Prioritization Framework: A Step Toward Cost-Effective Verification, Validation, and Accreditation

S. K. Numrich, Robert A. Zirkle, James R. Ayers, Forrest R. Smith, and Anna Vasilyeva

Verification, validation, and accreditation (VV&A) is too often done as an afterthought to the development process of a simulation at a time when most of a project's resources have been exhausted. VV&A may also be done when some higher authority threatens the existence of a program because the simulation tools used have not been subjected to VV&A. Under such circumstances, the process becomes a tax on already burdened programs and a headache for the program manager. The method discussed here was developed to assist a program manager in performing verification and validation (V&V) on an existing simulation within budget and without sacrificing application performance. Prioritization based on the intended use of the simulation seemed the most beneficial route to performing a cost-effective V&V process (Department of Defense 2007).

The framework gives the user the ability to provide rational, repeatable, and documented evidence for decisions concerning where to focus V&V resources.

Terms

The purpose of verification is to determine whether the equations or computational models used to represent the entities in a simulation are encoded properly so that the software accomplishes what the model developer intended it to accomplish. In validation, the essential question is whether or not the encoded *representation* corresponds to the measure of the physical (or real) world it is supposed to represent. The measure of the real world, called a *referent*, encapsulates an understanding of the segment of the real world to be captured in a simulation. Evaluation of how well the representation corresponds to the referent could be accomplished using comparisons to measured data or to a commonly accepted mathematical relationship. Lacking either of those means, developers may seek the considered opinion of subject matter experts.

To accomplish V&V of a simulation, the *V&V agent*—the organization, group, or person performing V&V activities—must have a viable set of *requirements* describing what the software is supposed to do, along with referents or acceptable standards of representation for those requirements. These requirements and referents should have guided the development of the model or simulation. Without the guidance provided by requirements and in the absence of good referents, any degree of performance could constitute acceptable correspondence to the real world, potentially leaving the user with software that is inappropriate, inadequate, or unusable.

Nonetheless, upon initiating the V&V effort, the V&V agent often discovers that the list of requirements is incomplete, thus failing to cover the user's requirement space and lacking in requisite specificity. Furthermore, referents frequently are not specified for any of the required representations. In a quest for supporting documentation, the V&V agent is likely to find that there is no conceptual model (and thus, no record of agreements and compromises between the user and developer), and the documentation on any verification and validation done during development is incomplete at best.

The V&V agent is often left with the task of refining a weak set of requirements and engaging the user in determining a potential set of referents. We developed a prioritization framework to help users determine the relative importance of representations and the *behaviors* contained in the refined requirements in light of the type of conditions under which a simulation was intended to be used. The goal was to allow users to make judicious choices as to which representations and behaviors had to undergo V&V now and which could be left to a future date, based on cost and priority.

Prioritization Framework

Asking a user about requirements and referents often ends in consternation on the part of both the user who never thinks in those terms and the V&V agent who works with them daily. The prioritization framework employs a scenario-based approach that allows the user to express requirements by examining the intended uses of the simulation and determining their relative importance. The importance could be based on frequency of use or on components of the scenario that cannot be tested in real-world exercises. The evaluations by the user are expressed in a probability tree and result in a weighting factor for each intended use. The V&V agent uses the same scenario descriptions, but focuses on each representation (an environmental factor, weapon, sensor, etc.) to determine how critical each specific representation is to the execution of the scenario and intended use. The final evaluation of priority uses both the user's weighting factors and the determination of criticality to compute the final priority. The computation is done using a spreadsheet. The user can easily change any of the probabilities and recompute the prioritization in minutes. The result is a simple system readily explained in terms the user understands and can defend. The following sections illustrate the use of the framework.

Scenarios

Our user had three significant missions for which the simulation was to be used: protection of a warehouse facility, defense of a convoy, and protection of a distribution center. By walking through each of these scenarios, the user was able to identify essential representations in the scenario and conditions under which the scenario would take place (time of day and weather conditions). While all three missions were assessed in the application of the prioritization framework, the following explanations show only two: protection of the warehouse and defense of the convoy. Additional missions can be added as branches on the probability tree.

Weighting Factors

The prioritization framework uses a tree structure virtually identical to a probability tree where the starting point is the set of problem scenarios, which are represented as mission areas in our example (Figure 1). Each mission area is assigned a percentage based on its importance or frequency of use (according to the user's preference). The percentages are used to assign a weighting factor between 0 and 1.

