An Approach to Simulation Effectiveness

Duarte Gonçalves
CSIR
PO Box 395
Pretoria, 0001
dgoncalv@csir.co.za

Abstract. Simulation is an important aspect of engineering complex systems. In the real world, numerous problems can prevent the effective use of simulation. This paper looks at the tough question: *When is a simulation effective*? How would we know? The context and purpose of simulation are important in answering the question.

If the simulation is viewed as a system, it follows that it has stakeholders and requirements originating from the creating system. An important result is that measures of simulation effectiveness include fidelity, time-to-answer, and resource usage. The importance of a referent (codified knowledge) in defining fidelity and related pitfalls are discussed. Simulation effectiveness assessment enables simulation designers to trade simulation effectiveness against cost and risk subject to constraints. A brief overview of how abstraction and simulation method selection can be used for this trade-off is given. The impact of simulation effectiveness on risk is discussed. The benefits are balanced simulations with risk that is better matched to the problem at hand.

Introduction

Modeling and simulation are essential for systems engineering and decision support. The modeling and simulation problems commonly encountered and a general solution are discussed. This background motivates the need to answer the tough question: *When is a simulation effective*?

The problem

Common modeling and simulation problems encountered in the engineering of systems and the provision of decision support in the real world are:

- Not enough time and money for simulation during an acquisition project (in the South African context)
- Heuristic system level design (trial and error)
- Model/simulation breadth/depth imbalance
- Value to the customer is only known when the system is delivered – this is too late.

Engineering is about making decisions. In all the problems listed above, the quality of decisions made based on models and simulations is compromised in some way. Undesirable outcomes follow poor decisions. The concept of decision quality has been considered in the context of technology management (Simpson 2002), but there does not appear to be any *integrated* work in the context of simulation.

The solution to these problems is complex, costs money, takes time and has three main parts:

- **Develop a relevant capability before acquisition commences.** Understanding and insight into new systems takes time. This requires management focus on an application area and a mandate from the customer or client. It also means that capabilities must be anticipated based on market gaps, technology trends, and research directions. *Models* *and simulations for a specific application area* ***and*** *the ability to build the models constitute the capability*.

- Ability to respond to acquisition projects with effective simulation.
- Updating the capability by feeding back knowledge gained from previous projects.

This paper focuses on the second part of the solution with the question: *When is a simulation effective*? In order to answer this question, the context and purpose of simulation must first be understood. In the remainder of the paper a framework for assessing simulation effectiveness is defined. From this framework, some requirements for the capability alluded to in the first part of the solution will also become evident. The relationship between risk and simulation effectiveness will be discussed before concluding.

The Context and Purpose of Simulation

Figure 1 illustrates the context of modeling and simulation. There are four systems in this diagram. Each has a different purpose:

- The creating system – satisfies a business need
- Research system – creates new knowledge
- Created system – satisfies the customer/client need.

The purpose of the *simulation system*, the fourth system, is typically:

- Effectiveness prediction – Answers the question: "Will the system work well enough?" and is a basic building block required for other tasks below.
- Validation of requirements and other types of validation
- Trade-off studies – Choosing from a set of alternatives
- System design and requirements analysis – determining the system parameter set that achieves a certain system effectiveness
- Robustness analysis, and
- Risk analysis.

In this paper the simulation system is the system of interest. The referent is a codified body of knowledge – this topic will be revisited in the next section. The research system is primarily focused on creating the referent by generating new knowledge and using existing literature. This is about building the capability. Modeling and simulation then supports the creating system in the purpose just described and could also support the research system although this has not been indicated in Figure 1. As the creating system operates, new knowledge becomes available through validation or from previous projects and the referent is extended.

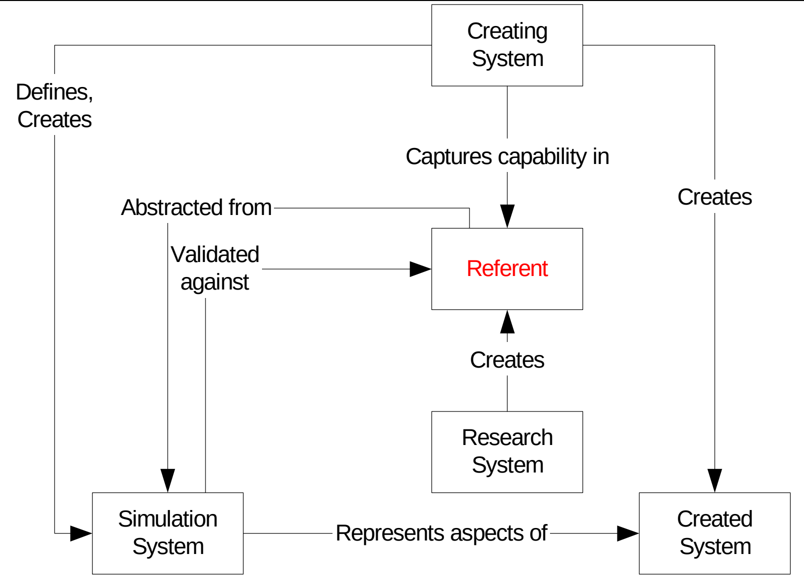

Figure 1 Context of simulation

Simulation is indicated as a system in Figure 1. It must satisfy the needs of the creating system. Specifically, simulation must satisfy certain effectiveness requirements. This also means that the simulation system has a life-cycle, sometimes extending over several decades. Consequently, system engineering life-cycle processes, for example as defined by ISO15288, are relevant.

