2007-03

Team 6: Joint Capability Metamodel-Test-Metamodel Integration with Data Farming

Beach, T.

Team 6: Joint Capability Metamodel-Test-Metamodel Integration with Data Farming

TEAM 6 MEMBERS

T. Beach, - Lead, Contact*
Joint Test and Evaluation Methodology, USA

D. Dryer, Ph.D.
H. Way
Joint Test and Evaluation Methodology, USA

S. Sanchez, Ph.D.
Naval Postgraduate School, USA

W. D. Kelton, Ph.D.
University of Cincinnati, USA

LTC J. Schamburg, Ph.D.
TRADOC Analysis Center-Monterey, USA

D. Martin,
EBR Inc., USA

INTRODUCTION

US adversaries are continuously seeking new ways to threaten US interests at home and abroad. In order to counter these threats, now more than ever, commanders must seek to leverage existing and emerging joint capabilities effectively in a variety of unique contexts. Achieving mission effectiveness in today's joint operational environment demands robust synergy among a wide array of mission-critical Service systems and capabilities.

Previously, forces and platforms were developed upon specific threat and scenario constructs. Requirements have often been developed, validated, and approved as stand-alone Service solutions to counter specific threats - not as participating elements in an overarching system-of-systems or joint capability.

Now, joint operations have become the mainstay of warfighting. The systems and capabilities warfighters employ must be tested and evaluated in a joint mission environment.

The challenge program managers and test organizations face is that effective testing and evaluation is becoming more difficult as individual platforms are increasingly being integrated into a complex joint mission system of systems. Ensuring that we test like we fight is the challenge Joint Test and Evaluation Methodology (JTEM) is working towards during this International Data Farming Workshop (IDFW).

JTEM IDFW 14 OBJECTIVES

Specific objectives of Team 6 sessions during IDFW 14 included:

1. Technical interchange involving the CEM measures framework, data farming, efficient design of experiment, and other selected visualization, modeling, analysis, and simulation (VMAS) capabilities relevant to JTEM.
2. Initial characterization of an Integrated Fires capability area using the CEM for further use in the Data Farming for Test Planning effort.
3. Front-end systems engineering of a CTM test design analysis environment incorporating an Integrated Fires capability use case and candidate VMAS solutions. This included review and refinement of the CTM Develop Test Design process descriptions.

JTEM CAPABILITY EVALUATION

A critical piece of JTEM's efforts entails the development of methods and processes to enable the evaluation of system of systems performance as it pertains to capabilities supporting joint missions. As part of this endeavor, a Metamodel-Test-Metamodel approach is being developed as part of JTEM’s Capability Test Methodology (CTM). In order to structure the underlying business rules and concepts in the CTM’s evaluation thread, a Capability Evaluation Metamodel (CEM) is being developed. The

* For more information contact: Timothy D. Beach, timothy.beach@jte.osd.mil.

CEM is called a metamodel due to its integration across multiple Department of Defense policies related to joint capability and due to its embedded use of design of experiment (DOE) models to describe the CTM test space. The CEM provides the underlying information concepts and relationships to dynamically model and distill the test space into a test design, drive the analysis plan, and systems engineering design of the live virtual constructive distributed environment.

An integrated visualization, modeling, analysis, and simulation (VMAS) environment is required to evolve CEM test design structures. Potential VMAS catalysts include test design visualization, efficient design of experiments, simulation model classes and hybrids, as well as simulation analysis and visualization techniques. Moreover, functional and integration requirements to enable effective and suitable distillation of a capability's test space as part of a capability test design need to be taken into consideration.

THE CAPABILITY EVALUATION METAMODEL

In order to provide conceptual consistency and an underlying business rule structure for the CTM, JTEM is employing an ontology approach. An ontology can be defined as "an explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them." [1]. In keeping with this definition, the ontology supporting the CTM evaluation thread incorporates a JTEM lexicon and capability evaluation metamodel (CEM) to provide underlying conceptual definitions and relationships for the CTM [2]. The JTEM lexicon is a cross-domain dictionary of CTM-relevant Department of Defense (DOD) terminology and definitions. Authoritative DOD sources are used, where possible, for JTEM terms and definitions. When modifications or additional terms are needed for the CTM, these are noted in the lexicon as proposed by JTEM, requiring feedback from JTEM to authoritative DOD lexicon sources.

