Developing an Integrated Toolset for the Tactical Human Integration Of Networked Knowledge ATO

Peter W. Foltz, Randall Sparks, Mark Rosenstein
Pearson Knowledge Technologies
Boulder, CO
[pfoltz,rsparks,mbr]@pearsonkt.com

Kathleen Carley, Michael Martin
Carnegie Mellon University, CASOS
Pittsburgh, PA
[kathleen.carley,mkmartin]@cmu.edu

Walter Warwick, Michael Matessa
MAAD Operation, Alion Science and Technology
Boulder, CO
[wwarwick,mmatessa]@alionscience.com

Noelle LaVoie, Ursula Lauper
Parallel Consulting
Longmont, CO
[lavoie.lauper]@parallel-consulting.com

***Abstract*— At present, no single tool or analytical technique can analyze and account for the wide range of data and contexts in complex networked environments. This paper describes an ongoing collaborative effort of Pearson's Knowledge Technologies (PKT) division, Carnegie Mellon University's Center for the Computational Analysis of Social and Organizational Systems (CASOS) and the MA&D Operation of Alion Science and Technology. The project builds on each group's separate strengths of modeling command and control systems, understanding teams through the analysis of cognitive, task and team communication, developing software monitors embedded in the network to analyze information flows, automated metrics for team performance, and providing tools for managing and visualizing information. The goal of the effort is aimed at developing an integrated toolset that supports different but complementary analysis methods.**

Keywords- networked teams; cognitive modeling; dynamic network analysis; communication analysis; team performance metrics; task network modeling.

I. INTRODUCTION

It has become commonplace to discuss military networks in terms of their communication, information and social "layers." While it is important to understand network performance, informational displays and re-structured chains of command, the ultimate issue is whether network-centric technologies are aligned with the Warfighter's task, situational contexts and cognitive abilities. Understanding human performance in network-centric environments thus depends on a more holistic view. At present, no single tool or analytical technique spans all three network layers or can analyze and account for the wide range of data and contexts in such environments.

Nevertheless, a number of tools have been developed that assess different, but complementary aspects of networked team knowledge. What is required are ways of combining information and modeling approaches from different tools. This paper describes a collaborative effort of Pearson's Knowledge Technologies (PKT) division, Carnegie Mellon University's Center for the Computational Analysis of Social and Organizational Systems (CASOS) and the MA&D Operation of Alion Science and Technology. The project builds on each group's separate strengths of modeling command and control (C2) systems, understanding teams through the analysis of cognitive, task and team communication, developing software monitors embedded in the network to analyze information flows, automated metrics for team performance, and providing tools for managing and visualizing information.

In particular, the toolset addresses three inter-related aspects of human integration of networked knowledge: CMU's tools for extracting and analyzing dynamic network structure; PKT's tool for understanding and assessing the content of communication over a given network; and Alion's tool for modeling and exploring network performance for given team structures, technologies and human cognitive behaviors. The work focuses, in part, on enhancing the existing tools by supporting interoperability at the data level so that each component tool can leverage the analyses and predictions provided by the other tools. In this paper, we provide an overview of the tools as well as describe progress toward initial interoperability and the types of analyses performed by the combined toolset given a common dataset. We further describe data collection efforts for development and testing of the tools in networked C2 environments.

A. *Project Objectives*

The overall goal of the project is to build on existing research program to perform research and development of an integrated toolset that will support more effective information sharing and automated assessment of teams and multimodal, multiplex networks. The integration of differing but complementary analysis methods can facilitate the development of a set tools that can cover a much wider range of the overall problem space of human integration of networked knowledge. The integrated toolset will extend research and development in network science, team performance metrics, cognitive modeling, task analyses, and team process analysis.