The Referent

There are several definitions of a referent in the context of modeling and simulation. Two such definitions are considered in this paper from two points of view:

- a referent for a generic capability, and
- a referent for validation and fidelity assessment.

On presenting such definitions it becomes clear that there is a circular relationship between the referent, the model, and fidelity (introduced in the next section). Since the simulation system interacts with the referent, the criteria for identifying, selecting and specifying a referent are considered briefly.

In the context of developing a generic capability we are concerned with using the referent as the *source* from which models are abstracted and validated in general. The following definition will be used, (based on Pace 2004):

Referent: A codified body of knowledge about the thing being simulated.

A more specific definition has been selected for use in the context of validation and fidelity assessment for a specific application, (also based on Pace 2004):

Referent: The referent is the best or most appropriate codified body of information available that describes characteristics and behavior of the reality represented in the simulation from the perspective of validation (or fidelity) assessment for the intended use of the simulation.

'Information' could include data, theories, results from other simulations (preferably validated simulations), human expert knowledge, etc. This information is appropriate if it has the right accuracy, scope, depth, context and cost for the intended purpose. Describing a reality that does not exist, such as an unprecedented system, may require several iterations to define and validate the referent as new information becomes available. There is thus a relationship between iterations of the referent and technology readiness levels.

The fundamental assumption is that there is a referent for modeling and simulation. Models can only be validated against the referent and not the real world, as illustrated in Figure 2. Both the referent and the model are representations of the real world and are incomplete to some degree. To illustrate this, consider validating a model against the real world by measurement. The process of making a measurement creates a new representation or model of the real world that is not complete. Herein lies the potential limitation with the concept of a referent. Since the referent is also a model, it has a certain fidelity. In order to evaluate fidelity, one needs to have a referent. The problem is defined recursively. The difficulty of measuring fidelity lies in the fact that fidelity is a relative measurement. There is no absolute measure of fidelity, with reality as the 100% mark. Currently, the only way out of this predicament is to reach consensus on what constitutes an acceptable referent for a particular purpose. It should also be remembered that a referent intended for fidelity assessment *need not have the same* scope and depth as a generic referent intended as a source for abstracting models. For many complex systems the cost of performing extensive measurements is prohibitive and maintaining an appropriate referent may be a cheaper option.

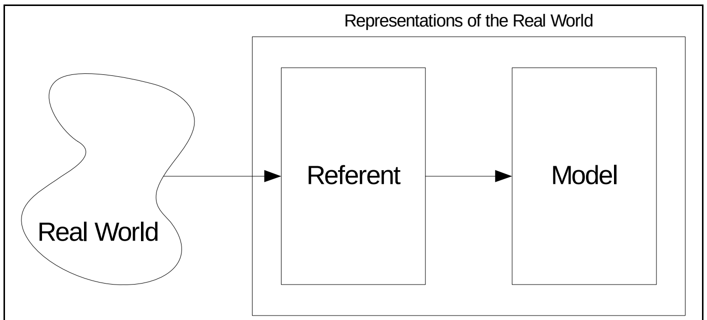

Figure 2 The Referent: bridge between the real world and models (Schricker et al. 2001)

Identifying a referent depends on the modeling and simulation requirements, i.e. which entities, interactions and environments are modeled, and the intended use. A referent may be selected based on (Pace 2004):

- Convenience (availability and accessibility),
- Cost, or,
- Proxy, i.e. for a system that does not exist, use knowledge of similar systems

In some cases the stakeholder might specify the referent. Any potential referent must be assessed in terms of (Pace 2004):

- Scope – the breadth of the parameters, elements, interactions or applications over which the model is applicable,
- Depth – the level of detail required, and
- Context – the conditions under which referent information is applicable. This can be conditions under which data is measured, assumptions or physical conditions.

These issues must extend to cover the multitude of intended uses within an application area when creating a referent as part of a capability. Specifying a required referent for a specific application would consider the following, in addition (Pace 2004):

- Domain Coverage – The domain required by a certain application or intended use in terms of parameters and underlying conditions. Coverage is the overlap of the referent and the extent required by the domain. For unprecedented systems, where a referent is selected by proxy, there may be little or no overlap.
- Attributes – The attributes contained in a referent which are relevant to the intended use or application.
- Parameter Uncertainty – The uncertainty of parameters contained in the referent.

Having identified and specified the referent one can move to defining simulation effectiveness.