The second tier uses time of day. For the example shown, only day and night were used; however, it would be possible to use day, night, dawn, and dusk as each of these times presents unique lighting conditions that affect sensors. The third tier refers to weather condition: clear, rain, or snow. At each tier, the sum of the weighting factors assigned within that tier must sum to 1.

Criticality Rating

Assisting the user in setting the weighting factors is the first step. The second step consists of examining the representations and behaviors and determining whether they are important or useful for any scenario at the specified time of day and weather condition. For simplicity, the following numerical assignments were made: 2 for critical to use at that time and under that condition, 1 for occasionally used at that time and under that condition, and 0 for not needed at that time and under that condition.

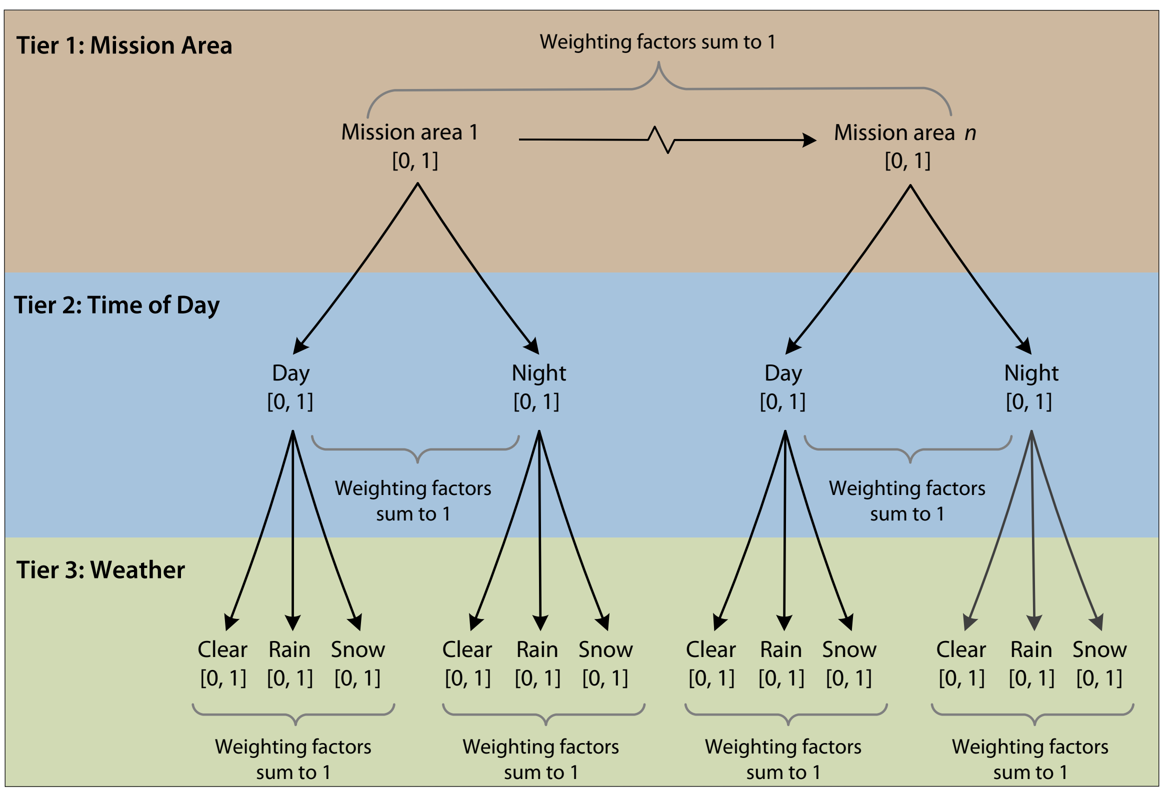

Figure 1. Probability Tree Structure Used to Determine Weighting Factors

Evaluation Matrix

As an example, consider a problem in which the user is examining the defense of goods being held at a warehouse and the protection of those goods when being transferred by convoy to a distribution center.

The user determined that for the initial set of evaluations, the warehouse mission area would be exercised about 60 percent of the time, and the convoy mission area, about 40 percent of the time. Attempts to attack at night represented the most significant threat to the facilities, leading to a weighting factor of 0.7 for the warehouse mission area at night. Convoys, on the other hand, are rarely planned for night, but extenuating circumstances might make it necessary to extend the duration of a convoy into hours of darkness. Thus, some weighting was placed on night operations for convoys (0.2).