II. ANALYSIS TOOLS FOR NETWORK CENTRIC OPERATIONS

Each team brings a unique but complementary analytical framework for understanding C2 in network centric operations: CMU brings tools for extracting network structure directly from text-based data sources and analyzing dynamic network structures; PKT brings tool for understanding and assessing the content of communication over a given network; Alion brings a framework for exploring the impact of human behavior on network performance for a given team structure. Each group's work spans a gap not addressed by the other two and suggests an approach to integration. For instance, Alion's task network modeling tool provides an environment for representing nominal human behavior and conducting analysis of command and control structures. TeamComm tools developed by PKT can be incorporated to automatically analyze the content of team communication and can provide inputs for message traffic, team performance predictions, and cognitive states for the task network model. Similarly Dynamic Network Analysis (DNA) performed by CASOS tools on entities such as personnel, knowledge, resources, tasks, and locations can be used as input to the task network model and can analyze the output to assess and visualize message traffic data. Below we provide an overview of the primary tools being used, the integration approach and benefits of the combined toolset

A. *Task Network Modelling in C3TRACE/HBA*

C3TRACE is the Army Research Laboratory's Human Research and Engineering Directorate's task network modeling tool. Like other similar tools (e.g., Micro Saint, Imprint), C3TRACE provides a framework for representing human behavior as a decomposition of operator goals or functions into their component tasks, which themselves can be further decomposed. This representation is supported by an intuitive graphical representation of the process being modeled. A discrete event simulator undergirds this graphical representation, allowing the modeler to specify branching logic, resource requirements and other state changes that might occur as a simulation executes.

C3TRACE has been developed to support the evaluation of different command and control structures and information technology on system and human performance. So, in addition to the standard task network graphical user interface, C3TRACE includes tools that allow operator characteristics (e.g., "aptitude," level of training, years in service, etc.) to be specified along with scenario definition tools that allow the modeler to generate a stream of communication traffic to be processed by the operators. A C3TRACE model comprises three components: a task network representation of the functions carried out given a particular command and control structure; a specification of the operators working within that structure, including their individual capabilities and their assignments to particular functions (i.e., tasks); and a scenario defined by the streams of communication traffic to be processed by the operators. At run time, individual, time-stamped messages, tagged according to type (e.g., a *low-priority* message about *enemy location*) flow through the task network and consume resources of the operators assigned to those tasks. Depending on resource constraints at that time and the operator's capabilities, command functions may be completed, or not, and messages then passed to the next function. Measures of performance are drawn from how effectively messages were processed (e.g., whether they were dealt with in a timely fashion), whether command functions could be completed (e.g., was the right kind of information available to complete a task) and whether operators were under- or overloaded with respect to specified workload thresholds.

C3TRACE is actively used by analysts within ARL. It is the most recent addition to a family of task network modeling tools that Alion has developed for both private sector and government use (e.g., IMPRINT). C3TRACE inherits its core functionality from Micro Saint Sharp. (See Swoboda, Kilduff & Kildoff, 2005; Hansberger & Barnette 2005; and Warwick, Archer, Hamilton, Matessa, Santamaria, Chong, Allender & Kelley 2008 for additional details and examples of its application.)

The Human Behavior Architecture (HBA) approach was motivated by the recognition that effective human performance modeling can be achieved at various levels of fidelity with a wide variety of tools. Developed under an Army Phase II SBIR, with continued support under the ARL ADA, C3TRACE/HBA provides a deep level of integration of the C3TRACE task network modeling tool and the ACT-R cognitive modeling tool. The two share graphical user interfaces for building, running and debugging models. Taking advantage of the integration between C3TRACE and ACT-R, we will use ACT-R to represent some of the cognitive limitations of operators as they handle various message streams (digital, voice, radio) in a networked C2 environment. In particular, the attention and memory mechanisms of ACT-R can be used to determine which messages are missed or forgotten. These consequences will feed back into the larger simulation and provide a more detailed analysis of command and control structures and the passage of information through the network.

B. *Team Communication Modeling*

TeamComm and TeamViz are a set of technologies developed at PKT to analyze networked communication data and generate performance metrics. The TeamComm tools use a machine learning approach to analyze semantic, syntactic, relational, and statistical features of the communication streams and automatically associate the features of the communication with aspects of good and poor individual and team performance. The tools are based on Latent Semantic Analysis and other natural language processing and statistical machine learning technologies developed for the assessment of team knowledge and performance.

Both verbal and written communication data are converted into a computational representation which includes a range of measures of the content (what team members are talking about), quality (how well team members seem to know what they are talking about) and fluency and flow (how well team members are talking about it and to whom). Under this approach a model of the relationship between features of the communication and the metrics of performance is built which can then be used to analyze any new communication stream and generate performance predictions. The system is initially trained on data from ratings made by SMEs, or based on objective performance measures, and then automatically learns to judge performance in the same way as the SMEs. Because the system is first trained on human rated data, it can provide predictions over a wide range of performance metrics, as well as provide judgments which raise alarms if performance falls below thresholded levels.

