builds the simulation in real time from a library of simulation models that can be easily combined to produce desired functionality (Kasputis and Ng, 2000). Syntactic verification and semantic validation are two key steps in component-based simulation model development (Petty and Weisel, 2003a). In syntactic composability, components have to be properly connected and must interoperate, which assumes common communication protocols, data formats, as well as a common understanding of the time management mechanisms employed. In semantic composability, the composition must be meaningful for all components involved, in terms of data exchange and context. Furthermore, the composition must be valid (Petty and Weisel, 2003a). Model Verification, Validation and Accreditation (VV&A) is one of the most important steps in the simulation life-cycle. Classical, that is, non component-based, validation methods include various stages and processes, and often result in a costly, time-consuming process (Balci, 1997; Department of the Navy, 2004). The validation cost is further aggravated in component-based modelling and simulation where the complexity of the models increases rapidly with size, and semantic composability validation remains a non-trivial problem (Davis and Anderson, 2003; Petty and Weisel, 2003a; Banks *et al*, 2005; Tolk, 2006).

Despite advances in simulation validation, the validation of semantic composability remains a non-trivial problem because of several issues. Firstly, composition is not a closed operation with respect to validation since valid components do not necessary form valid compositions (Balci, 1997). Secondly, reused components are developed for different purposes and when composed may result in emergent properties (Gore and Reynolds, 2008). This implies that the overall behaviour of the composed model cannot be obtained as a union of the individual behaviours of its constituents. Similarly, the context in which a reused component was developed and validated might differ from the new context of the composed model (Bartholet *et al*, 2004; Tolk, 2006), thus influencing the component interaction in unspecified ways. Thirdly, there exist various validation perspectives on the component interactions over time, for example, model behaviour aspects such as deadlock, safety, and liveness, temporal aspects, and formal measures of the validity of compositions, also called 'figures of merit' (Kasputis and Ng, 2000). Lastly, the composed model must be similar or close to the real system it abstracts (Balci, 1997). Very often, the implementation of an automated process to evaluate this similarity is difficult, if not impossible (Balci, 1997; Davis and Anderson, 2003; Tolk, 2006).

**Correspondence: C Szabo, Department of Computer Science, Faculty of Engineering, Computer and Mathematical Sciences, University of Adelaide, Room 4.17, Level 4, Innova21, North Terrace Campus, Adelaide, South Australia 5000, Australia.*
E-mail: claudia.szabo@adelaide.edu.au

In recent years, several automated approaches have been proposed for the verification and validation of composed simulation models, most of which target the validation of semantic composability (Gustavson and Root, 1999; Wainer *et al*, 2002; Petty and Weisel, 2003a; Traore, 2006; Trojet *et al*, 2009). In this paper, we have selected for evaluation a Z-based specification for validating Discrete Event System Specification (DEVS) models (Traore, 2006), an automated input/output transformation validation for CD++ and DEVS models (Wainer *et al*, 2002), a formal theory of composability (Petty and Weisel, 2003a), and our proposed deny-validity approach (Szabo *et al*, 2009; Szabo and Teo, 2009). These approaches focus on various aspects of validity, with several underlying assumptions and limitations that influence their outcome and applicability. An important issue remains their underlying computational cost, especially since in the context of large-scale composed simulation models this cost increases exponentially (Kennedy, 2003; Szabo and Teo, 2009). This is because simulation components are complex entities with a variety of attributes and complex behaviour, and their composition often leads to state space explosion. In particular, the validation of the entire model space as proposed by model checking or theorem proving approaches (Traore, 2006) becomes unfeasible once a certain number of components, connections, and/or interactions is achieved. Similarly, the cost of comparing the composed model with a reference model grows exponentially with the number of components and their attributes and states. Besides problem size, other factors can influence the computational cost of semantic composability. These factors can be classified in two main categories: (i) simulation problem characteristics (ie, number of components, the composition structure, etc), and (ii) validation approach characteristics (abstractions, validation steps etc). We discuss these factors in detail in the following sections.

To distinguish between existing validation approaches that can be employed in the simulation life-cycle, a common metric to evaluate the cost of validation is needed (Pace, 2004). In this paper, we propose to quantify the computational cost of semantic composability validation. We first analyse the validation cost of existing semantic validation approaches in terms of simulation problem characteristics such as the number of components and composition structure. We next analyse the scalability of existing approaches using composed models with a large number of components. Next, we evaluate and quantify the trade-off between validation accuracy and computational cost. Our main contribution is a comprehensive, quantitative analysis on the computational cost of semantic validation in terms of problem characteristics and validation accuracy.

This paper is organised as follows. Section 2 presents an analysis of the computational cost of semantic composability validation. In Section 3 we evaluate current validation approaches in terms of computational cost, number of components, and validation accuracy. We compare and contrast our work with existing evaluation studies in Section 4. Section 5 concludes this paper and discusses future work.

2. Cost of semantic composability validation

The factors that influence the computational cost of semantic composability validation can be classified in two main categories: (i) simulation problem characteristics, and (ii) validation approach. From a component-based perspective, measurable simulation problem characteristics include the number of components, the description of the components, the composition structure, and the degree of interaction between components. Validation approach characteristics include the validation techniques and abstractions employed. These factors are classified in Table 1. The number of components in the composed model describes the size of a component-based simulation model. More fine-grained measures of the composition size are determined by the way in which components are described, in terms of the number of attributes, and the number of states per component. A particular importance, depending on the validation approach, has the number of time delay attributes. The composition structure also influences the computational cost of validation. If a component-port paradigm is adopted, such as in DEVS (Zeigler *et al*, 2000), CoDES (Teo and Szabo, 2008), or OSA (Dalle, 2006), the composition structure is best described using the number of connectors. A large number of connectors, in particular those with complex semantics such as fork/join, can lead to state-space explosion even for composed models with a small number of components, when model checking or theorem proving validation approaches are employed (Szabo and Teo, 2009). This is also the case when component interaction, expressed in terms of the number of events per unit time, is frequent. Validation approach characteristics that influence computational cost include validation techniques and validation abstraction. The validation techniques employed can have a major influence on the computational cost of validation. Classic examples are model checking techniques, which are not feasible when large models, in terms of size and complexity, are validated (Kennedy, 2003). The type of abstraction employed by the validation approach has the potential to reduce the computational cost. For example, if time is not considered in a timeless validation abstraction, the computational cost can be reduced, as we will show in Section 3 for the Petty and Weisel approach.

