Toward an Integrated Model Checking, Theorem Proving and Simulation Framework for Analyzing Authority and Autonomy

Ellen J. Bass, University of Virginia
Matthew L. Bolton, San José State University
Karen M. Feigh, Georgia Institute of Technology
Elsa L. Gunter, University of Illinois at Urbana-Champaign
John Rushby, SRI International

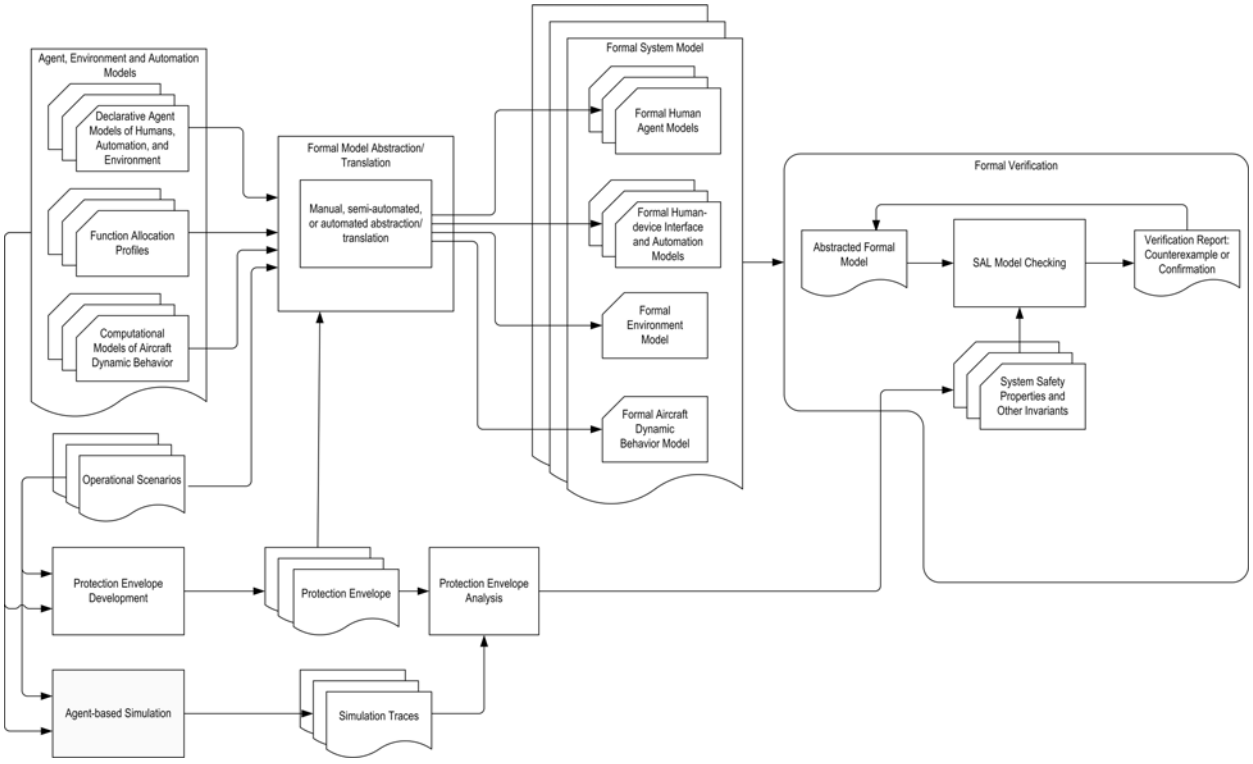
ORIGIN AND UNDERLYING PRINCIPLES

In complex systems, human operators are responsible for a wide array of activities including monitoring the system during normal operations, making minor adjustments when the operational requirements change, diagnosing problems when unusual situations arise, programming any associated automation, and taking over when abnormal situations and emergencies occur. In some domains, roles and responsibilities may shift between human and automation based on environmental situations, regulations, and procedures. New methods must be able to analyze concepts of operation for distributed autonomous and semi-automated systems including their human operators.

No single analysis framework can address the combinatorial explosion resulting from such system complexity. Agent-based simulation has shown promise toward modeling such complexity but requires a tradeoff between fidelity and the number of simulation runs that can be explored. Model checking techniques can verify that the modeled system meets safety properties but they require that the components are of sufficiently limited scope. Thus leveraging these types of analysis methods can help to verify operational concepts addressing the allocation of authority and autonomy.

To make the analyses using these techniques more efficient, we claim that common representations for model components, methods for identifying the appropriate safety properties, and techniques for determining the set of analyses to run are required. In addition, automated tools to create appropriate inputs and to interpret outputs are necessary. Methods to move between levels of abstraction and from one analysis technique to another are also required. Finally methods to ensure that the techniques are addressing their analysis goals are necessary.

Our work begins to address these needs. By developing agent, environment and automation modeling languages, by developing protection envelope-based methods to define and refine system safety properties, by developing an agent-based simulation architecture, by analyzing simulation traces to ensure that the simulation's design meets the intended



analysis goals, by developing abstraction and related methods so that model checking analyses can provide useful information, and by creating associated analysis support tools, our work focuses on verification methodologies and techniques that support human-automation interaction analysis.