Weather conditions were then assigned for day and night. The user expected to exercise the simulation for the warehouse half the time for clear weather conditions during daylight hours, resulting in an environmental weighting factor of 0.5 for day, and less frequent use of the simulation during inclement weather (weightings of 0.3 and 0.2, for rain and snow, respectively). At night, however, the more dangerous conditions at the warehouse facilities were operations during rain when the effectiveness of some of the sensors protecting the site would be reduced. In this case, rain was therefore assigned a weighting of 0.5, while clear weather and snow were assigned lesser values (0.3 and 0.2, respectively). In this manner, the user was able to provide weighting values for the probability tree in Figure 1 by considering the circumstances under which the simulation would be used to provide assessments.

Convoys are typically planned for daylight hours with a limited number having to extend into nighttime operation; therefore, the convoy's weighting factor for day was set at 0.8 with a corresponding value of 0.2 for night. Similarly, convoys are planned for clear weather, although daytime operations can more readily tolerate rain than can nighttime operations. Thus, the weighting factors assigned for daytime operations were 0.6 for clear weather, 0.3 for rain, and 0.1 for snow. Nighttime operations are less tolerant of inclement weather; therefore, the weighting factors set were 0.8 for clear weather and 0.1 each for rain and snow.

The evaluation matrix in Figure 2 shows these weighting values in the top four rows, which are color coded to correspond to the tiers of the probability tree in Figure 1. The cumulative weighting factor is determined by multiplying the three weighting factors (mission area, time of day, and weather) for each branch of the tree. In Figure 2, these cumulative weighting factors are found in the fifth row, the first white box under the colored rows, which represent the individual weighting factors.

The next step was conducted by the V&V agent's technical experts and involved the assignment of a criticality rating (2, 1, or 0) to each combination of mission area and condition. To illustrate this process, Cumulative weighting factor

Criticality rating

Individual priority score (product of cumulative weighting factor and criticality rating)

Overall priority for a given representation considering all scenarios (sum of individual priority scores across all scenarios)

		Warehouse = 0.6						Convoy = 0.4						
		Day = 0.3			Night = 0.7			Day = 0.8			Night = 0.2			
		Clear	Rain	Snow	Clear	Rain	Snow	Clear	Rain	Snow	Clear	Rain	Snow	
		0.5	0.3	0.2	0.3	0.5	0.2	0.6	0.3	0.1	0.8	0.1	0.1	
Category	Requirement	0.09	0.05	0.04	0.13	0.21	0.08	0.19	0.10	0.03	0.06	0.01	0.01	
Environment	Berm	2	2	2	2	2	2	1	1	1	1	1	1	
	Vegetation: grasses	2	2	1	2	2	1	2	2	1	2	2	1	
	Precipitation	0	2	2	0	2	2	0	2	2	0	2	2	
Sensors	Light intensification devices	0	0	0	2	2	2	0	0	0	2	2	2	
	Passive IR devices	1	1	1	2	2	2	1	1	1	2	2	2	
														SUM
Environment	Berm	0.18	0.108	0.072	0.252	0.42	0.168	0.192	0.096	0.032	0.064	0.008	0.008	1.6
	Vegetation: grasses	0.18	0.108	0.036	0.252	0.42	0.084	0.384	0.192	0.032	0.128	0.016	0.008	1.84
	Precipitation	0	0.108	0.072	0	0.42	0.168	0	0.192	0.064	0	0.016	0.016	1.056
Sensors	Light intensification devices	0	0	0	0.252	0.42	0.168	0	0	0	0.128	0.016	0.016	1
	Passive IR devices	0.09	0.054	0.036	0.252	0.42	0.168	0.192	0.096	0.032	0.128	0.016	0.016	1.5

Figure 2. Evaluation Matrix

we first separated prioritized values into categories of representation types: weapons, platforms, sensors, human behaviors, and environmental factors. Every representation was evaluated for each scenario; however, for purposes of illustration, we selected requirements from two different categories of representation. We chose three environmental representations: berms, vegetation in the form of grass, and precipitation. Note that precipitation here refers to the model requirement to represent precipitation and its effects, while rain and snow are the weather conditions under which the model is expected to be used. We also chose two representations from the sensor category, light intensification devices and passive infrared (IR) devices.