The validation of semantic composability includes a comparison between the composed model and the reference model, and an evaluation of the closeness between them. The outcome of this evaluation depends on the formalism used to abstract the two models to facilitate reasoning. In this paper we analyse the influence of a time-based and a timeless formalism on computational cost. Another factor

Table 1 Factors that influence the computational cost of validation

Category	*Factor*	*Comment*
Simulation Problem	Number of components	Each composed model consists of several components. *n*—number of components
	Component description	Metrics that describe components eg, number of attributes, number of states *a*—number of attributes; *s*—number of states/component
	Composition structure	Number of connectors of each type, ie, one-to-one, fork, join *c*—total number of connectors
	Degree of interaction	The frequency of interaction between components #events/unit time
Validation Approach	Process	Focus on overall model properties or on comparison between composed model and reference model.
	Abstraction	How the components & composed model are abstracted, eg, using time *versus* timeless
	Techniques	The validation techniques involved, eg, model checking, I/O testing, visualisation, etc

that influences the comparison process is the validation window size, defined as the interval during which the comparison between the composed model and the reference model is performed. Intuitively, given a single validation window, a larger window size ensures that deviant behaviour of the composed model from the reference model can be identified. However, we have found that an increase in size of the validation window beyond a certain knee value comes at very high computational cost. This is similar with the relationship between accuracy and cost suggested in Sargent (2000), in which after a certain knee value, increases in accuracy come at disproportionately larger increases in cost.

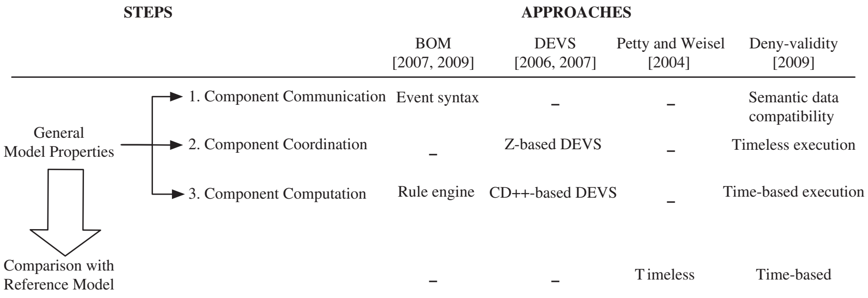
3. Cost analysis

In this section we propose to evaluate the computational cost of four validation approaches, namely, the CD++ DEVS-based validation approach (Wainer *et al*, 2002), the DEVS-based Z specification validation approach (Traore, 2006), the Petty and Weisel formal validation approach (Petty and Weisel, 2003a), and our deny-validity approach (Szabo and Teo, 2009). Our selection was based strictly on the availability of the implemented code, or in the case whereby the code was not available, on the completeness of the descriptions in the published papers to facilitate faithful implementation. The Z specification language has been proposed to formally validate models represented in the DEVS formalism (Zeigler *et al*, 2000; Traore, 2006; Trojet *et al*, 2009). The atomic DEVS model is represented in Z in a time-less manner, and a theorem proving tool based on Z such as Z/EVES (Saaltink, 1997) is used to verify the model and prove certain properties. However, the Z specification language limits the applicability of this approach to coupled DEVS models. Wainer *et al* proposed an input/output transformation validation tool that inputs certain data in the DEVS coupled model and expects specific data through an output point (Wainer *et al*, 2002). This lightweight approach treats a DEVS coupled model as a blackbox and is easy to use since it employs two well-known DEVS/CD++ constructs, that is, Generator to generate input, and Acceptor to accept input. However, only a single output point can be tested and only primitive data types such as real and integer are considered. In the formal theory proposed by Petty and Weisel a composed simulation model is modelled as the composition of mathematical functions that represent components over one-dimensional integer domains (Petty and Weisel, 2003a). The simulation of the composed model is represented as a Labelled Transition System (LTS), where nodes are model states, edges are function executions, and labels are model inputs. Using several metrics, the distance between simulations of the composed model and a perfect model is calculated. However, time is not modelled and the approach is not feasible for compositions with feedback loops and fork and join connectors.

In our previous work, we proposed a deny-validity approach in which the composed model was subjected to a battery of tests aimed at discarding it as semantically invalid[1] (Szabo *et al*, 2009; Szabo and Teo, 2009). Informally, we attempt first to eliminate models in which components cannot communicate and coordinate meaningfully. While properties such as communication and coordination fall into the general definition of simulation verification (Balci, 1997), they are included in the definition of semantic composability (Petty and Weisel, 2003a; Tolk, 2006) and as such they are validated in this layer. Next, models with invalid semantic composability are also those that have valid model properties, but whose execution is not close to that of the real

---

[1] We ensured that models with invalid syntax were eliminated using an approach based on compositional grammars expressed in EBNF (Teo and Szabo, 2008).



STEPS		APPROACHES BOM [2007, 2009]	DEVS [2006, 2007]	Petty and Weisel [2004]	Deny-validity [2009]
General Model Properties	1. Component Communication	Event syntax	_	_	Semantic data compatibility
	2. Component Coordination	_	Z-based DEVS	_	Timeless execution
	3. Component Computation	Rule engine	CD++-based DEVS	_	Time-based execution
Comparison with Reference Model		_	_	Timeless	Time-based

Figure 1 Landscape of recent validation approaches.

system the composed model abstracts (Balci, 1997). To eliminate models that are not similar to the real system being abstracted, we compare between the composed model and a reference model. In contrast to current static perfect model validation (Petty and Weisel, 2003a), our proposed time-based formalism represents dynamic component behaviour, can represent fork and join connectors, and is applicable to a composed models with wide variety of topologies and interactions. Furthermore, we are able to quantify the similarity between the composed model execution and the reference model execution using our defined semantic metric relation.

We divide our evaluation process in two stages. Firstly, we evaluate the cost of validating general model properties. Secondly, we evaluate the cost of comparing between the composed model and a reference model. The factors that influence the computational cost of validation form a complex parameter space in which the influence of each individual factor is difficult to identify. Thus for simplicity, we employ a single-server queue model with a varying number of components. A more complex example is analysed in Section 3.4.