The CEM provides a conceptual model to relate key CTM test and evaluation lexicon concepts, including capability, system of systems, mission, task, and various types of measures. Key concept hubs of the CEM are represented in Figure 1 as boldly outlined rounded rectangles. A central CEM concept hub is Joint Capability and it is expanded in Figure 1 to show its main relationships. Capability is defined in the DOD Joint Capabilities Integration and Development System (JCIDS) instruction as "the ability to achieve a desired effect under specified standards and conditions through combinations of means and ways to perform a set of tasks" [3]. This definition is reflected in the CEM's Joint Capability hub relationships.

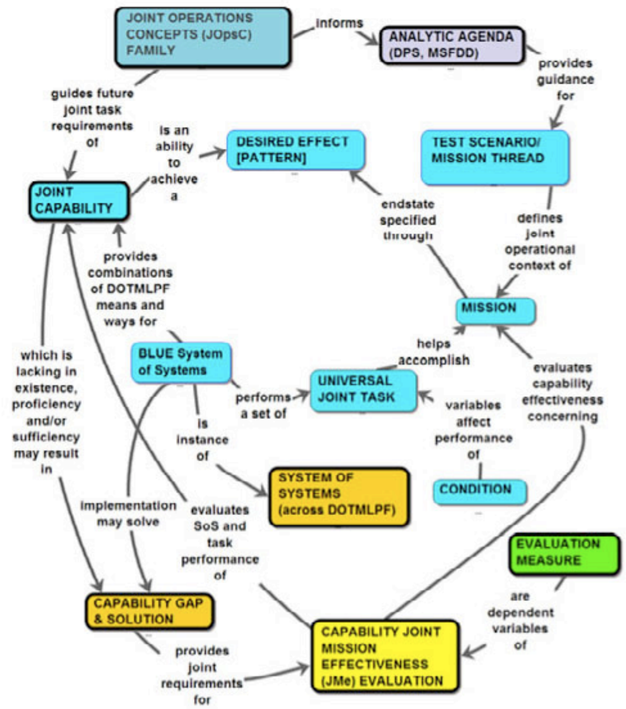

Figure 1. Capability Evaluation Metamodel (CEM) Concept Hubs

A Blue (Friendly) System of Systems (SoS) provides the means and ways for a Joint Capability to perform a set of Universal Joint Tasks. Such Universal Joint Tasks help accomplish Missions, whose Endstate is specified through Desired Effects. The JCIDS capability definition also mentions Conditions, which can be related as variables (e.g., environmental, disparate forces) affecting the performance of Universal Joint Tasks. Although not mentioned in the JCIDS capability definition, the concept of mission is important to relate Universal Joint Tasks to Desired Effects. Joint Capability hub relationships complete with Joint Capability being an ability to achieve Desired Effects. The Blue SoS identified in the Joint Capability hub is an instance of the System of Systems concept hub, which incorporates non-materiel and materiel aspects across the resource construct of doctrine, organization, training, materiel, leadership and education, personnel, and facilities (DOTMLPF).

The joint operational context for evaluating Mission Desired Effects is to be defined by a Test Scenario or Mission Thread and guided by authoritative sources from the DOD Analytic Agenda concept hub involving Defense Planning Scenario (DPS) and more detailed Multi-Service Force Deployment Document (MSFDD) descriptions. DOD also develops the Joint Operations Concepts (JOpsC) Family, including subordinate Joint Operating Concepts (JOC), supporting Joint Functional Concepts (JFC), and detailed Joint Integrating Concepts (JIC) that amplify a portion of a JOC or JFC. JOpsC Family Concepts inform the development of Analytic Agenda products and guide joint task requirements of future Joint Capabilities.

JTEM USE CASE EXERCISE

Based on the Joint Capability hub, a use case was developed to instantiate the concepts presented in the CEM. This use case provided a Joint Operational Context for Test based on an Integrated Fires (IF) joint capability provided by two systems (Network Enabled Weapon [NEW] and a Fire Support Platform [FSP]) within a Systems of Systems (SoS). The mission of the IF capability in the use case is to block the advance of enemy forces into a main Joint Forceable Entry Operations (JFEO) operations area. This IF use case is graphically presented in Figure 2 as a DoDAF OV-1 view that incorporates Mission, Effects and End State Attributes, the SoS structure and the tasks that provide the contribution to achieving the mission end state.