The performance predictions generated by the TeamComm tools can then be incorporated into visualization tools to provide commanders and Soldiers with applications such as automatically augmented AARs and debriefings, near-real-time alerts of critical incidents, timely feedback to commanders of poorly performing teams, and graphic representations of the type and quality of information flowing within a team.

The TeamViz visualization toolset builds on the TeamComm tools to monitor performance, improve collaboration and support planning and decision making in large-scale team exercises. TeamViz automatically analyzes the content and patterns of information flow of the networked communication and provides network visualization tools to improve situation understanding of team members (See Pierce, Sutton, Foltz, LaVoie, Scott-Nash & Lauper, 2006).

Components of the TeamComm/TeamViz technologies have been previously developed and tested on a number of types of command and control communication data. Using human and ASR transcriptions of team missions in diverse applications such as Air Force and Naval command and control, planning Stability and Support Operations in simulated Bosnia, Singapore and Darfur scenarios, a number of performance metrics were successfully predicted.

For example, the technology was used to automatically monitor and predict performance as well as provide performance alarms for convoy teams in the DARWARS Ambush! simulator system and in live STX lane convoy training conducted at NTC (see Foltz, LaVoie, Oberbreckling, Rosenstein, Psotka & Chatham, 2008). Components of this technology accurately predicted both objective team performance scores and SME ratings of performance including measures of Command and Control, Situation Understanding, and Leadership. Prediction reliability was within the range of SME reliability (see Foltz, LaVoie, Oberbreckling, & Rosenstein, 2007). The methods can further detect "critical incidents" during exercise (incidents which significantly changed the course of the operation). In convoy exercises, the system could detect approximately 80-90% of the critical incidents, with false alarm rates of only 15-20%. Additional data collection and analysis is currently underway for FCS-sponsored work at Aberdeen Proving Ground to evaluate team performance metrics in FCS-specific contexts.

The TeamViz tools have been tested for monitoring performance and supporting planning and decision making in realtime in a U.S.-Singapore simulation exercise. The exercise was designed to evaluate collaboration among joint, interagency, and multinational forces conducting combat and stability operations (Pierce, Sutton, Foltz, LaVoie, Scott-Nash & Lauper, 2006). The tools provide teams and evaluators ways of monitoring performance in large collaborative environments using a set of visualization tools built on the communication analysis toolset.

C. Dynamic Network Analysis Tools

The capabilities at the CMU CASOS Center stem from a coherent theoretical framework, Dynamic Network Analysis (DNA), for representing, analyzing, and evolving multi-mode, multi-plex networks of personnel, knowledge, resources, tasks, locations and other entities of interest to C2 applications. The operational core of DNA is a set of interoperable technologies for representing different types of networks in a unified manner (DynetML), extracting networks from disparate sources of text (AutoMap), analyzing and visualizing networks (ORA), and simulating the dynamic, nonlinear effects of actions on multi-agent, multi-level societies at various levels of fidelity. CASOS personnel have used this operational capability to successfully support and enable the application of DNA to real-world problems including counter-narcotics, counter-terrorism, and insurgencies.

The CMU CASOS tool suite supports the awareness and the selection of appropriate courses of action by enabling decision-makers to identify key actors, locations, or resources that can be exploited to gather information, influence others, suppress undesired behavior, assess the immediate and near term impact of removal of nodes and links in the underlying networks, infer changes in performance, information access, ability to mobilize resources, tension, and other forms of behavior. Key forecasting techniques include network healing, belief formation, and attitude propagation, identification of command and control structures, and sub-task analysis. Specific information on the research, development, testing, and implementations of the software tools can be found in Carley and Reminga (2004)

The DNA tools have been researched and developed over a long period of time and have been previously integrated into other DOD efforts. For example, ORA has existed for ten years, and is in use at SOCOM, SKOPE, and many HIDTA's. It is upgraded continuously and new releases are provided approximately every six months. There are annual training sessions, integrated help, sample data, etc. All metrics have been validated and optimized. AutoMap has existed for about 15 years. It is less widely used than ORA by the DOD but it is upgraded continuously and new releases are provided approximately every six months. There are annual training sessions, integrated help, sample data, etc. as with ORA. Construct has been completely rebuilt to use modern agent technology. All sub-models have been validated. It has been used to assess the impact of various interventions on changes in beliefs, social networks, and knowledge networks. DyNetML is the graph-based xml interchange language and version 2 will be released soon.