3.1. Evaluation methodology

To better understand the landscape of semantic composability validation, we summarise existing validation approaches in Figure 1. An important observation is that most component-based verification and validation approaches focus on two main aspects, namely, on the validation of general model properties, and on the comparison of the composed and the reference models. General model properties are validated by the CD++ DEVS-based approach (input/output transformations), the DEVS-based Z specification approach (component coordination), and BOM-based approaches (component communication and input/output transformations). Comparison between the composed model and a reference model is performed by the Petty and Weisel approach. Lastly, our deny-validity approach targets both validation stages.

For our study, we were able to obtain the code for the CD++ DEVS-based approach. We implemented the second DEVS-based approach following the complete descriptions given in Traore (2006). However, the Z specification does not permit the representation of coupled or connected objects, and as such it cannot be used to specify component-based models. Instead, we employed the Object-Z formalism (Duke *et al*, 1991), which can be easily adapted in a similar manner as that suggested in Traore (2006), to cater for DEVS coupled models. However, model checking based on Object-Z is still in its early stages. Nonetheless, we were able to reach the outcome described in Traore (2006), that is, syntax, type, and inconsistency checking by employing the Wizard checker for type and syntax checking on the composed model (Smith, 2000). Ongoing work exists to include the Object-Z specification into theorem proving tools such as Isabelle/HOL (Smith *et al*, 2002), but these were beyond the scope of our study. The advantage and major improvement of our implementation is that it caters for component-based models. We implemented the Petty and Weisel approach following details from the published papers (Petty and Weisel, 2003a, b; Petty *et al*, 2005). A fundamental assumption in the Petty and Weisel approach is that a simulation component can be transformed into a mathematical function over integer domains. Nonetheless, no details about how this transformation should be performed are provided. This transformation is crucial when computing the mathematical composability of the functions that represent components, which by definition translates to verifying integer domain inclusion. In our implementation, we employed a brute-force transformation by mapping every component output into an integer number (iteratively for each output) and assuming that the functions are mathematically composable. This is not a limitation in our study of computational cost because checking for mathematical composability is not a major component of the validation cost in the Petty and Weisel approach.[2]

Our study follows the two main validation steps described in Figure 1. Firstly, we evaluate the computational cost of

---

[2] In particular, provided that the transformation from the meta-model to integers is sound, the algorithm to determine mathematical composability could employ a Radix sort algorithm (Knuth, 1997) to first sort the integer values, followed by an inclusion check, resulting in O(kn) steps, where n is the number of elements in the interval, and k is the average element length.

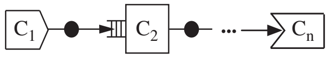


Figure 2 Simple queueing model.

Table 2 Computational cost of validating general model properties

#Components	*Runtime (s)*					
	*DEVS*1 *P*2	*DEVS*2 *P*3	*Deny-validity*			
			*P*1	*P*2	*P*3	*Total*
10	<0.1	0.2	<0.1	5.3	51.1	56.5
100	<0.1	0.5	0.2	146.6	43.5	190.3
500	0.2	4.5	0.3	193.6	67.0	260.9
1000	0.7	16.7	0.5	330.4	130.9	461.7

approaches that aim to validate general model properties, namely, the CD++ DEVS-based validation approach (Wainer *et al*, 2002), the DEVS-based Z specification validation approach (Traore, 2006), and our deny-validity approach (Szabo *et al*, 2009; Szabo and Teo, 2009). We employ a simple queue model with varying number of service unit components as shown in Figure 2. We have chosen this queue model because of its relative simplicity with respect to factors such as the number of attributes per component and composition structure. Besides the need for simplicity, we have also excluded these factors because they are not considered by most validation approaches. Secondly, we study two approaches that compare the composed model with a reference model, namely, Petty and Weisel's validation approach (Petty and Weisel, 2003a) and our deny-validity approach. We evaluate the variation of cost with model size and analyse these approaches in terms of accuracy and cost. We first show the impact of timeless abstractions, such as that employed by Petty and Weisel, on the validation process. Next, we evaluate the variation of cost with the validation accuracy, expressed using the size of the validation window during which the composed model is compared with the reference model. All experiments were executed on a Dell PowerEdge SC1430 Dual Quad Core Server, with Intel Xeon, 1.83 GHz, and 4GB RAM. Because the validation approaches are deterministic in their execution, variation in execution time is caused by the operating system. As such, we analyse 10 execution runs and present their minimum.

3.2. General model properties

The validation of general model properties focuses on three main properties, namely, component communication (P1), component coordination (P2), and component computation (P3), in the composed model. The validation of component communication aims to check data compatibility between connected components in the composition. Current approaches ensure that components follow a common reference model (Tolk *et al*, 2008), verify syntax in terms of event parameters and data types (Moradi *et al*, 2007), or validate semantic compatibility using a component-based ontology (Teo and Szabo, 2008). The validation of component coordination aims to check that interleaved executions of components in the model are correct (Defense Modeling and Simulation Office (DMSO), 1996). This is usually done using model checkers, and in general by abstracting time in instantaneous transitions (Kennedy, 2003; Szabo and Teo, 2009). Lastly, the validation of component computation aims to check that the components can execute during the composition run. For simplicity, the Z-specification DEVS-based and the CD++ approach will be hereafter referred to as DEVS1 and DEVS2 respectively.

Table 2 presents the variation of computational cost when varying the number of components from 10 to 1000. While the computational cost increases proportionally with the composed model size, the increase is insignificant in the DEVS1 and DEVS2 approaches, in which the computational cost approaches one second and 17 s for 1000 components respectively. In contrast, the deny-validity approach takes close to 7 min for the same model. The reason for this discrepancy becomes evident when we look at the cost components of the deny-validity approach, which is the only approach that validates all general model properties. The cost of validating property P1 is insignificant for this model, but increases with the number of data types that a component outputs. This is because P1 is validated by establishing data compatibility between components using our proposed ontology, which is queried for every pair of connected components (Teo and Szabo, 2008). Space constraints prevent us from showing these results here. The validation of property P2 incurs the largest cost of the three properties. The validation is done using the SPIN (Ben-Ari, 2008) model checker to validate a specification of the composed model, and becomes highly unfeasible when the composed model size increases beyond 250 components. As such, flags that limit the search space are employed for validation of models with 500 and 1000 components. In particular, we use the '−w' flag to reduce the depth of the search tree, and '`DMEMLIM`' to cap the amount of memory used. Property P2, component coordination, is also validated by the DEVS1 approach. However, this approach focuses only on type and syntax checking, and does not look at all possible interleaved execution states, like the deny-validity approach.

