Purpose-Driven Metadata for Verification and Validation of HSCB Models

Alicia I. Ruvinsky, Andrew Shilliday, Janet Wedgwood, and John Welsh
Lockheed Martin Advanced Technology Laboratories

ABSTRACT
From inception, to implementation, to utilization, a computational model’s lifecycle information is studied and identified with contributing theories, responding metrics, evidential data sets, and robustness testing. Capturing these facts as metadata elements may facilitate invaluable mechanisms, such as model verification and validation, supporting effective utilization and trust of HSCB models. This work explores the purposeful identification of model metadata for the support of verification and validation (V&V) techniques of HSCB models. Capturing information that supports model V&V enables mechanisms such as model comparison, federation, combination, and corroboration. The purpose-driven metadata identification will provide an infrastructure for model comparison that will ultimately support the identification of (1) families of models producing output that may be corroborative and (2) combining models so the output from one model produces input to another model. Both of these techniques will contribute to the effective federation of models. In a system of federated models, the user may be unaware of the existence or functionality of an appropriate model set, but the system, upon characterizing the problem space and data requirements, may identify models and produce useful results.

PRIMARY TRACK
Valid Model Use and Validation

SECONDARY TRACK
Hybrid Models

DESCRIPTION
Metadata for HSCB models can support tools and users for model validation and verification (V&V). Model developers can use metadata during model development to make a model’s assumptions about problem domain, scope, and theoretical grounding explicit and transparent, supporting reflection during design time, when it is easier to correct problems. Model V&V staff can use metadata to capture results of model validation and verification tasks such as tests, results, data sets, etc. Model consumers can use metadata to find and select models relevant to their operational needs, taking into consideration the model’s intended purpose(s), prior validation, prior usage and performance, etc. Metadata can be used to match data sources to model inputs. Using metadata can help assess models for their compatibility, interoperability, and composability.

The processes and challenges for V&V of models and simulations of physical systems have been studied. In general, model validation determines “the degree to which a model … provides an accurate representation of the real world from the perspective of the intended uses of the model” [1]. Model verification determines “that a model implementation … accurately represents the developer’s conceptual description and specifications” [1].
The unique characteristics of human systems models present additional challenges, making many of the techniques applied to V&V of physical systems insufficient. For example, HSCB models lack “clear and generally accepted conceptual models” [2]. The underlying social science theories tend to be “softer”: e.g., often qualitative vs. quantitative, less formal (mathematical), often having many exceptions to the rule, etc. Thus validation against real-world data is harder for HSCB models. Predictive HSCB models often use historical data for validation. This can be useful (e.g., our DARPA ICEWS models were validated in this manner), but often insufficient since history presents a limited set of experimental conditions with one set of data points. Conversely, a model’s utility cannot be rejected because it does not resolve to an accurate point prediction as more relevant factors operate below the analytic horizon in HSCB models than in physical models.

Experience has provided an understanding of common pitfalls associated with metadata use for information (document) markup in operational settings. “One size fits all” (monolithic) approaches to metadata often fail. Vocabularies should be small, extensible, and utilized in flexible combinations. It is hard to achieve consensus as to what the metadata elements should be. Agreements (standards, etc.) must be modified as the world changes. Thus metadata standards for validation and verification must be (1) non-monolithic, (2) require little or no extra effort by information producers, curators, and consumers, and (3) be machine- and human understandable.

Verification of a model addresses issues regarding the model’s implementation with respect to functional and theoretical accuracy. Did the modeler implement the model correctly? How well are the modeler’s desired theoretical relationships manifested in the model?

Metadata may be used to verify a model’s mapping of real world variables to the theoretical space in line with the modeler’s intent. For a given model, the metadata will capture all theoretical relationships being represented by the model. For each theory, the metadata will capture the relationship between theoretical concepts being expressed and the real world variable proxies for measuring the theoretical concepts.

An example of a verification technique that uses purpose-driven metadata is auditing [3]. An audit is a type of informal verification technique where the adequacy of the model’s current utilization is based on established standards or guidelines documented for the model. In the envisioned metadata standard for HSCB models, the theoretical relationships captured in a model will be represented with the boundaries for which the theories have been experimented and tested. By reasoning over the metadata, the auditor may extract the tested boundaries of theories informing a model in question, as well as an assessment of the appropriateness of the current parameter specifications with respect to the tested boundaries, facilitating the audit process.

Metadata may be used to identify appropriate vs. inappropriate mappings between real world and theory. If a model has been used to predict behaviors in populations of children between five and ten years old, is it appropriate to use it to model adult males? The metadata for a theory will not only identify relationships between variables; it will capture studies supporting the theorized relationship such that real world variables representing theoretical constructs are documented and characterized. This way, the boundaries may be identified and tested by a model’s underlying theories.

Validation of a model assesses the model’s relevance in the real world. Does the model make correct claims about ground truth according to verified metrics? Did the modeler build a useful model? How well is the performance of the model mapping to patterns observed in the world? How robust are the generated outcomes of the model?
An example of a validation technique using purpose-driven metadata is the sensitivity analysis to assess model robustness. Sensitivity analysis is the systematic changing of model input variable values over a defined range where the model behavior produced is assessed and documented [3]. Sensitivity analysis is designed to identify sensitive model behavior variables, as well as the range of sensitivity. As sensitivity tests are performed, they are captured as model studies where parameter input settings output behavior are captured. Reasoning over these studies will expose variable and range sensitivities of the model.

Purpose-driven metadata will provide an infrastructure for model comparison, supporting the identification of (1) families of models producing corroborative output, and (2) combining models so the output from one model produces input to another model. A framework for managing theory and model metadata may enable both of these techniques, contributing to the effective federation of models. In a system of federated models, the user may be unaware of the existence or functionality of an appropriate model set, but the system, upon characterizing the problem space and data requirements may identify such models, producing useful results.

ACKNOWLEDGEMENTS
Work supported in part by ONR MESA program (N00014-10-C-0314), as well as the DARPA ICEWS program (FA8650-07-C-7749).

BIOGRAPHY
Dr. Alicia Ruvinsky is a Senior Software Engineer at Lockheed Martin Advanced Technology Laboratories. Before working with LM ATL in March 2009, Alicia completed her dissertation at University of South Carolina in computer science with a focus in multiagent systems. At LM ATL she has been involved in providing transparency for social science models in the DARPA funded ICEWS project. Alicia has also been developing a formal representation of social science theory, work that has been leveraged in the ONR funded MESA project developing a suite of tools for validation and verification of HSCB models.

REFERENCES