Defining Simulation Effectiveness

The quantitative measure of simulation effectiveness is critical in understanding the quality of the decisions that can be made based on simulation results. Furthermore, simulation trade-offs can be performed at the simulation system level, illustrated in Figure 3. The creating system defines the simulation requirements for the simulation system. Models are abstracted from a referent. The simulation, and hence any trade-offs, are subject to constraints. For a balanced simulation, a trade-off space for large simulations must consider (Felix 2004):

- Effectiveness,
- Cost, and
- Risk.

Effectiveness in this context is the ability of the simulation system to satisfy the needs of the creating system. Although the focus in this section is on simulation effectiveness, does not mean that other types of simulation requirements are not important. Because a simulation is specialized software, i.e. it is software that represents the created system, the effectiveness is also specialized. Typically, simulation effectiveness is decomposed by:

- Fidelity,
- Time-to-answer,
- Resource usage, and
- Other application specific measures of effectiveness.

Simulations fundamentally support decisions. Thus, *fidelity* is the effectiveness aspect relating the validity of information used for decisions. *Time-to-answer* is an indication of the relevance of information and *resource usage* the efficiency in obtaining information from the simulation. The referent is the reference for assessing fidelity.

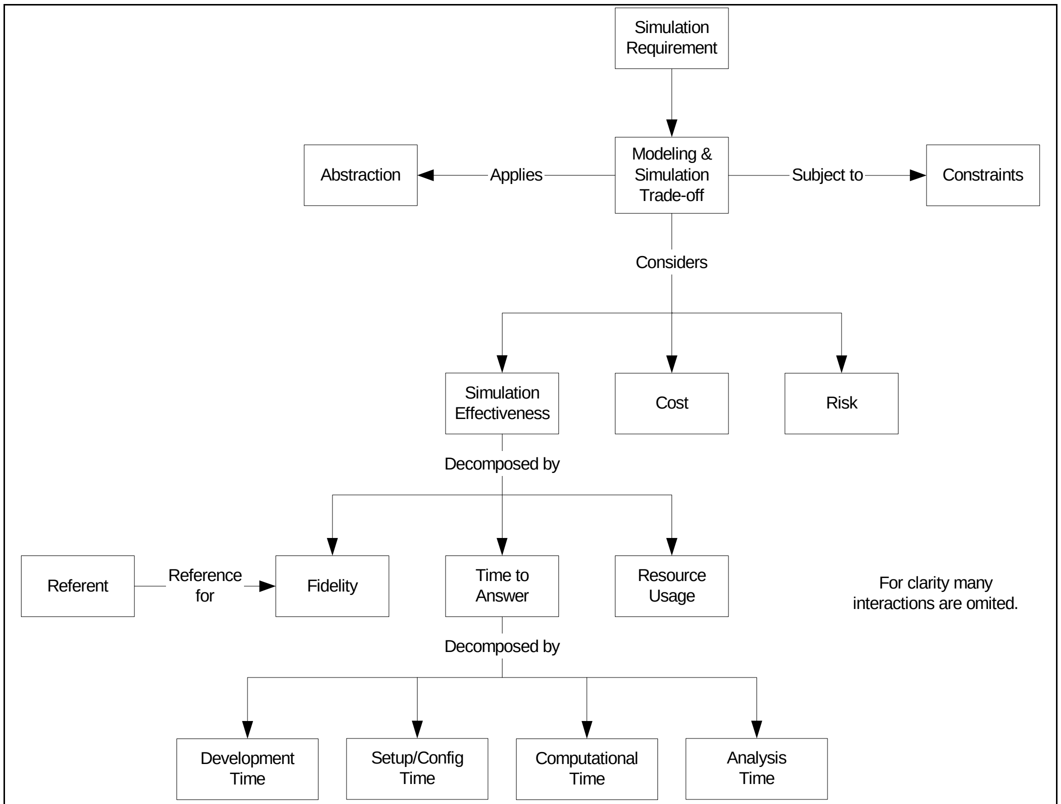

Figure 3 Modeling and simulation trade-off space

Simulation effectiveness is important because it is a framework within which the systems engineer can specify how well the simulation must perform. It also provides the criteria against which the simulation can be assessed. For safety critical applications where safety analysis or survivability analysis are required, for example, simulation effectiveness is an important simulation quality measure.

The three components of simulation effectiveness, fidelity, time-to-answer, and resource usage, are described in more detail in the following sections and an example is presented to illustrate some of the concepts.

