Formal Performance Guarantees for Human in the Loop Robot Missions^{*}

D. M. Lyons Senior Member, IEEE, R. C. Arkin Fellow, IEEE, S. Jiang Student Member, IEEE, M. O'Brien, F. Tang, P. Tang

Abstract—A key challenge in the automatic verification of robot mission software, especially critical mission software, is to be able to effectively model the performance of a human operator and factor that into the formal performance guarantees for the mission. We present a novel approach to modelling the skill level of the operator and integrating it into automatic verification using a linear Gaussians model parameterized by experimental calibration. Our approach allows us to model different skill levels directly in terms of the behavior of the lumped, robot plus operator, system.

Using MissionLab and VIPARS (a behavior-based robot mission verification module), we present a comparison of our predicted performance guarantees for two missions in which a teleoperated quadrotor identifies a target for an autonomous ground robot to intercept: one mission in which the operator flies the quadrotor by line of sight to locate the target and one where the operator flies the quadrotor using its video feed. We demonstrate the effectiveness of our approach by comparing predicated performance to experimentally measured performance.

I. INTRODUCTION

Deploying an effective team of robots to search for, identify, and neutralize a hidden chemical, biological, or nuclear weapon of mass destruction will likely involve the use of multiple robots of different capabilities directed in some by human operators. Due to the great potential for loss of life in such a situation, it is important to be able to establish a-priori performance guarantees for the mission. A key challenge here is to model the performance of the human operator and factor this into the formal performance guarantees for the mission.

While research in automatic verification of robot software has typically followed the methods used in general purpose software verification field [1], we have developed an novel approach to efficiently establish probabilistic performance guarantees for behavior-behavior based robot software operating in uncertain environments [2]. A software module based on the approach, *VIPARS*, has been used to establish probabilistic performance guarantees for multiple-robot missions [3], for missions with probabilistic localization software [5]. We have also studied the usability of our system [6].

In this paper, we address the challenge of establishing formal performance guarantees for a multirobot mission in which a quadrotor is piloted by a human operator to search an area for a target, and once it is found, a ground robot is automatically directed to acquire the target. We present a novel approach to modelling the skill level of the operator and integrating it into automatic verification, and we present a comparison of our predicted guarantees with validation results calculated from experimental trials.

The next section presents a literature review. Section III describes the multiple robot mission, the role of the operator in the mission, and the required performance guarantees. Section IV begins with a brief overview of the verification approach we have developed, and then presents our novel approach to operator modelling for formal verification. Section V compares predicted and measured performance using this approach. Section VI discusses these results.

II. LITERATURE REVIEW

The issues involved in how human operator direction or monitoring can be integrated with autonomy/semi-autonomy are a topic of much research. This problem becomes even more difficult if it is necessary to obtain formal performance guarantees for a system that includes a human in the loop (see [7] for review). Much work in that area falls into the category of verifying the human-machine interface (e.g., for medical equipment); however, our work in this paper falls into the category of verifying properties of a system that includes a human operator component. Webster et al. [8] model a human patient that is the uncertain environment in which their Caro-Obot patient care robot works. Their principal concern is to formally guarantee the behavior of the robot with respect to its patient. In our application mission, the operator directly controls a quadrotor searching for the target. Thus, although there is uncertainty associated with the human operator's actions, the operator is an integral part of the robot mission attempting to achieve mission goals, and the effect of the operator on the mission is what must be modeled. Bolton et al. [9] use a Task Analytic Model to model the possible actions of a human driver using cruise control. In our work, our focus is on representing not the possible action of the operator, but rather the skill level and resultant uncertainty with which the operator pilots the quadrotor. As such a key element is collecting performance data on the operator; Driggs-Campbell et al. [10] describe a simulation testbed for collecting such operator performance date for vehicular applications.

III. QUADROTOR/GROUND-ROBOT MISSION

To illustrate the effectiveness of our proposed verification

^{*}This research was supported by the Defense Threat Reduction Agency, Basic Research Award #HDTRA1-11-1-0038.

D.M. Lyons, F. Tang and P. Tang are with the Dept. of Computer & Information Science, Fordham University, NY 10458, USA (Ph: 718-817-4485,

Em: dlyons@fordham.edu). R.C. Arkin, S. Jiang and M. O'Brien are with the Mobile Robotics Laboratory, Georgia Institute of Technology, GA 30332, USA (Em: arkin@cc.gatech.edu)

system in providing performance guarantees for missions involving heterogeneous human-robot systems (Figure 3.1), we developed a simple biohazard search mission where the humanrobot team is tasked to search and locate a biohazard within a given environment.