The validation of property P3 is performed in a similar manner for DEVS2 and the deny-validity approach, by checking several types of data at a connection point in the composition. However, the DEVS2 approach can only process a limited number of events and requires the user to input the exact moment in time when the output is expected. Moreover, the DEVS2 approach validates input at the last connection point in the composition. Lastly, only primitive data types such as real or integer can be validated. In contrast, in the deny-validity approach the user can specify desired data of any type, including domain specific, at any connection point in the composition. Additional liveness properties are also validated (Szabo and Teo, 2009).

3.3. Comparison with reference model

Comparing the simulation model with a real or referent system is a traditional validation approach (Balci, 1997). This comparison is done in component-based simulation by the Petty and Weisel (Petty *et al*, 2005) and the deny-validity approaches (Szabo and Teo, 2009). These two approaches follow a similar validation sequence, which can be separated into three major steps: (i) transformation of components into formalism; (ii) transformation of composition formalism into an LTS (Srba, 2001); and (iii) comparison with reference model. Both approaches rely on a formal comparison between the simulation of the composed model and that of a reference model. The major difference between the two approaches lies in that Petty and Weisel employ a static formalism, in which a component is represented as a function over integer domains, whereas the deny-validity approach proposes a time-based formalism in which a component is represented as a function over a three-coordinate domain containing time, state, and input/output. The Petty and Weisel approach offers a high level of abstraction, which permits reasoning about closeness under composition and has reduced computational cost as we will show below. However, it is difficult to transform component representations into integer values automatically, and the approach assumes that model properties, such as syntactic composability, safety, liveness, are validated beforehand. Moreover, the comparison process orders the component functions based on the location (left to right) of the components in the composition, which does not permit the validation of compositions with fork and join connectors and feedback loops. This is because the functions representing components on the fork/join branches need to be ordered during execution, which is not possible using only integer value domains. The same applies to feedback loops, where a mathematical composability ordering based on the position of the components cannot be deduced. Figure 3 highlights the difference between timeless and time-based ordering. As it can be seen, the LTS that represents the composed model has an equal number of states in both approaches, but a different sequence of function execution. The deny-validity approach orders the functions based on the time moment when the component interacts with its neighbours, resulting in the sequence $f_1, f_2, f_3, \ldots$ In contrast, the Petty and Weisel approach orders the functions based on the position of the components in the composition, resulting in the sequence $f_1, f_2, f_3, \ldots$. Although there are instances where a position ordering is useful, for example, when reasoning about validation closeness under composition, the latter time-based ordering paints a more realistic picture of the execution of the composed model. However, the time-based formalism employed in our deny-validity approach incurs an additional validation overhead as we discuss below.

The computational cost of validation can also be divided into three components, namely, (i) f, the formalism cost, as the cost of transforming from the component representation to the chosen formalism; (ii) p, the process cost, as the cost of transforming the composed model into an LTS using the chosen formalism; and (iii) c, the comparison cost, as the cost of comparing between the two LTS, representing the composed model and the reference model respectively. These cost components can be further refined as functions of various component and composition characteristics, as shown in Table 3. An important point to highlight is that the cost of obtaining or constructing the reference model is not included. This is because the reference model is assumed to exist *a priori* in the Petty and Weisel approach, whereas in our approach the reference model is automatically derived based on generic descriptions of reference components (Szabo *et al*, 2009; Szabo and Teo, 2009).

Table 3 Cost components

Cost component	*Petty and Weisel*	*Deny-validity*
Formalism—f	at least $f(n, \tau)$	$f(n, s, t, \tau)$
Process—p	$p(n, \tau)$	$p(n, s, t, \tau)$
Comparison—c	$c(n, \tau)$	$c(n, \tau) + c_{eps}(n, a, \tau)$

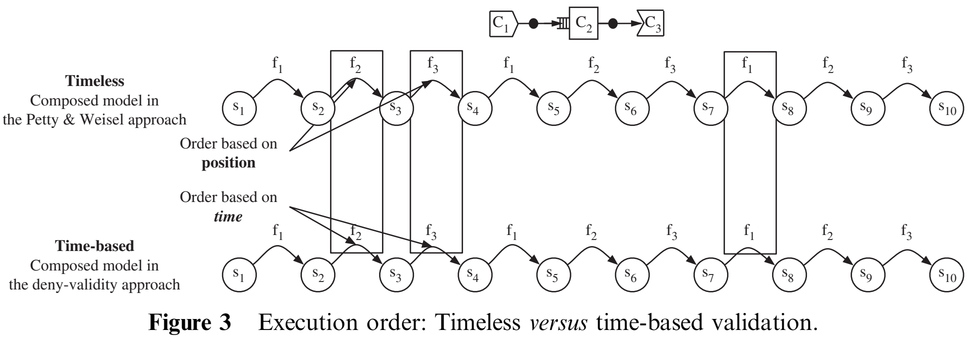


Figure 3 Execution order: Timeless *versus* time-based validation.

Our results presented below show that the number of components (n) drastically influences the computational cost. This is because the fundamental unit of each validation approach is the mathematical function, which represents a component. Other parameters, such as the average number of states per component (s) and the average number of attributes per component (a), influence the cost of the deny-validity approach but not the cost of the Petty and Weisel approach. This is because the Petty and Weisel approach deals only with integer representations. Nevertheless, the influence of the number of states and attributes on the computational cost in the Petty and Weisel approach should be more evident in the formalism cost, f, because the transformation from the component representation to unique and meaningful integer values should consider states and attributes as well. However, Petty and Weisel do not discuss details about how this transformation is performed. Another parameter that influences the computational cost is τ, the size of the validation window during which the composed model is compared to the reference model. This translates into the number of simulation steps in the Petty and Weisel approach, and into the unfolding degree in our deny-validity approach. For example, for the simple model in Figure 3, $\tau = 3$, resulting in three simulation steps and 10 states for the LTS representing the composed model. The parameter τ can be seen as a measure of the accuracy of the validation process if we agree that the longer the interval under which the composition is observed as compared with the reference model, the more accurate is the validation result. We next evaluate the computational cost of semantic validation with the composition size in terms of the number of components, n and analyse the trade-off between the computational cost and the validation window size τ.