III. Integration Approach

In order for each modeling tool to leverage off of information from the other modeling tools, a consistent approach to integration is required. The first step to integration, which is currently underway, is to identify the points of interoperability between the modeling approaches. This step involves identifying the key metrics that can be analyzed by each tool and the methods that can be applied by the tool. Then for each tool, we identify the levels of analysis, view of the networked system, key dependencies, and information inputs and outputs required to share information with the other modeling tools.

A. Examples of Information Sharing Among the Tools

Analyzing the content of what is being communicated allows the characterization of who has particular knowledge in the network and where it flows. Figure 1 illustrates an example of how information can be exchanged among the models for combining both hand-tagging and automated tagging of the message content (e.g., characterizations of what is being communicated). An initial C3TRACE model uses hand-coded messages tagged with information about who is communicating about what, where, why, and how. These communication tags can be fed to the PKT and CMU tools which can use information on the message tags to perform additional, automated tagging across much wider sets of communication data as well as refine the hand-coded tags. The tagging from the PKT and CMU tools can feed back to the C3TRACE model in order to permit the development of enhanced C3TRACE models that do not require hand-tagging of the content. This can improve both the fidelity and the speed of development of C3TRACE models. The tagging predictions from the PKT and CMU models further can feed directly into tools for predicting aspects of networked team performance such as situation awareness, trust or the identification of specific team problems. Prior work has shown that automated tagging can provide accurate indications of team performance metrics (See Foltz et al., 2005).

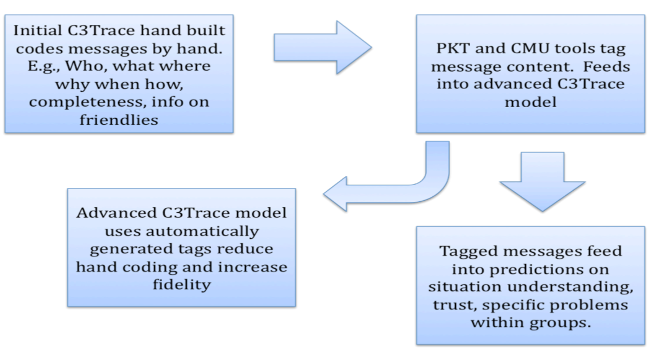

Figure 1. Predictions based on message tags

As a second example, in network centric operations, it is critical that a commander's orders be effectively communicated, understood and operated upon. Thus, methods for monitoring the flow of commander's intent in a network can provide important feedback on the quality of the network and team structures. Figure 2 illustrates how each of the modeling tools can generate complementary analyses of aspects of the commander's intent in the network. PKT's tools generate measures of how commander's intent is being communicated by different operators and groups by time. CMU's tools measure the commander's intent relative to the actual structure of the network. Finally, Alion's tools can compare the measures generated by PKT and CMU against the expected "flow" of intent generated by a C3TRACE model of the doctrinally correct representation of the C2 structure..

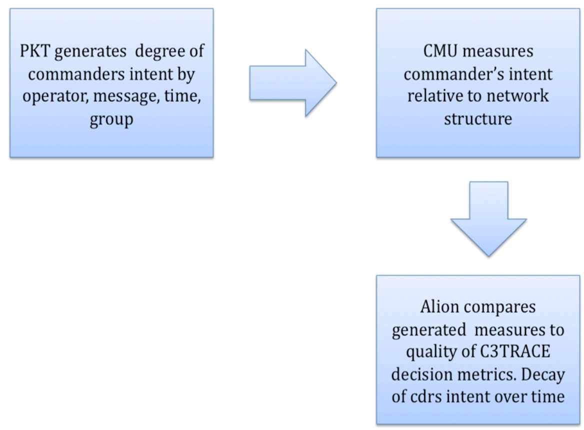

Figure 2. Tracing flow of commander's Intent

B. Integration and Fusion of Model Data

In order to combine the measures generated by the separate models, a second step in integration is the development of a shared dataspace for integration and fusion of the data and model predictions. The shared dataspace development includes using information about the dependencies among the models to create a unified data format both for input data (e.g., communications, task events) and for output predictions (e.g., communication patterns, team SA, critical events). The shared dataspace will contain team state information, mission contexts and communication, as well as model predictions about cognitive, social, and team process metrics. An API will be developed to permit each tool to post to and retrieve data from the dataspace.