Fidelity

A definition of fidelity is crucial to any discussion and common understanding of this concept. Ideally such a definition is theoretically sound and practically useful. The following definition is from (Gross 1999):

*Fidelity***:** *1. The degree to which a model or simulation reproduces the state and behavior of a real world object or the perception of a real world object, feature, condition, or chosen standard in a measurable or perceivable manner; a measure of the realism of a model or simulation; faithfulness. Fidelity should generally be described with respect to the measures, standards or perceptions used in assessing or stating it. 2. The methods, metrics, and descriptions of models or simulations used to compare those models or simulations to their real world referents or to other simulations in such terms as accuracy, scope, resolution, level of detail, level of abstraction and repeatability. Fidelity can characterize the representations of a model, a simulation, the data used by a simulation (e.g., input, characteristic or parametric), or an exercise.*

It is essential to emphasize that fidelity must relate to the purpose of the simulation. Thus high levels of detail do not imply high fidelity, when it is not related to the purpose of the simulation. These definitions also suggest that it is possible to quantify fidelity numerically.

The approach to exploring the dimensions of fidelity for a new problem is (Gross 1999):

- Enumeration of entities in terms of scope (breadth) and depth (level of detail)
- Identify significant relationships between these entities (Beware of making implicit assumptions of independence), and
- Identify contributing factors, which could include materials and components used, algorithms, and parameters related to system measures of effectiveness.

Assessing fidelity is based on the concept of differential fidelity proposed in (Gross 1999) and illustrated in Figure 4. The fundamental assumption is that there is a fidelity referent relating to an application area. A model is abstracted from the referent possibly using methods suggested in Figure 8. The creating system defines the required fidelity based on the intended application. For quantitative fidelity measures this can be a *tolerance* on accuracy, error, resolution or uncertainty. The model fidelity is the difference between the referent and the model under assessment. There are a few subtleties here, however. When assessing fidelity, the purpose is not to evaluate problem variation but to evaluate model deviation. Where a model parameter has uncertainty relating to its measurement then this will translate to model fidelity. If the model fidelity is within tolerance, then it is *valid and fit for the intended purpose*. Fidelity will be assessed under specific test conditions as required by the problem at hand. However, without adequate coverage, i.e. evaluating over the model input space, the fidelity assessment may be misleading.

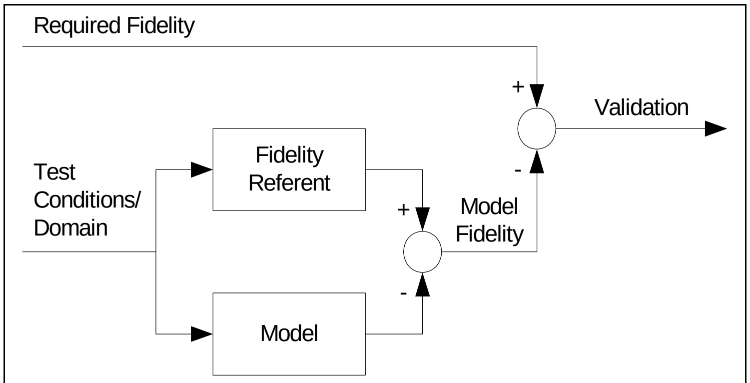

Figure 4 Differential fidelity model

The fidelity is assessed on the criteria illustrated in Figure 5. A qualitative and a quantitative approach to fidelity are essential. Fidelity measures the level of abstraction. A qualitative approach is based on the *existence* of inputs, attributes or characteristics in a model which may have been abstracted from the referent. A hierarchy of models is possible for physical, structural, behavioral and functional models (Figure 6). The simplest model is at the top of the hierarchy with increasing refinement down the hierarchy. The existence of a certain level in this hierarchy indicates the fidelity of a specific model.

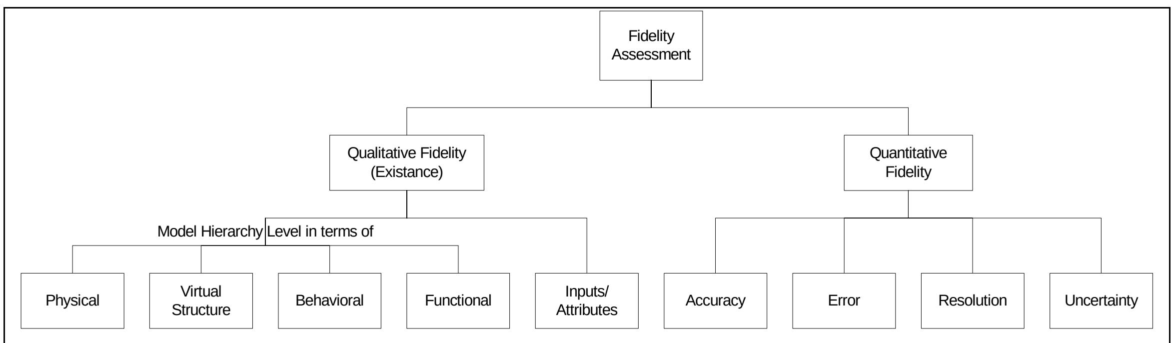

Figure 5 Assessment of qualitative and quantitative fidelity

For quantitative fidelity, measures or metrics can be defined for accuracy, error, resolution, or uncertainty, depending on the problem at hand. In the case of error, a typical metric might be mean square error.