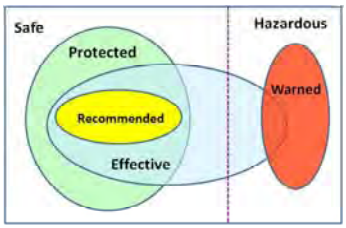
MODELED RELATIONSHIPS

Declarative models with common representation of agents semantically represent the relationships between the model elements to specify the structure of and the interrelations between systems components. A simulation framework, WMC (for Work Models that Compute) models the complex, heterogeneous dynamics of systems that include physical systems, humans, and automated agents [12-13]. Human work is a response to the situation, with strategies chosen based on conditions in the physical environment; the allocation of responsibility within the team; and agent status including expertise, the demands placed on it, and available resources such as time and information. Actions are organized using an abstraction hierarchy [14-15]: at the bottom are the resources and actions, and at higher levels, more aggregate functions provide descriptions that relate the detailed actions to the specific goals of the work.

Enhanced Operator Function Model with Communication (EOFMC) is an XML-based language for describing task analytic models with human-human coordination and human-automation interaction [2,6]. Each human operator model is a set of task models that describe goal-level activities. Activities decompose into lower level activities and eventually atomic human actions. Decomposition operators specify the cardinality of and temporal relationship between the sub-activities or actions. EOFMC models teamwork as shared tasks: coordinated group activities undertaken by two or more human operators while allowing for human to human communication. The EOFMC language has formal semantics specifying how an instantiated model executes [2]. We have developed tools to translate instantiated EOFMs into formal models capable of being evaluated by the model checkers in the Symbolic Analysis Laboratory (SAL) [7] and the theorem prover Isabelle [10].

Human operator knowledge in EOFMC is embedded in task structure (including strategic knowledge specifying when activities can execute). To support model checking analyses with modeled human knowledge of the automation, we have modified the relational abstraction approach of [16] so that we can assert a relation as our model. This is sufficient for our purposes because our relations are conservative (i.e., admit more behaviors than would an accurate model). We construct very approximate models to begin with, then, if we discover an interestingly anomalous scenario (e.g., one in which the pilot's mental mode is "descend" but the airplane is climbing), we refine the model until the scenario becomes realistic or is found to dissolve as an artifact of excessive approximation. Once we have developed a realistic scenario with the model checker, we attempt to reproduce it in a high-fidelity simulation.

To identify and model safety properties, we use the protection envelope [8, 18-19]. Safe sequences are those in which the actions of the operator and system never lead to a domain-dependent concept o f loss. Sequences that are not safe are hazardous. Effective sequences are ones in which a domain-dependent concept of progress is accomplished. Among these are the recommended sequences in which the operator follows the steps in the task description. There may be ways to make progress that are not recommended, perhaps because the recommended procedures describe one of many ways to meet the goal or because other ways may be hazardous. Another set of sequences are the warned sequences. Warning sequences are not always hazardous; they are often aimed at making a sequence non-hazardous by enabling the system designer to make key assumptions about operator behavior. The protected sequences are ones in which the operator may vary from recommended or effective procedures without straying into hazardous territory. We envision this "protection envelope" as an engineered set of properties of the system that form a specified subset of safe behaviors—that is not safe by luck but rather safe by design.



The protection envelopes can be succinctly specified by using logical properties. Currently we use Linear Temporal Logic (LTL) [11] formulae for this purpose. LTL formulae are usually checked against finite state automata. However we need a model that can identify different entities, the actions performed by different entities in different situations and is able to capture the evolution of the system through the combined actions of the entities. The model offering all these characteristics is the Concurrent Game Structures (CGSs) [1], a type of automata that moves from state to state according to the actions of a set of agents. Specifying the protection envelope using LTL allows us to verify inclusion of a behavior in the protection envelope by simply checking whether the CGS corresponding to that behavior satisfies the protection envelope property [18].

With respect to simulation trace analysis, we can formally encode and analyze traces to assess safety and effectiveness requirement conformation [17]. Our work demonstrates that, with the help of faithful abstractions, we can obtain valuable insights about simulated traces from the formal verification procedures irrespective of the size of the simulation traces. The combination of simulation trace generation and formal verification provide feedback that may (i) assess the appropriateness of the requirement specifications, (ii) suggest possible infidelity in the simulation modules and (iii) delineate design error of the original safety-critical system.

PROBLEMS ADDRESSED

While some analyses are exhaustive with respect to possible human action choices, others focus on representative and/or well established patterns of human operator deviations from normative behavior.