Results and analysis. We implemented the Petty and Weisel approach based on its description (Petty and Weisel, 2003a, b; Petty *et al*, 2005). Since the details of mapping each component into a function over integer domains are not documented in Petty and Weisel approach, we used a simple and fast heuristic as discussed in Section 3. Next, the composed model LTS was created as shown in Figure 3. Lastly, we implemented the comparison between the composed model LTS and a reference model LTS using the BISIMULATOR tool (Garavel *et al*, 2007), which is the same one we employ in comparing between the LTS of the composed model and reference model in the deny-validity approach.

Table 4 presents the variation of computational cost when varying the number of components from 10 to 1000, with a validation window of size $\tau = 3$.

Table 4 Comparison with reference model

#Components	*Runtime (s)*	
	Petty and Weisel	*Deny-validity*
10	1.40	3.18
100	3.30	14.34
500	8.42	33.67
1000	13.40	75.70

As it can be seen in Figure 4(a), the computational cost of the Petty and Weisel approach is reduced, for example, to less than 15 s for a simple queueing model with 1000 components. In contrast, our deny-validity approach has a runtime of around 1 min for the same model. This is because in our approach, a larger number of components implies that a larger number of component executions have to be ordered towards achieving the time-based ordering presented in Figure 3. To evaluate the trade-offs between computational cost and accuracy of the validation process, we evaluate the runtime cost of validating a queueing model with 1000 components while varying the values of the validation window size τ to 3, 10, 15, and 20. Our results are presented in Figure 4(b).

The increase of the computational cost with the validation window size τ is insignificant for the Petty and Weisel approach. As it can be seen in Figure 4(b), for $\tau = 25$, the minimum runtime of the Petty and Weisel approach is on 35.5 s for an M/M/1 model with 1000 components. In contrast, there is an evident trade-off in the deny-validity approach. Specifically, the validation runtime increases from around minimum 2.5 min for $\tau = 20$, to around 18 min for $\tau = 25$. The explanation for this decrease in performance is the following. Our validation process includes a time-ordering module which orders all functions based on the time moment in which components interact with neighbours. This implies that a correct time ordering of components is necessary. The computation of this ordering is a constraint satisfaction problem, requiring a constraint solver to solve several equations. For $\tau = 25$, the LTS of the composed model has around 25 000 states. This translates into 25 000 constraints over the entire positive integer domain (since time values are integers) that have to be solved in order to determine the time ordering of the function executions, as required by our validation process. This operation has to be performed twice, for the composed and the reference model respectively. In solving these constraints, we employ the Choco constraint solver (Choco Constraint Programming System, retrieved January 2010), which for our constraint types has a theoretical complexity of $0(c^3)$, where c is the number of constraints. This problem does not appear in the Petty and Weisel approach, because time is not modelled and as such a time-based ordering is not necessary. For the curve in Figure 4(b), we have calculated the knee values beyond which increases in validation window size come at disproportionate increases in validation cost. For the deny-validity approach, for models with 500 components, we have calculated the knee value as $\tau_{\text{knee}} = 15.42 \approx 15$. For models with 1000 components, $\tau_{\text{knee}} = 19.59 \approx 20$.

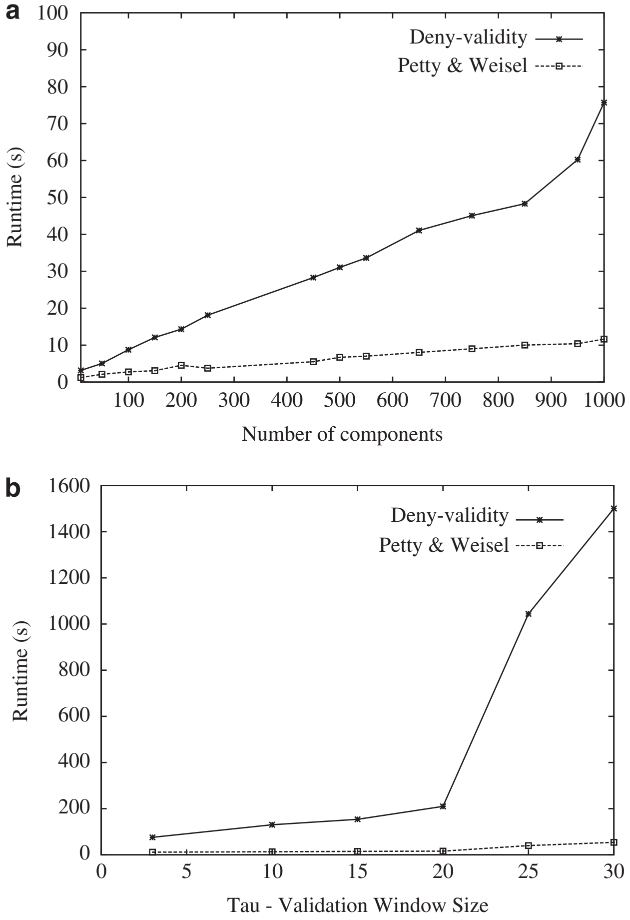


Figure 4 Comparison between Petty and Weisel and deny-validity: (a) cost *versus* size; (b) cost *versus* accuracy.

3.4. Further insight

In the above, we limited our evaluation study to only two factors, namely, the number of components n and the validation window size τ. Another important factor that influences the cost of validation is the composition structure, described by the number of one-to-one, fork, and join connectors. Existing validation approaches cannot validate models with fork and join connectors. However, this is possible in our deny-validity approach. We analyse the cost of validating composed models that contain fork and join connectors as shown in Figure 5.

Figure 4 presents a breakdown of the validation cost for composed models that contain only one-to-one connectors. Figure 4(b) presents a breakdown of the validation cost for composed models that contain a ratio $r = 10\%$ of fork and join connectors, for example, for 1000 components there are 10 fork and 10 join connectors.