The concepts presented are illustrated using a digital elevation map (DEM) as an example. The resolution of the model is the ground sampling distance, typically in meters. This might vary from 1000m to 30m. This limits the ability of the model to represent rough terrain. The model may have measurement uncertainty originating from the instrument used to collect the data. The larger the instrument uncertainty the less precision the model will have. In addition the data may be rounded to the nearest meter, resulting in quantization uncertainty.

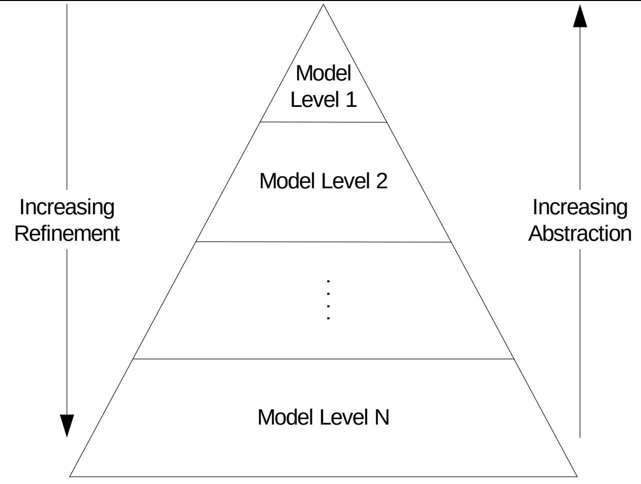

Figure 6 Model hierarchy

Apart from difficulties in defining a referent, another potential problem is in propagating the fidelity for the entire simulation down to sub-models. This is necessary in order to specify these sub-models. An obvious approach is one based on sensitivities - this is discussed in (Gross 1999), section 3.2.

Time-to-Answer

Time-to-answer is the total time from when a question is asked to when the answers are available during the acquisition phase at the required fidelity level, assuming resources are fully available. It is a measure of the ability of the simulation capability to support decisions in a timely manner. Without considering time-to-answer, one might be tempted to build models at a very high level of detail and breadth. The components of time-to-answer are:

- Development time (during acquisition)
- Simulation setup/ configuration time
- Computational time, and
- Post-processing analysis time.

When there are new personnel using existing simulations, an additional component of time-to-answer may include time-to-first-use by an individual, given a certain education/experience level. If the time-to-first-use is too high relative to the perceived benefit, new personnel will not make use of the simulation tool. Similarly if the time-to-answer is too high relative to the perceived complexity, the simulation as a whole may be abandoned.

Resource Usage

The use of resources indicates the efficiency of the simulation. In this context time is not considered as a resource. Resource usage is also important from the point of view of specifying or quantifying the modeling capability. Typical resources that might be considered are:

- Computing which includes number of processors and their processing speed, volatile and non-volatile storage space, and
- Level and extent of skill (manpower and personnel issues).

The resources impact directly on cost, but may be constrained separately from cost.

An Electro-optical Surveillance Example

Consider for example an electro-optical surveillance system. An overview of the generic referent is illustrated in Figure 7. Each category has one or more models and associated parameters, papers, books and a person competent on the topic. The surveillance system includes a human observer (class 4.3) which makes modeling more challenging, especially in evaluating measures of effectiveness such as detection, recognition and identification range within the surveillance context. The surveillance system (class 4) observes an entity, which could be a target, in the environment (class 3). This referent is incomplete in some ways because of the project scope. For example, it does not include an 'Induced Environment' under environment nor does it consider the counter-surveillance problem, i.e. where the surveillance system becomes the entity being observed. The implied context of the referent is tactical surveillance in the wavelength range 0.4 - 12µm.