APPLICATIONS

We are using Continuous Descent Arrival (CDA) scenarios to drive our work. A CDA procedure allows aircraft to continuously descend from high altitude directly into the ILS glide slope without any level flight segment at low altitude. Conventional approach procedures typically employ periods of constant altitude and speed. While these constant segments simplify the air traffic control tasks of spacing and sequencing traffic by providing periods of well-defined vertical and speed behavior, a CDA aims to eliminate the level altitude segments and their associated thrust transients at low altitude. This keeps the aircraft higher and at lower thrust prior to intercepting the ILS, thereby reducing noise exposure on the ground below.

LIMITATIONS AND DEVELOPMENT OPPORTUNITIES

The work described herein is part of on-going research. Not only do some of the methods require more development in order to become standalone tools, the integration of the methods into a coherent analysis framework requires more attention.

ACKNOWLEDGMENTS

This work was funded in part by NASA award NNA10DE79C and NSF Award 0917218. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NASA and NSF.

REFERENCES

1. Alur, R., Henzinger, T.A. & Kupferman, O. (2002). Alternating time temporal logic. *Journal of the ACM (JACM), 49(5)*, 672–713.
2. Bass, E.J., Bolton, M.L., Feigh, K.M., Griffith, D., Gunter, E., Mansky, W., & Rushby, J. (2011). Toward a multi-method approach to formalizing human-automation interaction and human-human communications. *2011 IEEE International Conference on Systems, Man, and Cybernetics*. October 9-12, 2011, Anchorage, Alaska, 1817-1824.
3. Bass, E.J., Feigh, K.M., Gunter, E. & Rushby, J. (2011). Formal modeling and analysis for interactive hybrid systems. *4th International Workshop on Formal Methods for Interactive Systems (FMIS)*, June 21, 2011, Limerick, Ireland.
4. Bolton, M.L. & Bass, E.J. (accepted). Using model checking to explore checklist-guided pilot behavior. *The International Journal of Aviation Psychology*.
5. Bolton, M.L. & Bass, E.J. (2011). Evaluating human-automation interaction using task analytic behavior models, strategic knowledge-based erroneous human behavior generation, and model checking. *2011 IEEE International Conference on Systems, Man, and Cybernetics*. October 9-12, 2011, Anchorage, Alaska, 1788-1794.
6. Bolton, M.L., Siminiceanu, R.I., & Bass, E.J. (2011). A systematic approach to model checking human-automation interaction using task-analytic models. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 41(5)*, 961-976.
7. De Moura, L., Owre, S., & Shankar, N. (2003). The SAL language manual. CSL Technical Report SRI-CSL-01-02 (Rev. 2). Menlo Park, CA: SRI International. http://sal.csl.sri.com./doc/language-report.pdf
8. Gunter, E. L., Yasmeen, A., Gunter, C. A., & Nguyen, A. (2009). Specifying and analyzing workflows for automated identification and data capture. In *Proceedings of the 42nd Hawaii international conference on system sciences* (pp. 1–11). Los Alamitos: IEEE Computer Society
9. Mansky, D. & Gunter, E. (in press). Using locales to define a rely-guarantee temporal logic. Accepted to *Interactive Theorem Proving (ITP) 2012*.
10. Paulson, L.C. (1986). Natural deduction as higher-order resolution, *Journal of Logic Programming, 3(3)*, 237-258.
11. Pnueli, A. (1977). The temporal logic of programs. *Proceedings of the 18th IEEE Symposium Foundations of Computer Science (FOCS 1977)*, 46-57.
12. Pritchett, A. R., Feigh, K. M., Kim, S. Y., and Kannan, S. (under review). Work models that compute to support the design of multi-agent socio-technical systems. Submitted to *IEEE Transactions on System Man and Cybernetics, Part A: Systems and Humans*.
13. Pritchett, A. R., Kim, S. Y., Kannan, S. & Feigh, K. M. (2011). Simulating situated work. *2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, Miami Beach, FL.
14. Rasmussen, J. (1979). On the structure of knowledge - A morphology of mental models in a man-machine system context. Risø-M-2192, 1979.
15. Rasmussen, J. (1983). On the structure of knowledge - A morphology of mental models in a man-machine system context. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13(3)*, 257-266.
16. Sankaranarayanan, S. & Tiwari, A. (2011). Relational abstractions for continuous and hybrid systems. In G. Gopalakrishnan and S. Qadeer (Eds.): Computer Aided Verification (CAV) Lecture Notes in Computer Science, *6806*, pp. 686–702.
17. Yasmeen, A.,Feigh, K.M., Gelman, G. & Gunter, E.L. (under review). Formal analysis of safety-critical system simulations. Submitted to the *2nd International Conference on Application and Theory of Automation in Command and Control Systems (ATACCS 2012)*.
18. Yasmeen, A. & Gunter, E. (2011). Automated framework for formal operator task analysis. *Proceedings of the 2011 International Symposium on Software Testing and Analysis(ISSTA '11)*. July 17th-21st, 2011, Toronto, ON, Canada, 78-88.
19. Yasmeen, A. & Gunter, E. (2011). Robustness for protection envelopes with respect to human task variation. *2011 IEEE International Conference on Systems, Man, and Cybernetics*. October 9-12, 2011, Anchorage, Alaska, 1809-1816.