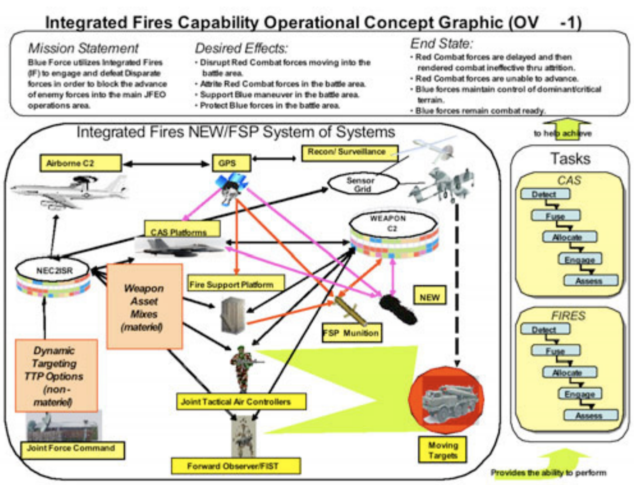

Figure 2. Integrated Fires OV-1 View

Team 6 participants presented and used the IF capability use case to populate CEM structures, which define input factors and levels of a capability test space. These test design dimensions are mission, system of systems, and mission conditions (including disparate force and environmental conditions). This exercise of applying joint mission-level capability concepts to the structure of an efficient design of experiment (DOE) [4] provided a basis for further configurations of a CTM test design.

CTM REFINEMENTS

During working sessions, Team 6 participants also reviewed and discussed refinements to initial modeling sections of the CTM. Based on this discussion, JTEM has proposed refinements to the CTM and added detail at lower process levels. The majority of the changes to the CTM were incorporated into the second step of the CTM process, the Plan Test phase.

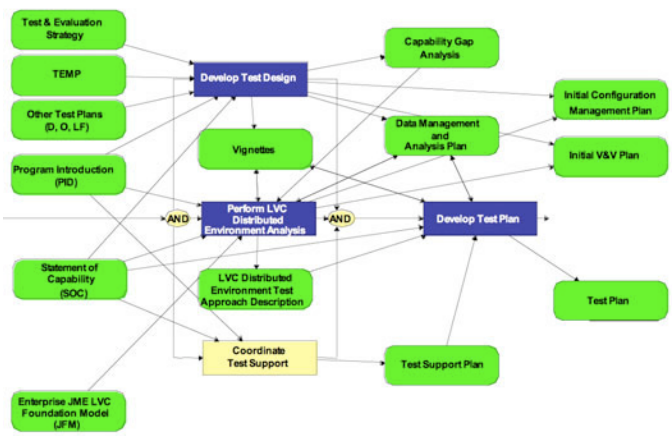

Figure 3. CTM 2: Plan Test Phase

As shown in Figure 3 below, the plan test phase takes test concepts contained in the program introduction document (PID) and statement of capability (SOC) and further develops them into a test plan. The Test Planning process includes developing the test design, performing Live, Virtual and Constructive (LVC) distributed environment analysis, and coordinating test support. Developing the test design involves producing test vignettes and a data management and analysis plan (DMAP). As part of the plan test phase, the parallel process of performing LVC distributed environment analysis produces a LVC distributed environment test approach description. The final part of this parallel process is test support coordination. This step ensures that the test has all the necessary products and range facility support and enables the development of the test support plan. Through completion of the development of the test design, performing the LVC distributed environment analysis, and coordinating the test support plan the test manager is able to develop the test plan. The Develop Test Plan process then synthesizes these processes to develop the overall Test Plan, which incorporates all identified products from this phase.