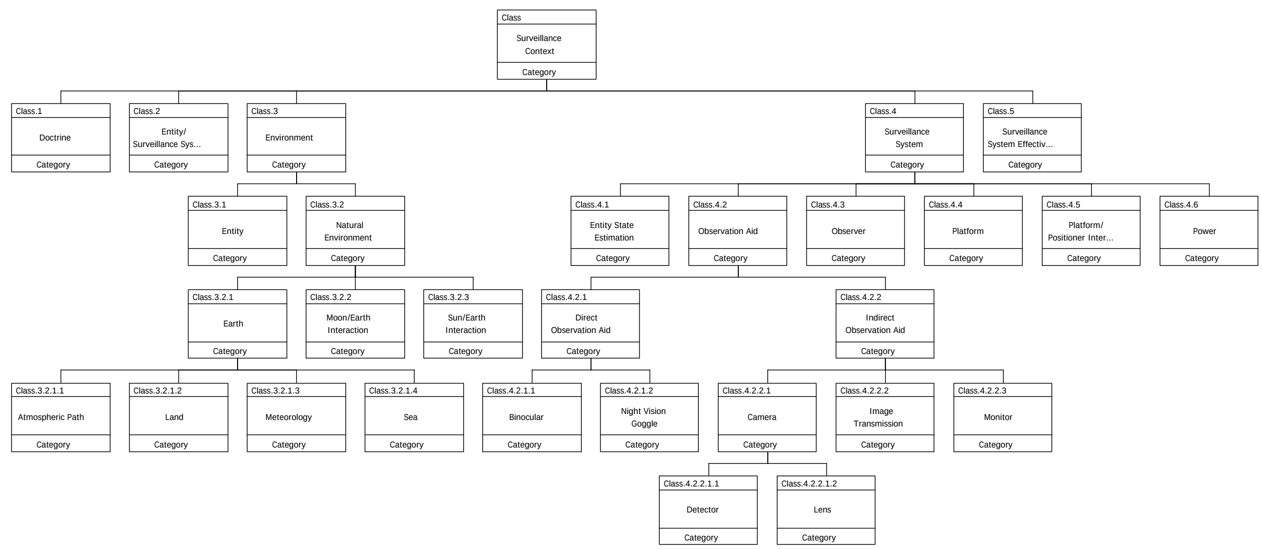

Figure 7 Example: Electro-optical Surveillance system referent overview.

A question typically asked about electro-optical surveillance systems is "How far can we 'see' with this system?" Upon further probing into this question one might find that the operator would like to know the detection range of a night vision goggle (NVG) used for driving. This is essentially effectiveness prediction. Suppose that the task being performed by the operator only requires knowledge of detection range to within ±10%. This defines the quantitative fidelity. From the referent in Figure 7 a model can be abstracted to determine detection range. For example, sub-models relating to the sun and sea are not relevant. A qualitative fidelity description would include sub-models relating to the moon, NVG and objects of interest. The more difficult questions are "Do we need an atmospheric path model (a sub-model of class 3.2.1.2.)?" and "Is a terrain model (a sub-model of class 3.2.1.2.) required?" The answer to these questions is not obvious. However, these are already covered by the quantitative fidelity requirement. Thus it is necessary to evaluate the impact of the atmosphere and terrain on detection range to determine whether these sub-models should be included.

Assessing quantitative fidelity of such a surveillance system is not a trivial matter. Firstly, a number of trained observers are required. Secondly, cost will normally permit evaluation over a small number of conditions such as weather, terrain, etc. These points may not always be representative of the set of conditions over which the system must operate over its life-cycle. Good experiment design is required to control or measure parameters which could impact the detection range.

Fidelity assessment can only be done reliably when a good understanding of the application has been achieved.

Simulation Trade-offs

In order to develop a balanced simulation, this section considers methods for trading fidelity, time-to-answer and resource usage. The most important class of methods is abstraction. The following definition of abstraction is adapted from (Gross 1999):

Abstraction: The process of selecting the essential aspects of a system to be represented in a model or simulation while ignoring those aspects that are not relevant to the purpose of the model or simulation.

A taxonomy of model abstraction techniques is presented in Figure 8 based on (Frantz 1995). Each one of these represents a method for trading off simulation effectiveness, cost and risk.

Careful consideration of simulation methods used to calculate the required statistics can reduce time. For example, under certain conditions it could be easier to use the expected value than to do a Monte Carlo simulation. A cost function dependant on fidelity, computational time and resource usage can be used to 'optimize' the level of abstraction.