As expected, for a composed model with $n = 1000$ and $\tau = 20$, the total validation time increases from 10 min for the simple composition in Figure 4(a) to around 17 min for the composed model with a more complex structure in Figure 4(b). Moreover, the percentage of validation cost also changes. In the composed model without fork and join connectors, the percentage of validation cost was distributed as 50:19:31% for component coordination, component computation, and validation by comparison with reference model respectively. In contrast, the presence of fork and join connectors leads to a 40:15:45% cost distribution, suggesting that as the number and complexity of connectors increase, the penalty incurred by the comparison with a reference model also increases. However, the cost of verifying component coordination becomes unfeasible as suggested before (Kennedy, 2003) and is capped as discussed in Section 3.2. An important point to highlight is that the total execution cost for an SSF implementation (Cowie, 1999) of the simple composed model with $n = 1000$ is a minimum of 49.57 s, and the cost of validating it is a minimum of 11.28 min. However, because of the layered nature of the validation approach, the cost incrementally increases and the user can stop at any validation layer as desired. The validation of component communication and coordination incurs an initial overhead of 5.5 min. The meta-simulation layer incurs and additional 2.14 min overhead. Finally, the comparison with a reference model incurs a final cost of 3.64 min. This increase in the cost of the validation by comparison with a reference model is due to the increase in the number of states in the LTS. For $n = 1000$ and $\tau = 3$, the total number of states for the model in Figure 5 is 19 000. The number of states increases to 39 960 when $n = 1000$ and $\tau = 20$ (Figure 6).

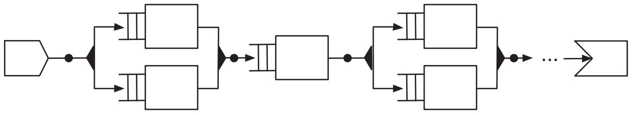


Figure 5 Queueing model with fork/join connectors.

The validation cost is also influenced by the complexity of the composition structure. Our next experiments analyse the variation of the validation cost when r, the percentage of fork connectors in the composition, increases from 10 to 30%. We analyse various problem and validation window sizes, $(n; \tau)$, and the problem size, in terms of $(n; \tau)$, increases from $(1000; 3)$ to $(1000; 20)$ as before. The results shown in Table 5 represent the minimum of five runs.

Table 5 shows the variation of the cost of validating general model properties and the cost of comparing with a reference model when the complexity of the composition structure in terms of the ratio of fork and join connectors to the total number of connectors increases from $r = 10\%$ to $r = 30\%$. As it can be seen, once r increases beyond 15%, the dominant cost becomes that of comparing with a reference model. This is because an increase in the number of fork and join increases the cost of computing

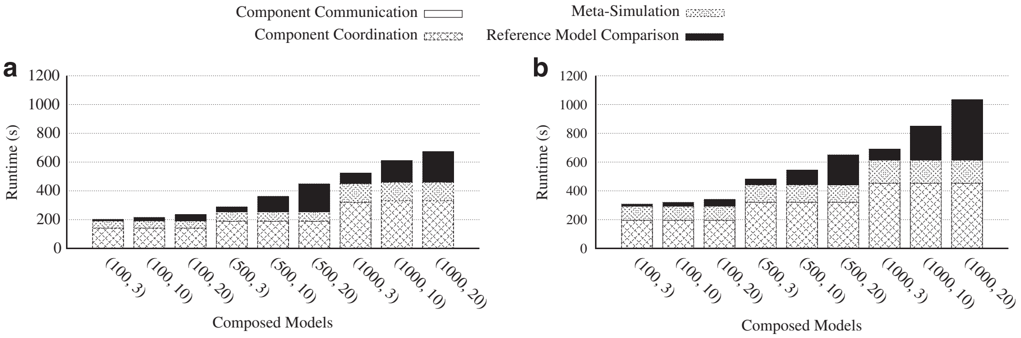


Figure 6 Total validation cost.

Table 5 Validation cost with composition structure

Problem size (n; τ)	*Composition structure r(%)*	*Runtime (s)*	
		Model properties	*Model execution*
(1000; 3)	15	321.6 (MS) + 160.8	72.59
	20	358.23 (MS) + 234.6	80.25
	25	408.13 (MS) + 265.3	93.53
(1000; 10)	15	320.1 (MS) + 160.8	150.25
	20	360.1 (MS) + 234.6	219.10
	25	408.3 (MS) + 265.3	239.61
(1000; 20)	15	320.1 + 160.8	427.32
	20	361. 2 (MS) + 234.6	500.12
	25	407.13 (MS) + 265.3	624.43
	30	452.32 (MS) + 280.2	731.36

the order of the LTS states, despite the number of states remaining constant.[3]

3.5. Discussion

This study has analysed four representative validation approaches, namely, the CD++ DEVS-based validation approach, the DEVS-based Z specification validation approach, the Petty and Weisel validation approach, and the deny-validity validation approach. The first two approaches propose to validate general model properties, such as input/output transformations among others. We have compared these approaches with the deny-validity approach, in which the first validation stage also looks at safety, liveness, and input/output transformations. Our study shows that the deny-validity approach is extremely sensitive to the model size in terms of the number of components, with a total validation time of 1 min for a model with 10 components, and increasing to 7 min for a composed model with 1000 components. The DEVS-based approaches have a significantly smaller validation time, with at most 15 s for a model with 1000 components. The reasons for the discrepancy in the validation costs for these approaches are twofold. First, the DEVS-based approaches do not analyse safety and liveness and other logical properties of the composed model. The validation of these properties is time consuming and suffers from state-space explosion, but provides increased confidence in the validity of the composed model. Second, the deny-validity approach performs a semantic validation of input/output transformations, in that it queries a simulation ontology to determine partial relationships between data exchanged by the components in the composed model. In contrast, the DEVS-based approaches only look at exact matching between input and output. This provides reduced computational cost but also decreases the knowledge about what data are exchanged between components.

We next compare between the Petty and Weisel approach and the deny-validity approach in the validation of the composed model as compared with a reference model. Our study has found that the introduction of time in the deny-validity approach *versus* a timeless validation approach as proposed by Petty and Weisel incurs significant trade-offs in the computational cost of validation. Specifically, a timeless Petty and Weisel validation approach incurs a validation cost of around 13 s for models with 1000 components, and seems to grow linearly with the composed model size. In contrast, the time-based deny-validity approach takes around a minute for models with 1000 components, and seems to grow exponentially with the composed model size. This exponential growth is further aggravated when the size of the validation window is increased. In contrast, the increase in validation window size has no effect on the Petty and Weisel approach. However, a timeless validation approach cannot validate models with complex structure, such as fork and join connectors. We further analyse the deny-validity approach. Our study finds that the number of fork and join connectors together with the validation window size have a direct effect on the cost of validation, which reaches around 10 min for large and complex models, with the dominant cost being the incurred by the time-based validation formalism.