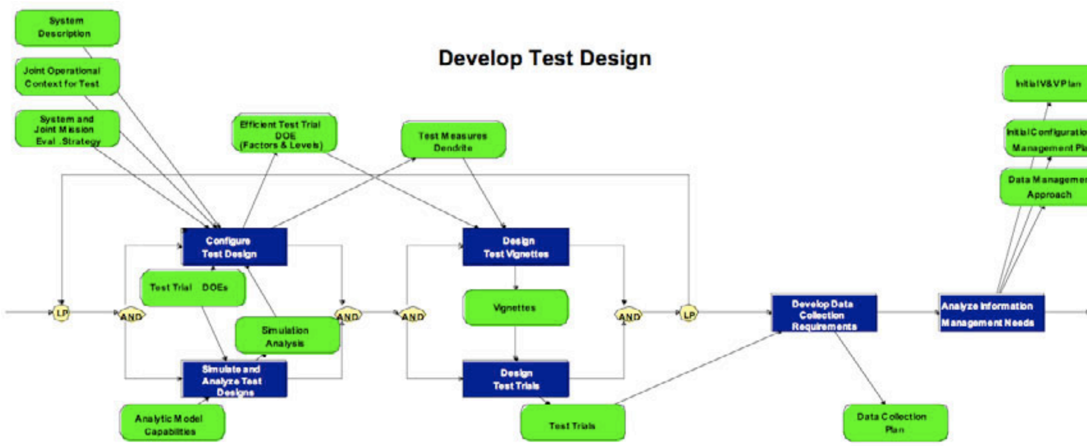

Figure 4. Updated CTM.2.1: Develop Test Design

As a result of IDFW discussions, the Develop Test Design process in the CTM Plan Test phase was revised. This revision further incorporated initial metamodeling and simulation of the Metamodel-Test-Metamodel evaluation approach into the CTM. The revised Develop Test Design process, shown in Figure 4, includes initial processes called Configure Test Design and Simulate and Analyze Test Designs. During Configure Test Design, it is necessary to analyze System Descriptions, the System and Joint Mission Evaluation Strategy (including the test scenario), and the Joint Operational Context for Test. Using these products the program manager can configure an initial test design involving CEM test space dimensions and CEM evaluation measures. The Configure Test Design produces Test Trial Design of Experiment (DOE) output. These Test Trial DOEs are then analyzed in the Simulate and Analyze Test Designs process using Analytic Model Capabilities to better determine suitability of candidate test designs. The Simulate and Analyze Test Designs process produces Simulation Analysis as an output. This simulation analysis is used to help validate test design DOE configurations. If necessary, the test design is then refined, based on the Simulation Analysis and subject matter expertise. These two processes can be iterated until an Efficient Test Trial DOE is produced, including test factors and levels. These processes also produce a Test Measures Dendrite, which contain test response measures.

Once the test design is well developed, Figure 4 shows the processes to Design Test Vignettes and Design Test Trials. From vignette descriptions, test trial matrices are created that specify levels for independent test factor variables and dependent test data collection measure variables. The Test Vignette and Test Trial output provide analysts with the ability to Develop Data Collection Requirements and produce a Data Collection Plan. The final process of developing the test design is the Analyze Information Management Needs process, which produces an Initial Verification and Validation plan, an Initial Configuration Management Plan, and an initial Data Management Approach.

CTM VMAS CATALYSTS

Team 6 participants also discussed an integrated set of VMAS catalysts required to evolve CEM test design structures during the CTM's Plan Test phase. Potential VMAS catalysts include; test design visualization, statistical design of experiments, simulation model classes and hybrids, as well simulation analysis and visualization techniques which can fill capability evaluation gaps in the front-end part of the CTM evaluation thread.

Dr. Kelton asserted that there is an early need to identify critical output measures (metrics) to guide development of models and decision criteria. He discussed automatic specification of empirical input probability distributions, extant simulation models, the need to reduce the dimensionality of the heavy factorial loads that will be associated with a SoS test matrix to foster more efficient, low resolution model designs and the limitations of traditional polynomial-regression-based metamodels. There are several classes of empirical input probability distributions (bounded, infinite right tail) that could aid in rapid modeling and re-modeling as additional data become available. There is the possibility to use extant simulation model as a tool to guide relative efforts on the various inputs. These inputs probably are not all equal and can use simulation as a sensitivity-analysis tool to assess relative importance.

Dr. Kelton discussed the miss-use of random numbers in stochastic simulations. Almost always you can get better simulation results (lower variance with no more computing) if you use random numbers more smartly. This could be largely automated, with the right software design (which doesn't exist yet), so users wouldn't have to think about it or do anything differently.