Once individual models are able to leverage off of predictions from the other models and data from the shared dataspace, algorithm development will focus on methods for fusing the information from each tool in the dataspace and resolving information conflicts. For example, when two models have differing predictions for the same entity, algorithms will need to be developed to generate a single optimal prediction.

As an example of how the tools can work together using a shared dataspace, spoken communication can be processed through Automated Speech Recognition (ASR) and then analyzed by the TeamComm tools. The TeamComm tools can post the communication to the shared dataspace with tags for each communication indicating predicted measures such as situation awareness, knowledge gap, degree of team consensus, etc. The C3TRACE tool can obtain the tagged communication data from the shared dataspace to populate the communication events in the task network model and align the communication to individual and team tasks, roles, knowledge needs and actions. Similarly, CMU's ORA can analyze network structure key entities such as personnel, knowledge, resources, tasks and locations that will be posted to the shared dataspace and used to help define network topologies in C3TRACE.

Finally, the information in the dataspace will serve as the basis for updating C2 visualization tools. Work will be performed to identify the most critical information provided across the models and determine how that information can be fused and best presented in visualizations and as performance alarms. This work will involve both identifying how prediction data can be represented using the existing C2 visualization tools in the toolset as well as identifying and developing additional visualizations.

For example, based on an operator's characteristics, his tasks and the communication content and amount, C3TRACE can detect whether information has been passed appropriately and whether an operator's workload threshold has been exceeded. These features can then be represented as performance scores and provided as alarms. Continuing with the example, if the integrated toolset then detects information flowing to an operator exceeding the operator's expected workload it can signal an alert or redirect or reprioritize information based on perceived goals. Analyses of network structures provided by ORA will similarly foster detection of significant changes in the network information flow and whether information has been, or should be, redirected.

C. Performance evaluation and data collection

Throughout the project, ongoing evaluation is being conducted to determine the success of the integrated toolset as well as how diagnostic the predictive measures are of key performance indicators. The subcontractor Parallel Consulting will work collaboratively with the team to develop and perform human and system performance evaluation throughout the project.. At the start of the project, a set of quantitative metrics that cover a number of different dimensions of performance success were identified, with the goal to generate and assess these measures throughout the project. Below we outline some of the key metrics and measures that can be used to determine how well the network analysis modeling and tools have achieved results.

Metric: Quality of modeling of team/individual performance of individual tools and tools within integrated toolset

Measures: Correspondence (e.g., correlation and agreement) of model predictions to objective and subjective measures of performance. Objective measures may include quality of team solution, time on task, degree of mission success, ACE reports. Subjective measures include SME, Soldier and team ratings of cognition (e.g., situation awareness, leadership, failures, errors), social variables (trust, team structure knowledge) and processes (e.g., team quality, consensus, correct information flow, information sharing).

Metric: Verification of network forecasting accuracy.

Measures: Identification of command and control structures, analysis of network healing quality, sub-network identification and analysis, and information flow.

Metric: System performance.

Measures: Speed of tool processing of incoming communication and network data. Degree to which data can be processed in real-time. Speed of data sharing and processing between tools.

Metric: Integrated toolset predictions.

Measures: Degree to which information/predictions are provided by more than one method and measures of their individual and joint contributions to overall predictions. Agreement among tool predictions.

Metric: Integration progress**.**

Measures: Number of key components integrated with each other and the interoperability of the integrated tools. Successful development of APIs, shared dataspaces, and unified modeling data structures.

Metric: Usability.

Measures: Usability and usefulness ratings by Soldiers and SMEs for visualization tools and prototype C2 interfaces. Measure of degree of system support needed for use by organizations. Amount of training time required for effective use of visualization tools by Soldiers.

Metric: Client acceptance.

Measures: Integration into ongoing programs (e.g., FCS, C4ISR OTM). Use by THINK leveraged programs. Use/Interest of DOD entities outside of THINK. Degree of research collaboration of project team with DOD program personnel. Spinoff technologies.

Performance evaluation will be conducted on a variety of datasets, including data that was previously collected and new data that is collected specifically to support the integration of the tools. This approach has two main benefits: first, it allows development to proceed without waiting for the time it takes to complete a new data collection effort, and second, it will insure that the integrated toolset generalizes beyond a single exercise.