Figure 3.1: The Human-Robot System

This mission with air and ground robots introduces beneficial heterogeneity into the WMD (e.g., biohazard) search missions we have addressed in our prior work [4]. Quadrotors have excellent speed and mobility, making ideal platforms for searching. Their small payload, however, means they are unable to carry counter-measures. In contrast, ground robots can easily carry counter-measures, but take significantly longer to survey an area. By utilizing these systems together, the search time can be drastically decreased, and the ground robot limited to travelling directly to the target. Our previous work has verified both search based missions and multi-agent missions. The primary research question for this mission is to see if we can extend the verification techniques developed to heterogeneous human-robot systems.



Figure 3.2: FSA for the Ground Robot

The biohazard search mission can be broken down into three steps. First, the quadrotor launches, piloted by a human operator, to search for the target. During this period, the ground robot tracks the quadrotor via a camera. Once an operator locates a target, they will fly the quadrotor directly over it. One of the cameras on board faces downward and can detect the biohazard bucket used as the target in our implementation. Once the operator/quadrotor detects that the target is underneath, a message is sent to the ground robot, signaling that the target is located. In our approach to verification, mission designers first code their mission using the graphical editing tools of *MissionLab* [11]. This is then translated to VIPARS for automated verification.

When the ground robot receives the notification from the quadrotor that a target has been located, the ground robot then uses the current quadrotor's position as an estimate of the target's position. As the quadrotor exits its task by flying back to home base piloted by the human operator, the ground robot will move towards the target location. The *MissionLab* FSA of the ground robot is shown in Figure 3.2. The *GoToTarget* behavior moves the robot toward a target and stops at a distance (R_{max}) from the target; and the input to the behavior is the estimated location of the target.

IV. FORMAL VERIFICATION OF HUMAN IN THE LOOP

A brief introduction to our approach to formal verification is presented first to set the context for our proposed human operator model. In prior work [4] [2] [12], Lyons et al. designed a framework for verifying the performance of autonomous behavior-based robot missions in uncertain environments. *MissionLab* [11] mission software is autotranslated [12] to a *process-algebra* notation *PARS* (Process Algebra for Robot Schemas) for analysis. The interaction of the mission software with an environment model is analyzed to determine if the software will meet a performance criterion.

A. Automatic Verification with VIPARS

A behavior-based program and its environment are modeled in PARS as a set of interconnected processes, where a process P is written as:

$$P\langle u_1, \dots, u_n \rangle (i_1, \dots, i_j) (o_1, \dots, o_k) \langle v_1, \dots, v_m \rangle$$
⁽⁴⁻¹⁾

where $u_1, ..., u_n$ are the *initial* values for the *process variables*, $i_1, ..., i_j$ and $o_1, ..., o_k$ are *input and output port connections*, and $v_1, ..., v_m$ are *final result values* of the process variables. Processes compute result values from initial values, and this computation may be influenced by any communications that occur over port connections. Processes can be defined as combinations of other processes using composition operators: parallel ('i'), disabling ('#') and sequential (';').

Bounded recursion is captured using tail-recursive process definitions, e.g.,

$$\mathbf{P}\langle x \rangle = \mathbf{Q}\langle x \rangle \langle y \rangle; \mathbf{P}\langle y \rangle$$
(4-2)

Process **P** first activates process **Q** with input value *x*. **Q** then delivers output value *y*, which is then used to recur **P**. A variable *flow function f*_P relates the values of variables at the start of each recursive step of **P** to those at the end. The flow-function for atomic processes are *specified a-priori*; those for composite process, those defined as compositions of other processes, e.g., (4-2), are *composed* from the flow functions of the component processes. This can be automated to generate flow functions given a set of processes [2] with complexity linear in the number of processes. Since any execution of eq. (4-2) is modeled by $f_{P}^{n}(x_0)$ for $n \ge 1$ and initial parameter value x_0 , we have a straightforward verification method. Unfortunately, not all processes are defined in this form.

The system to be verified is expressed as the parallel, communicating process composition (e.g., **Sys** below) of a robot controller (e.g., **Ctr** with variable r1), and an environment model, (e.g., **Env** with variable r2), written:

$$\mathbf{Sys}\langle r_1, r_2 \rangle = \mathbf{Ctr}\langle r_1 \rangle(a)(b) \mid \mathbf{Env}\langle r_2 \rangle(b)(a)$$
(4-3)

$$= \operatorname{Sys}'(r_{1}, r_{2}); \operatorname{Sys}'(f_{2}, r_{2}) \rangle$$

$$= (f_{0} + (r_{1}, r_{2}), f_{0} + (r_{1}, r_{2})) \rangle$$

$$(4-4)$$

$$= (f_{0} + (r_{1}, r_{2}), f_{0} + (r_{1}, r_{2})) \rangle$$

$$(4-5)$$

$$f_{Sys}(r_1, r_2) = (f_{Sys, r_1}(r_1, r_2), f_{Sys, r_2}(r_1, r_2))$$
(4-5)