---

[3] However, the validation of component coordination using model checking remains capped to reduce state space explosion.

4. Related work

Verification, Validation, and Accreditation (VV&A) has been a principal focus of research in the simulation community since the late 1970s (Sargent, 1979), with work such as that by Balci (Balci and Sargent, 1981, 1982; Balci, 1997) and Sargent (Balci and Sargent, 1981, 1982; Sargent, 2000, 1979) among others, providing detailed guidelines and process organisation. However, very few works quantify the cost of validation, more so in the new area of component-based modelling and simulation. This is mainly because validation cost is highly dependent on many factors, such as characteristics of the simulation problem, model and application domain among others, and thus is difficult to estimate. Balci and Sargent propose a methodology for cost-risk analysis in the statistical validation of simulation models, by looking at the relationship between data collection cost, acceptable validity range, and model builder and model user risks, when performing statistical hypothesis testing (Balci and Sargent, 1981). A validity measure is proposed and the trade-offs between data collection budget and model user's risk are analysed. However, there is no estimate of the actual cost of validation. Historical data from the US Defense Modeling and Simulation Office suggest that VV&A activities account for 5-17.5% of the total modelling and simulation budget (Defense Modeling and Simulation Office (DMSO), 1996). A clear estimate of the validation cost and the key factors that influence it is also missing in the simulation industry. Reports from industry suggest that validation cost is between 5 and 19% of the total project cost (Love and Back, 2000) but no cost model is provided. The same applies to the validation of component-based simulations. While there are many works that focus on verifying and validating general model properties (Wainer *et al*, 2002; Traore, 2006; Moradi *et al*, 2007; Trojet *et al*, 2009) or on comparing the composed model with a reference model (Petty and Weisel, 2003a; Szabo and Teo, 2009), existing work lacks an evaluation of the applicability and cost of proposed approaches. In contrast, in this paper we evaluate four existing approaches. We present a classification of the factors that influence computational cost and analyse the cost of existing validation approaches in terms of the number of components and composition structure.

5. Conclusion

Industry practice suggests that the cost of traditional simulation validation ranges from 5 to 19% of the entire project cost (Love and Back, 2000). In component-based modelling and simulation, semantic validation cost can be significantly higher. To the best of our knowledge, we present the first quantitative study of the cost of validating semantic composability and its trade-offs with validation accuracy. This study compares the cost of four main validation approaches, namely, CD++ DEVS (Wainer *et al*, 2002), Z specification-based DEVS (Traore, 2006), Petty and Weisel formal theory (Petty and Weisel, 2003a), and deny-validity (Szabo *et al*, 2009; Szabo and Teo, 2009). A queueing model with one thousand components is used and validation measurement is performed on a Dell Pow-erEdge SC1430 compute server. The key factors that influence the computational cost are simulation problem characteristics, including the composition size and the degrees of interaction, and validation approach characteristics, including the techniques used and the levels of abstraction. Current validation approaches focus on validating general model properties, and on comparing between the composed model and a reference model. In general, the cost of validating model properties grows, as expected, with the number of components. However, the actual cost is vastly different among approaches. The computational cost ranges from less than 1 s in CD++ DEVS to more than 7 min in deny-validity. The cost depends mainly on the underlying abstractions and assumptions employed.

The cost of validating a composed model against a reference model in the time-based deny-validity approach is seven times more than the timeless Petty and Weisel approach. The time-based formalism and ordering in deny-validity provides increased accuracy at the expense of runtime. Moreover, the cost of time-base ordering in deny-validity explodes when the validation window size is increased. In the Petty and Weisel approach, increasing validation window size has no significant impact on cost. In contrast, a time-based ordering facilitates the validation of composed models with complex structures including feedback loops and fork and join component connectors. Specifically, a 25% increase in validation window size results in a five-fold increase in validation cost.

As the deny-validity approach is one of the most comprehensive process for semantic validation, allowing different degrees of validation accuracy and cost, we analyse it further. We first analysed the cost of the different validation layers in the deny-validity approach using a simple queueing model. Next, we analysed the influence of the composition structure, in terms of the number of fork and join connectors, on the validation cost. Our study shows that while it takes a minimum of 49 s to execute the SSF implementation of a composed simulation model with 1000 components, it takes minimum 17 min to validate it. The cost of validating general model properties represents 55% and the cost of validating model execution accounts for the remaining 45%. A 10% increase in the number of fork and join connectors in the composition comes with a 50% increase in the validation cost.

References

Balci O (1997). Verification, validation and accreditation of simulation models. In: Andradóttir S, Healy KJ, Withers DH and Nelson BL (eds). *Proceedings of the Winter Simulation*

Conference, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 135–141.

Balci O and Sargent RG (1981). A methodology for cost-risk analysis in the statistical validation of simulation models. *Communications of the ACM* **24**(4): 190–197.

Balci O and Sargent RG (1982). Some examples of simulation model validation using hypothesis testing. In: Chao YW, Madrigal O and Highland HJ (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 621–629.

Banks J, Carson JS, Nelson BL and Nicol DM (2005). *Discrete Event System Simulation*. Prentice-Hall: USA.

Bartholet RG, Brogan BC, Reynolds PP and Carnahan JC (2004). In search of the philosopher's stone: Simulation composability versus component-based software design. In: Smith J, Peters B, Ingalls RG and Rossetti MD (eds). *Proceedings of the Fall Simulation Interoperability Workshop*, (PCE) Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA.

Ben-Ari M (2008). *Principles of the Spin Model Checker*, Springer Verlag: Berlin.

Choco Constraint Programming System. http://sourceforge.net/projects/choco/, accessed January 2010.

Cowie J (1999). Towards realistic million-node internet simulations. In: Arabnia HR (ed). *Proceedings of International Conference on Parallel and Distributed Processing Techniques and Applications*, CSREA Press: Las Vegas, USA, pp 2129–2135.

Dalle O (2006). OSA: An open component-based architecture for discrete-event simulation. In: Burotzky W (ed). *Proceedings of the 20th European Conference on Modeling and Simulation*, Society of Modeling and Simulation Europe: Leipzig, Germany.