Initially, the tools are being used to model data collected from a U.S.-Singapore exercise designed to evaluate collaboration among joint, interagency, and multinational forces conducting combat and stability operations (Pierce, Sutton, Foltz, LaVoie, Scott-Nash & Lauper, 2006). 130 Singapore and U.S. military personnel participated in the exercise which ran over 12 days. Email and chat communication were recorded, as well as portions of the commander's briefings. These communications and command structures form the initial data set to be used for integration.

Collection of additional data is anticipated in 2010, and will ideally be drawn from an exercise that has the following characteristics:

- realistic

- full spectrum scenario
- generates extensive communication among participants
- involves complex social networks spanning echelons
- based on well-defined scenarios and tasks
- ability for SMEs to rate performance reliably

Working closely with BCBL, C2D, ARL and ARI we have identified a few potential candidates for future data collections that meet most of these criteria, including OmniFusion, Battle Command Training Program (BCTP) National Guard exercises, and the Digital Warfighter Exercise.

IV. Conclusions

The current work leverages several modeling efforts and works to combine the techniques to provide methods and tools that can account for a much wider picture of performance in network-centric operations. The key goal of the work, to develop the ability of the combined tools to use embedded software monitors to turn networked communication, cognitive states, task and personnel information into performance metrics and alarms. These can be applied both for modeling of performance to test new network-centric operational structures as well as to perform near real-time analysis of team performance. Within the integrated framework, the combined ORA, C3TRACE and TeamComm tools should be able to detect such factors as: critical changes in network structure, changes in human performance (e.g., exceeding workload threshold), changes in team performance (e.g., recognizing breakdowns in particular team performance metrics) and changes in communication or behavioral patterns.

This effort is designed not only to advance the science by providing a more holistic view of networks but also to provide the foundation for real-time improvements for C2 in network-centric environments. In fact, the same software techniques used to model the interactions among communication, information, social and cognitive processes for a given data set can be extended to process incoming data from C2 applications and enable cognitive and task-based mission analysis, automatic generation of performance metrics (collaborative, cognitive, metacognitive, network-based), automated performance alerts, and tools for visualization of network structures and collaboration. The toolset can result in enhanced information sharing as well as timely and relevant feedback to teams and their commanders.

V. References Cited

[1] Carley, Kathleen & Reminga, Jeffrey. (2004). ORA: Organization Risk Analyzer. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Technical Report CMU-ISRI-04-106

[2] Foltz, P. W., LaVoie, N., Oberbreckling, R. & Rosenstein, M. (2007). Tools for automated analysis of networked verbal communication. Network Science Report, Volume 1, 1 pp 19-24, United States Military Academy Network Science Center.

[3] Foltz, P. W., LaVoie, N., Oberbreckling, R. & Rosenstein, M , Psotka, J. & Chatham, R. (2008) DARCAAT: DARPA Competence Assessment and Alarms for Teams. Interservice/Industry Training, Simulation and Education Conference (I/ITSEC 2008), Orlando, FL., December 1-4, 2008.

[4] Foltz, P. W., Martin, M. A., Abdelali, A., Rosenstein, M. B. & Oberbreckling, R. J. (2006). Automated Team Discourse Modeling: Test of Performance and Generalization. In *Proceedings of the 28th Annual Cognitive Science Conference*

[5] Hansberger, J. T. & Barnette, D. (2005). Human performance modeling for operational command, control and communication. Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. pp1182-1185.

[6] Pierce, L., Sutton, J., Foltz, P. W., Lavoie, N., Scott-Nash, S., Lauper, U. (2006). Technologies for Augmented Collaboration. Proceedings of CCRTS, San Diego, CA, July.

[7] Swoboda, J. C., Kildoff, P. W. & Katz, J. P. (2005). A platoon-level model of communication flow and the effects of Soldier performance. Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting. pp1210-1214.

[8] Warwick, W,. Archer, R., Hamilton, A., Matessa , M., Santamaria, A., Chong, R. S., Allender, L.,& Kelley, T. (2008) Integrating Architectures: Dovetailing Task Network and Cognitive Models. Proceedings for the Seventeenth Conference on Behavior Representation and Simulation. Providence, RI SISO.

VI. Acknowledgements

The work described is supported by contract W15P7T-09-C-8324 awarded by CERDEC-C2D under the THINK ATO.