In eq. (4-3), the input of **Ctr** is connected to the output of **Env**, (*a*), and the input of **Env** is connected to the output of **Ctr**, (*b*). If only (4-3) were a sequential composition like (4-2) then we

could extract flow functions for the combined interaction of controller and environment. So, in [2] a static analysis algorithm *Sysgen* was developed to rewrite parallel compositions of the form (4-3) into a sequential composition (4-4) where **Sys'** is referred to as the *system period*. Once *Sysgen* analysis is complete, a *system flow function* can be extracted from **Sys'**.

Random variables e.g., r_1 , r_2 , are represented as multivariate mixtures of Gaussians, and operations on random variables are automatically translated by VIPARS into operations on distributions [13]. Flow functions relate variable values at recursion step t of **Sys'** to those at t+1. In the final phase of VIPARS processing, extracted flow functions are converted to conditional probabilities. These are then the basis of a Dynamic Bayesian Network used to carry out forward propagation of probability distributions, to determine whether the combination of controller and environment will meet a performance specification. We demonstrated that this approach is fast and accurate when validated against physical executions (most recently [4]).

B. Modeling a Human Operator

The combination of a specific human operator, with a specific skill level, guiding the quadrotor while searching for a target object, and the quadrotor motion to the target location, will be modelled as *a single, lumped environment process*. This is part of the environment model in verification because both the quadrotor and the human operator represent an external environment to which control signals are sent (the quadrotor) and from which input is taken (sensor input and the signal that the target has been found). The skill level of the operator can then be reflected in the quality of that result.

The environment model in PARS is

Env = Geometry | LumpedQH | Robot | Laser | Time. (4-6)

Other than the lumped quadrotor/operator model **LumpedQH** and the global time process **Time**, this environment is very like those we have used in prior work [5]. The **Geometry** process includes the probabilistic model of the space in which the mission is carried out, including any map information or knowledge of obstacles. The **Robot** and **Laser** processes are the experimentally calibrated models for the ground robot and laser sensor used in the mission. (The port connections and variable initializations have been omitted here for clarity.)

The lumped model **LumpedQH** encapsulates the behavior of the operator in locating the target, and its output is a signal to the ground robot to proceed to the target. If *H* is the set of such operator/quadrotor lumped models, and if $h \in H$ is one model, corresponding to a specific operator and quadrotor, then the lumped model captures the accuracy with which that operator eventually identifies the target, *p*, and the time it takes that operator to identify the target, *t*, as the set

LUMPED(h) = {
$$\phi = P(p), \ \theta = P(t)$$
 } (4-7)

where P(.) is a probability density. We propose the time t taken by this operator to find the target spatial error in location of target are Normally distributed. While there may be many contextual parameters that influence these, we have chosen to start by modelling the distance to the target d as the principal parameter:

$$\phi = N(\mu_{\phi}, \Sigma_{\phi}) = N(\mu_{\phi}(d), \Sigma_{\phi}(d))$$
(4-8)

$$\theta = N(\mu_{\theta}, \sigma_{\theta}^2) = N(\mu_{\theta}(d), \sigma_{\theta}(d)^2)$$

We will consider $\mu_{\phi}(d)$, $\Sigma_{\phi}(d)$, $(\mu_{\theta}(d), \sigma_{\theta}(d))$ to be linear functions of *d*, yielding a linear Gaussian model, and also consider that Σ_{ϕ} is diagonal with σ_{ϕ_x} and σ_{ϕ_y} .

C. Human in the Loop Quadrotor Mission

Figure 3-2 shows the ground robot component of the mission, respectively. The Quadrotor/Operator mission is completely encapsulated in the lumped model process in eq. (4-6), **LumpedQH** $\langle h \rangle$ ()(dT, dP). The result of the Quadrotor/Operator mission is the time of detection, written to the port dT, and location of the target, written to the port dP. The initial parameter h identifies the operator being modeled.

In prior work we have autotranslated [12] and probabilistically verified and experimentally validated [4] [5] sensory triggered sequences of robot motions through environments with obstacles similar to this mission, so we will focus here only on the aspects of it that relate to the novel lumped Quadrotor/Operator model. The *MissionLab* mission software is mapped to the following VIPARS **Mission** processes:

As Section III explains, the ground robot runs the **TrackQuadrotor** behavior to track the quadrotor visually. The detection of the target by the human operator is signaled by the termination of the **TargetDetected** process and its transmission of the detection time on the port connection tT.