Berms are built as barriers against incursion by unwanted visitors. They are not likely to be found along a convoy route, but they might be present as artifacts of prior events. For example, fortifications built along coastal roads on both the East and West Coasts of the United States for fixed gun batteries have defensive features in common with berms and could be represented as such. Thus, berms have a criticality rating of 2 for the warehouse mission area and 1 for the convoy mission area, across all conditions. Grasses are found around both warehouse sites and along the roadside; however, during snow, they are likely to be weighted down and, hence, less important as potential cover for threats. Criticality for grasses is rated as 2 for clear and rainy weather and 1 for snow for both the warehouse and convoy mission areas. Precipitation is irrelevant for clear days or night and thus rates a 0 under clear weather conditions. Light intensification devices are used during low light conditions and are thus rated as irrelevant (0) during the day and critically important at night (2). IR devices may have some use during the day, but are critically important at night, as reflected by their scores of 1 and 2 for those circumstances.

Once the cumulative weighting factors and the critical factors are determined, the prioritization can be computed, first for each individual scenario and then summed for all scenarios to find the overall priority. For purposes of illustration, the representations used in the above computations are replicated below the grey row in Figure 2. The values represented in each of the cells in the lower part of the matrix result from multiplying the cumulative weighting factor by the criticality rating. The overall priority, shown in the column on the right labeled *SUM*, is the sum of all the values for that row—effectively the sum of the priority for that representation for each scenario summed for all scenarios.

The simulation has hundreds of individually listed representations, making it desirable to group them into categories of similar entities. (The term *entity* is used in simulation to mean the thing, behavior, or condition being represented.) The representations are listed in Figure 2 under the term *requirement* because these are the required entities. While all the representations could be listed in a single set in priority order using the values computed in the SUM column, the results made more sense when the ordering was done within the category.

The computational framework presented above lacks the ability to account automatically for interdependence among representations. For example, to use a weapon successfully, the actor in the simulation might have to be able to assume different positions and seek cover. If the use of the weapon had a high priority, the accompanying behaviors on the part of the actor would have to also have that high priority, even if seeking cover was not a high priority when considering the actor's behaviors in isolation. However, having all the representations ordered within their respective categories facilitates cross-category comparisons for detection of such interdependencies. The ability to determine interdependencies is important when resources for validation are limited. The cut-off points for investment have to include all the related representations needed for coherent operation.

Conclusion

The use of this prioritization framework is readily understandable from the perspective of the user and technologist, and it allows the user to establish needs in clear terms. The framework gives the user the ability to provide rational, repeatable, and documented evidence for decisions concerning where to focus V&V resources, thereby providing increased confidence that the simulations selected adequately portray the conditions appropriate to the intended use.

The prioritization framework presented here is easy to implement and is based on user needs and intended use of the simulation. While the method was developed to support a V&V effort directed toward potential acceptance of a simulation developed for other users, it can be adjusted for use in managing investment in any new simulation tool. Prioritization is not a definitive assessment, but a triage that can help the user determine the final selection of requirements to be validated for a given model. It also provides a rational, defensible, and repeatable process for choosing what to validate and what to leave out of the V&V process.

Reference

Department of Defense. 2007. *Modeling and Simulation Verification, Validation and Accreditation Recommended Practices Guide*. On-line Reference Guide, Washington, DC: Office of the Secretary of Defense.

S. K. (Sue) Numrich (second from right), a Research Staff Member in the Joint Advanced Warfighting Division (JAWD) of IDA's Systems and Analyses Center, holds a doctorate in physics from the American University.

Bob Zirkle (second from left), a Research Staff Member in the Strategy, Forces and Resources Division (SFRD) of IDA's Systems and Analyses Center, holds a doctorate in political science from the Massachusetts Institute of Technology (MIT).

Jim Ayers (far left), a Research Staff Member in JAWD, holds three master's degrees, the first in strategic studies from the Air War College, the second in air mobility from the Air Force Institute of Technology, and the third in business from Webster University.

Forrest Smith (far right), a Research Associate in SFRD, holds a master's degree in operations research from North Carolina State University.

Anna Vasilyeva (not pictured), a former Research Associate in the Information Technology and Systems Division of IDA's Systems and Analyses Center, holds two master's degrees, one in technology policy from the University of Cambridge Judge Business School and another in aeronautics and astronautics from MIT.

The full paper "Prioritization Framework: A Step Toward Cost-Effective VV&A," was published in the online proceedings of the 2016 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), held in Orlando, Florida, November 28–December 2, 2016.