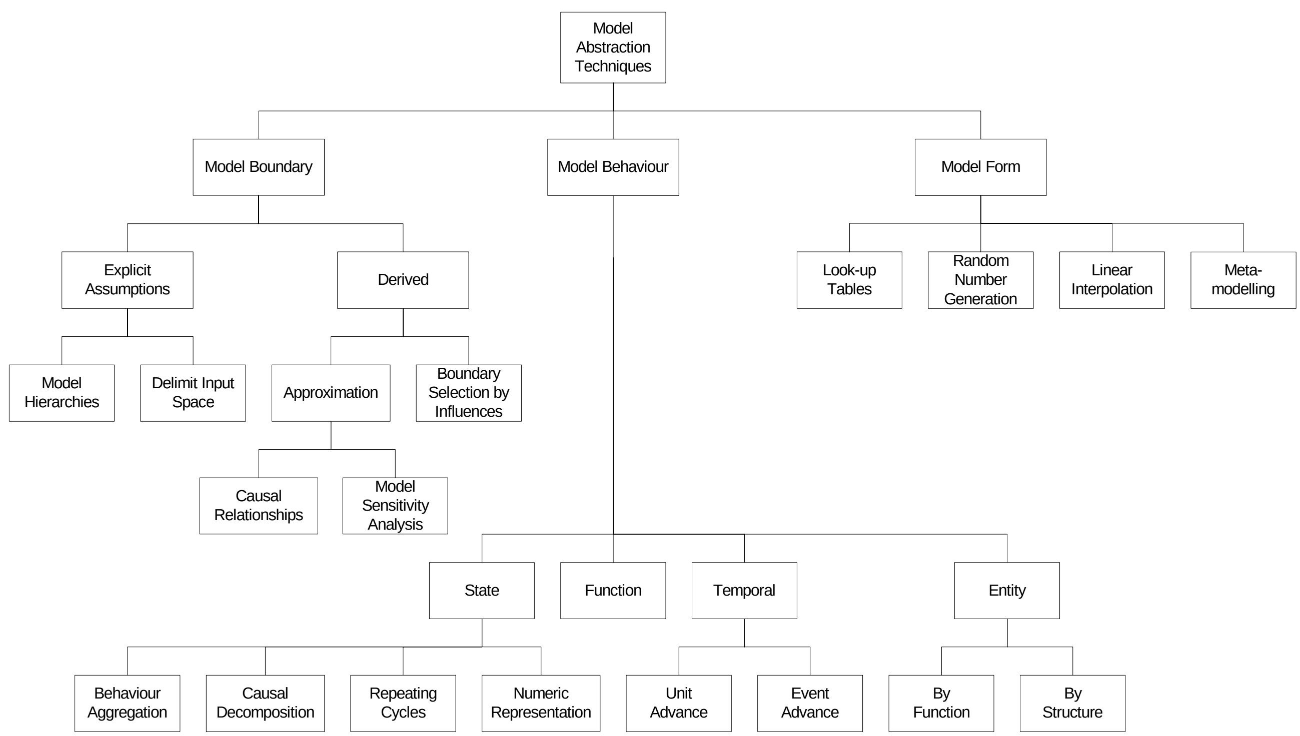

Figure 8 Taxonomy of Model Abstraction Techniques (Frantz 1995)

Relationship between Risk and Simulation Effectiveness

Simulation risks arise from inappropriate fidelity, a time-to-answer that is outside schedule constraints, inadequate resource levels, or cost overruns (refer to Figure 3). In this discussion, the consequences of a risk are limited to the simulation system. These will however impact the creating system in terms of decisions made and the created system in terms of physical consequences.

Consider total risk as a function of model fidelity shown in Figure 9 and given an infinite time. The acceptable fidelity or required fidelity of Figure 4 defines the modeling and simulation stop criterion. Models for a specific application need a certain level of *relevant* detail to achieve an acceptable fidelity. A certain amount of residual risk remains due to, for example, an imperfect referent, or an incorrectly selected referent. If fidelity is not measured or the question to be answered is not kept in mind, then excessive detail is added that is not relevant to the model purpose. Using incorrect levels of abstraction is one cause of this problem. However, detail beyond a certain point does not increase fidelity. This detail will certainly increase total risk because we expose ourselves to schedule overruns and additional cost.

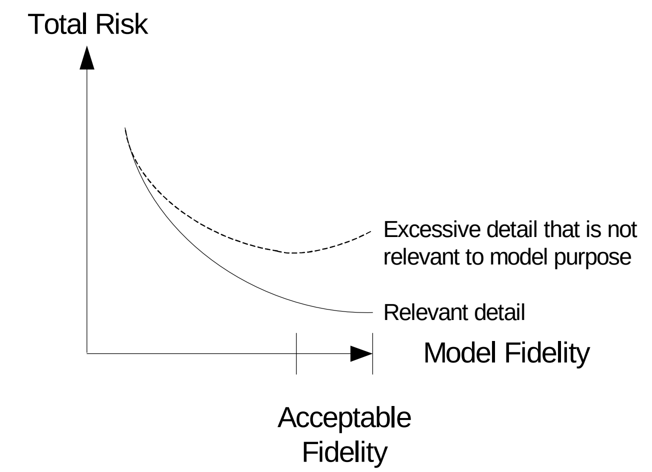

Figure 9 Total risk, fidelity and model detail

We now investigate the capability, which includes the referent and access to resources, assuming that only relevant detail is considered in the simulation.

The achievable fidelity as a function of time-to-answer is presented in Figure 10. Again, there is a range of acceptable fidelity for the intended purpose. Time-to-answer is almost always constrained by schedules in the real world. The 'Total risk' axis of Figure 9 is the third axis extending out of the page in Figure 10. However in Figure 10, risk will be indicated in color. Three risk levels are identified:

- **Low risk** – The fidelity is within the required tolerance for the intended application, simulation work can stop.
- **Moderate risk** – The fidelity is approaching an acceptable level, but it is not there yet.
- **High risk** – The fidelity is far from an acceptable level.