Davis PK and Anderson RH (2003). Improving the composability of Department of Defense models and simulations.

Defense Modeling and Simulation Office (DMSO) (1996). Verification, validation, and accreditation recommended practices guide. U.S. Department of Defense. Office of the Director of Defense Research and Engineering.

Department of the Navy (2004). *Modeling and Simulation Verification, Validation, and Accreditation Implementation Handbook*. Department of the Navy, http://nmso.navy.mil/NMSODigitalLibrary.aspx?Command = Core_Download&EntryId = 3988.

Duke R, King P and Rose G (1991). *The Object-Z Specification Language: Version 1*. Technical Report, http://www.dc.uba.ar/people/materias/isoft1/papers/tr91-1.pdf.

Garavel H, Lang F, Mateescu R and Serwe W (2007). CADP 2006: A toolbox for the construction and analysis of distributed processes. In: Werner D and Holger H (eds). *Proceedings of the 19th International Conference on Computer Aided Verification*, Springer Verlag: Bonn, Germany, pp 158–163.

Gore R and Reynolds PP (2008). Applying causal inference to understand emergent behavior. In: Mason SJ, Hill R, Moench L and Rose O (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 712–721.

Gustavson P and Root L (1999). Object model use cases: A mechanism for capturing requirements and supporting BOM reuse. In: Hieb M and Blalock J (eds). *Spring Simulation Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA.

Kasputis S and Ng HC (2000). Composable simulations. In: Joines JA and Barton RlR (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 1577–1584.

Kennedy KE (2003). Formal methods in the verification and validation of simulation models. In: Mason SJ and Hill R (eds). *Proceedings of the Spring Simulation Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA.

Knuth DE (1997). *The Art of Computer Programming, Volume 3: Sorting and Searching*. Addison-Wesley: Reading, MA.

Love G and Back G (2000). Model verification and validation for rapidly developed simulation models: Balancing cost and theory. In: Davidsen PI, Ford DN and Mashayekhi, AN (eds). *Proceedings of the 18th International Conference of the System Dynamics Society*, Bergen, Norway.

Moradi F *et al* (2007). A rule-based approach to syntactic and semantic composition of BOMs. In: Roberts DJ, Theodoropoulos GK and Saddik AEl (eds). *Proceedings of the 11th IEEE Symposium on Distributed Simulation and Real-Time Applications*, Systems Dynamics Society, IEEE Computer Society: Crete, Greece, pp 145–155.

Pace DK (2004). Modeling and simulation verification and validation challenges. *Johns Hopkins APL Technical Digest* **25**(3): 163–172.

Petty MD and Weisel EW (2003a). A composability lexicon. In: Mason SJ and Hill R (eds). *Proceedings of the Spring Simulation Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA, pp 181–187.

Petty MD and Weisel EW (2003b). Basis for a theory of semantic composability. In: Mason SJ and Hill R (eds). *Proceedings of the Spring Simulation Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA.

Petty MD, Weisel EW and Mielke R (2005). Composability theory overview and update. In: Hill R and Rose O (eds). *Proceedings of the Spring Simulation Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA.

Saaltink M (1997). The Z/EVES system. *Lecture Notes in Computer Science* **1212**(1): 72–85.

Sargent R (2000). Verification, validation, and accreditation of simulation models. In: Joines JA and Barton RlR (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 50–59.

Sargent RG (1979). Validation of simulation models. In: Highland HJ, Spiegel MG and Shannon R (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 497–503.

Smith G (2000). *The Object-Z Specification Language*. Kluwer Academic Publishers: Dordrecht, MA.

Smith G, Kamm?ller F and Santen T (2002). Encoding object-z in Isabelle/HOL. In: Bert D, Bowen JP, Henson MC and Robinson K (eds). *Proceedings of the International Conference of B and Z Users*, Springer Verlag: London, UK, pp 82–89.

Srba J (2001). On the power of labels in transition systems. In: Larsen KG and Nielsen M (eds). *Proceedings of the 12th International Conference on Concurrency Theory*, Springer Verlag: Aalborg, Denmark, pp 277–291.

Szabo C and Teo YM (2009). An approach for validation of semantic composability in simulation models. In: Hill R and Moench L (eds). *Proceedings of the 23rd ACM/IEEE/SCS Workshop on Principles of Advanced and Distributed Simulation*, IEEE Computer Society: New York, USA, pp 3–10.

Szabo C, Teo YM and See S (2009). A time-based formalism for the validation of semantic composability. In: Rossetti M, Hill R, Johansson B and Dunkin A *et al* (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 1411–1422.

Teo YM and Szabo C (2008). CODES: An integrated approach to composable modeling and simulation. In: Roberts DJ and

Karatza H (eds). *Proceedings of the 41st Annual Simulation Symposium*, IEEE Computer Society: Ottawa, Canada, pp 103–110.

Tolk A (2006). What comes after the semantic web—PADS implications for the dynamic web. In: Turner S and Riley G (eds). *Proceedings of the 20th Workshop on Principles of Advanced and Distributed Simulation*, IEEE Computer Society: Singapore, pp 55–62.

Tolk A, Diallo SY and Turnitsa CD (2008). Mathematical models towards self-organizing formal federation languages based on conceptual models of information exchange capabilities. In: Mason SJ, Hill R, Moench L and Rose O (eds). *Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 966–974.

Traore MK (2006). Analyzing static and temporal properties of simulation models. In: Lawson B, Liu J, Perrone F and Wieland F *et al* (eds). *Proceedings of the Winter Simulation Conference*, Institute of Electrical and Electronics Engineers Inc.: Piscataway, New Jersey, USA, pp 897–904.

Trojet MW, Frydman C and Hamri ME-A (2009). Practical application of lightweight Z in DEVS framework. In: Spiegel MG and Shannon R (eds). *Proceedings of the Spring Simulation Multiconference*, Simulation Interoperability Standards Organization: Orlando, USA.

Wainer G, Morihama L and Passuello V (2002). Automatic verification of DEVS Models. In: Liu J and Hill R (eds). *Proceedings of the SISO Spring Interoperability Workshop*, Simulation Interoperability Standards Organization: Orlando, USA.

Zeigler BP, Praehofer H and Kim TG (2000). *Theory of Modeling and Simulation*. Academic Press: New York.

Received 22 September 2011;
accepted 20 April 2012 after two revisions