Figure 4-1: The Sys process including both environment processes and the Mission processes, showing selected port connections related to the lumped model.

The quadrotor then communicates the relative position of the target to the ground robot. The ground robot transforms the relative position of the target by the tracked position of the quadrotor relative to the ground robot, to get the relative position of the target to the ground robot. This is transmitted by **TrackQuadrotor** on the port connection *cT*. Of course, there are errors associated with the tracking process φ_{tk} and with the transformation φ_{tr} . These are experimentally determined parameters to the **TrackQuadrotor** process. The processes and connections we have discussed are shown graphically in Figure 4-1.

D. Continuous Time Step

Our prior verification work [2] has represented time as a

discretized time step, but represented other variables as continuous; for example, the robot location or spatial accuracy. For this application, however, time needs to also be modelled as a continuous distribution. In this paper, the value of time will be given by another continuous random variable within the environment model, and consequently within the Dynamic Bayesian network slice for each time. This propagates time as a distribution (since random variables are represented in the DBN framework as Gaussian mixture distributions) from one DBN slice to the next. Both temporal uncertainty and time steps can vary between slices. Furthermore, the mixture model allows multimodal time steps. For example, a slice might generate a mode at t=5 with some variance and a second mode at t=10 with some variance, indicating the slice ends with high probability with one of these times.

V. RESULTS

A. Validation

Figure 5.1 illustrates a trial run of the search mission in a laboratory environment, where the target is represented by a red bucket and there are other non-target objects (e.g., boxes). A box is placed on the path between the ground robot and the target such that it occludes the robot's direct observation of the target. The mission starts with the human operator teleoperating the quadrotor to the target location. Once the target is detected, a signal is sent to the ground robot. The quadrotor is then flown back to a safe location. The ground robot uses the signaled location of the target to move toward; it stops when it is within 1.5m of the target.



Figure 5.1: Biohazard Search Mission within a Laboratory Environment

The ground robot is a *Pioneer 3-AT* equipped with a camera and a front facing SICK laser scanner for obstacle avoidance. The ground robot is controlled by the FSA (Figure 3.3) presented earlier. The quadrotor is an *Ascending Technologies' Hummingbird*. The quadrotor has two onboard cameras. One camera faces forward and is used for flying by video, while the other camera faces downward and is used for target detection. The object is detected when the biohazard is within the center field of view of the downward facing camera onboard the quadrotor.

We examine two modes of teleoperation for the validation experiment: the quadrotor is teleoperated by a human operator either by line of sight (LOS) or through the video stream from the quadrotor (FBV). The experiment is run 50 times for each operation mode. The performance of mission is evaluated with spatial and time criteria:

- 1. \mathbf{R}_{max} success radius; the ground robot is required to be within this radius (e.g., 1.5 *m*) of the target
- T_{max} mission deadline; the mission needs to be completed under this time limit (e.g., 60s)

For each trial, the time it takes for the team to complete the mission (t) and the distance of the ground robot from the target (r) are measured. The mission is only considered successful when both of criteria are met:

Success =
$$(\mathbf{r} \le \mathbf{R}_{\max})$$
 AND $(\mathbf{t} \le \mathbf{T}_{\max})$ (5-1)

B. Calibration

To derive the linear Gaussian models for the lumped operator and quadrotor system, a series of calibration experiments were conducted for the Line of Sight (LOS) and Fly by Video (FBV) cases. In each case an operator was asked to fly the quadrotor from a known start position and find a target. The distance to the target, time to detection, and the spatial error in detection were recorded. A single operator was used for all trials. The linear Gaussian models derived from these calibration measurements were as follows (where θ denotes the time distribution and ϕ the spatial error distribution as in eqns. (4-9)):

Line of Sight Model(LOS)	Fly by Video Model (FBV)
$\mu_{\theta}(d) = 0.4930d + 6.0135 s$	$\mu_{\theta}(d) = 0.4966d + 7.7245 \ s$
$\sigma_{\theta}(d) = 0.0396d + 1.7016 s$	$\sigma_{\theta}(d) = 0.8092 \ s$
$\mu_{\phi x}(d) = 0.0177d + 0.0084 \ m$	$\mu_{\phi x}(d) = -0.0020d - 0.041 \ m$
$\sigma_{\phi x}(d) = 0.2494 \ m$	$\sigma_{\phi x}\left(d\right)=0.1591\ m$
$\mu_{\phi}(d) = 0.0087d - 0.0572 \ m$	$\mu_{\phi y}(d) = 0.0064d + 0.004 \ m$
$\sigma_{\phi y}(d) = 0.0879 \ m$	$\sigma_{\phi y}(d) = 0.1435 \ m$

These models are used within lumped model process to generate the univariate Gaussian time and bivariate spatial error given the target distance (d) in the mission to be verified.