The risk levels increase relative to the model fidelity as time-to-answer approaches the schedule constraint. In other words, if the model fidelity is low at the start of the simulation project,

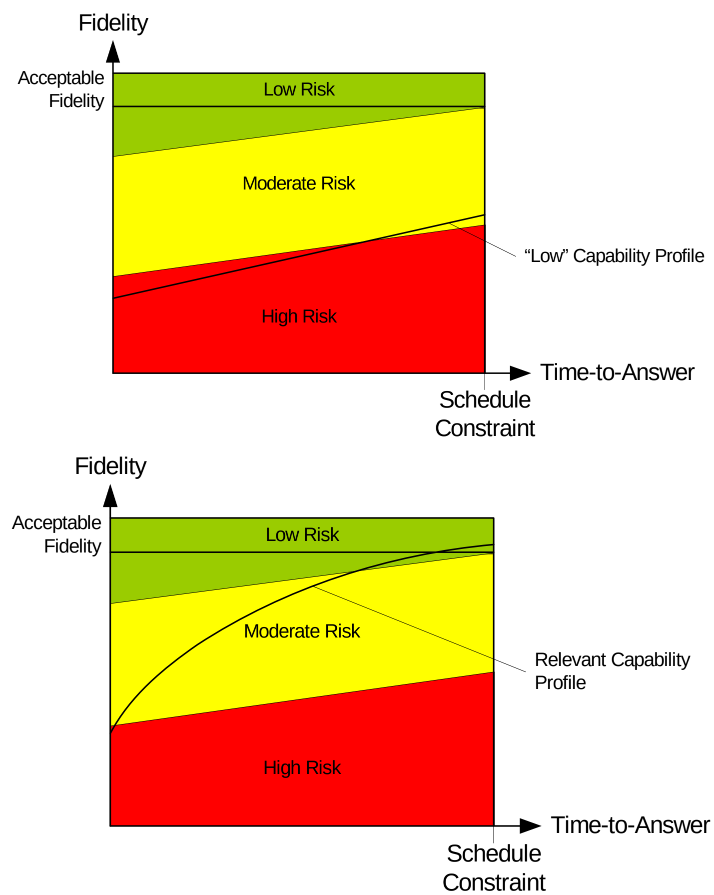

Figure 10 Conceptual capability/risk profiles for a given fidelity and time-to-answer. Top: "Low" capability, Bottom: relevant capability.

because one has time to do more work, this represents a lower risk compared to the same fidelity reached at the deadline. The risk levels used here are illustrative and are not intended to convey exact relationships.

A conceptual capability/risk profile is shown for a 'low' capability and a relevant capability. For a 'low' capability we start with a lower fidelity baseline than for a relevant capability. As time-to-answer is increased, so fidelity increases. For a 'low' capability the acceptable fidelity range is never reached within schedule constraints. When new questions about a complex system are asked in an application area where a relevant capability exits, the time-to-answer is not zero but the questions are more likely to be answered within schedule constraints. Thus, having a relevant capability compresses time-to-answer for a specific question.

Conclusions

The simulation effectiveness framework presented is a step towards quantitative evaluation of a simulation and answering the question: *When is a simulation effective*? The relationships between fidelity, time-to-answer, resource usage, cost and risk were developed. By trading these, a better simulation balance can be achieved leading to higher decision quality.

Other insights arise relating to the capability required for effective simulation. The need for two types of referents was shown:

- one which is part of a generic capability used as source for abstracting models, and,
- one for validation/fidelity assessment for a specific application.

It would seem that the referent in Figure 1 is actually a component of a Knowledge Management System (KMS). There may be several benefits to having a KMS in place, such as reduced time-to-answer when using existing knowledge and possibly reduced long term cost. This is a topic for further research. If the referent is missing, the project may need to be managed differently because of the risk profile. Such a project is in the realm of research and not systems engineering. However, difficulties remain in defining and assessing the required breadth and depth of the referent.

References

Felix, A., "Standard Approach to Trade Studies", *INCOSE 14th Annual International Symposium Proceedings*, 2004.

Frantz, F. K., "A Taxonomy of model abstraction techniques", *Proc. 1995 Winter Simulation Conference*.

Gross, D.C., Redactor, "Report from the Fidelity Implementation Study Group", 99S-SIW-167, *1999 Fall Simulation Interoperability Workshop Papers*, 1999.

Pace, D.K., *The Referent Study Final Report*, https://www.dmso.mil/public/transition/vva/evolvingconcepts, 2004, Accessed June 2005.

Schricker, B.C., Franceschini, R.W., Johnson, T.C., "Fidelity Evaluation Framework", *Proc IEEE Annual Simulation Symposium*, 109-116, 2001.

Simpson, J.J., "Innovation and Technology Management", *INCOSE 12th Annual International Symposium Proceedings,* 2002.

Acknowledgments

I would like to thank (in alphabetical order) Carel Combrink, Derrek Griffith, and Alwyn Smit for patiently discussing and reviewing the material which appears in this paper. I also acknowledge the useful feedback from the anonymous INCOSE reviewer.

Biography

Duarte Gonçalves holds a B. Eng in Electronics and a M. Eng in Computer Engineering. He has extensive experience in electro-optical systems, ranging from modeling the environment and electro-optical observation systems, to signal and image processing. He holds a full patent in the area of imaging spectrometers. He has also been involved in engineering surveillance systems for the South African DoD. More recently, he has consulted to the Karoo Array Telescope project, the South African technology demonstrator for the Square Kilometer Array (SKA). In 2005 he won the best paper prize at the INCOSE SA conference.