The basic idea Dr. Kelton would like to see implemented in simulation software using random-number generators (RNGs) is to foolproof their implementation, in three steps:

1. Use a modern underlying RNG with astronomical cycle length (like 10^57, not the measly 10^9 in old RNGs that are still, for some mysterious reason, in wide use) and excellent tested statistical properties. Such RNGs have existed for approximately 8 years.
2. Provide starting seeds for widely separated streams within the RNG. The stream number should be user-definable, but in addition to that the software should automatically increment the random-number stream each time a source of randomness is dropped into the model.
3. Within each stream, provide widely separated substreams that are automatically incremented for each statistical replication for all streams (i.e., substream 1 within all streams for replication 1, substream 2 within all streams for replication 2). With the right underlying RNG and the right seeds, exhausting even a substream (let alone a stream) would take thousands of centuries on today's computers (and, even under Moore's law, it will be a few hundred years before we need to worry about it again).

What this would do (automatically) for users goes a long way toward "synchronization" of random-number use in simulations. In simulation projects involving comparison of several competing policies or alternatives, which is most simulation projects, this results in the reduction of estimates in the difference between policies and alternatives, sometimes dramatically. Therefore, you get more precise results with no extra computing needed (extra computing will always reduce variance). With the large combat simulations, where a single replication can take hours, this could really help run times.

Potential CTM simulation output analysis techniques were also discussed. Dr. Kelton pointed out that when dealing with capability use cases involving complex adaptive systems (CAS), there are limitations in use of traditional polynomial-regression-based metamodels for output analysis. This is due to the need to identify potentially important response discontinuities or "tipping points" in such systems. LTC Schamburg reviewed his Advanced Response Surface Methodology (ARSM) approach to simulation analysis, which has relevance to CTM analysis problems during the Plan Test phase [5] and addresses discontinuities so they are not simply “paved over” by the data plot. These problems involve a larger number of input variables, multiple measures of performance, and complex systems relationships. The ARSM approach capitalizes on the underlying learning philosophy of the traditional RSM while benefiting from other knowledge discovery concepts and data mining techniques. Furthermore it does not require the restrictive assumptions of the traditional RSM nor does it restrict the analyst to the traditional RSM techniques.

Dr. Sanchez presented a candidate comparative analysis technique to address the multiple response analysis problem which occurs during CTM simulation analysis. Once each test measure response is analyzed with respect to test input factors and levels, a matrix can be created to compare results across multiple responses. An example was given for the comparison of two measures, where the following cases could occur: Test factor treatments could just show significance for a single response, or treatments could show significance for both responses. If a factor is significant for both responses, the treatment levels could agree, which would not require further analysis, or disagree, identifying where further tradeoff analysis is needed.

CONCLUSIONS

There are innovative techniques and methodologies that fit within the structural umbrella of the CTM. These techniques are currently being incorporated into CTM test planning processes. Data Farming techniques can provide an important contribution to the definition and evaluation of CTM metrics through Design of Experiment (DOE) techniques and evaluation methods such as ARSM.

The way ahead is to continue development of the efficient DOE relative to the Integrated Fires use case developed as part of the JTEM test sequence. VMAS catalysts and best practices will be identified and incorporated into the JTEM Analyst Guidebook and models designed using efficient DOE and the IF use case will be run during IDFW15.

REFERENCES

[1] Ontology. Dictionary.com. *The Free On-line Dictionary of Computing*. Denis Howe.

http://dictionary.reference.com/browse/ontology (accessed: December 15, 2006).

[2] Dryer, D., Bock, T., Broschi, M., Beach, T. 2007. DoDAF Limitations and Enhancements for the Capability Test Methodology. In *Proceedings of the Spring Simulation Multiconference*, 170-176. Norfolk, VA, 25-29 March.

[3] Department of Defense. 2005. "Joint Capabilities Integration and Development System. Chairman of the Joint Chiefs of Staff Instruction, CJCSI 3170.01E." Washington, DC, 15 May.

[4] Sanchez, S. M. 2005. Work smarter, not harder: Guidelines for designing simulation experiments. In *Proceedings of the 2005 Winter Simulation Conference*, eds. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 69–82. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. Available via http://www.informs-sim.org/wsc05papers/009.pdf [accessed April 24, 2007].

[5] Schamburg, J & Brown, D. 2004. A generalized multiple response surface methodology for complex computer simulation applications. In *Proceedings of the 2004 Winter Simulation Conference*, eds. *R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters*, 958-966. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. Available via http://www.informs-sim.org/wsc04papers/122.pdf [accessed April 24, 2007].