The result of the VIPARS verification of the mission described in the validation section is obtained by inspecting the final values of the time and ground robot position state variables to generate a probability of completion time graph and a spatial error probability graph.

C. Results

Figures 5.2 & 5.3 show the verification and validation results of the probability of mission success for time (T_{max}) and spatial (R_{max}) criteria for the LOS operation mode, where the quadrotor is teleoperated by line of sight. The results are further divided into three regions based on the probability of success generated by VIPARS: 1) High Confidence (Unsuccessful), 2) Uncertain, and 3) High Confidence (Successful). This division provides a succinct representation of the performance guarantee of the mission that is more readily comprehensible to the mission operator. Since the goal of verification is to ensure mission success before deployment, this information helps the mission operator in making such decision. The high confidence regions are where the probability of success is either 0 or 1.0; that is, the mission is either guaranteed to fail or succeed. The uncertain region lies between the high confidence regions, where the probability of mission success is between 0 and 1.0. The mission operator should avoid operating in this region when designing a mission. Furthermore, most of the discrepancies between verification and validation lie in the uncertain region, where the verification error is nonzero.



Figure 5.2: Verification and Validation of the Probability of Mission Success for *Time Criteria* (T_{max}) with Quadrotor Teleoperated by Line of Sight (LOS)



Figure 5.3: Verification and Validation of the Probability of Mission Success for *Spatial Criteria* (R_{max}) with Quadrotor Teleoperated by Line of Sight (LOS)

For instance, Figure 5.2 shows the VIPARS verification and experimental validation results for the performance guarantee of the time criterion $P(t \le T_{max})$, the probability that the cooperative search mission will be completed under the time limit, T_{max} . Based on this result, the mission operator can easily discern that the mission is guaranteed to be successful for $T_{max} > 60$ seconds, or has high confidence that the mission can be completed within 60 seconds. However, if the criterion is too

stringent (e.g., $T_{max} = 30$ seconds), then the mission would be unsuccessful, and the operator could modify mission parameters such as speeding up the robot, using a different robot, or even abandon the mission. Figure 5.2 shows that the VIPARS verification of the mission closely resembles the actual performance of the mission in experimental validation.

Similarly, Figure 5.3 shows the VIPARS and experimental results for the mission performance with the spatial criterion P(r $\leq R_{max}$), the probability that the ground robot reaches within R_{max} radius of the target. We observed that while most of the verification errors lie within the uncertain region, some discrepancies between verification and validation exist near its boundary with the high-confidence (successful) region. Nonetheless, the validation result is of still a relative high confidence with probability of mission success greater than 0.95.







Figure 5.5: V&V of Spatial Criterion at various Rmax

Since the mission is deemed successful only if both criteria of spatial and time are met, the VIPARS verification and experimental validation of both criteria $P(t \le T_{max}, r \le R_{max})$ are summarized in Figure 5.4. The result allows the robot operator to query the verification system for probability of overall mission success for different combination of performance criteria.

To further analyze the results, we examine the effects of the one performance criterion on the other (R_{max} and T_{max}) in terms of performance guarantees and verification errors in Figures 5.5 and 5.6. Figure 5.5 shows how the probability of success for the

time criterion (T_{max}) is affected by various specifications of the spatial criterion (R_{max}). We observed that for R_{max} within its own uncertain region, the various R_{max} 's significantly impacted the performance guarantee of the time criterion $P(t \le T_{max})$. Specifically, it caused $P(t \le T_{max})$ to plateau at probability values other than 1.0 for different R_{max} 's. Moreover, it also introduced significant verification errors. Similar observations are made in Figure 5.6 for the verification and validation of the spatial criterion $P(r \le R_{max})$ at various values of the time criterion, T_{max} .





Figure 5.7: Verification and Validation of the Probability of Mission Success for (a) *Time Criteria* (T_{max}) and (b) *Spatial Criteria* (R_{max}) with Quadrotor Teleoperated in FBV Mode.

Lastly, we also performed verification and validation of the cooperative search mission where the quadrotor is teleoperated with video streaming from a camera onboard the quadrotor (FBV operation mode) instead of line of sight (LOS). The results are summarized in Figure 5.7. For the time criterion (Figure 5.7a), while most of the verification errors lie within the uncertain region, there is some significant error near its boundary with the high-confidence successful region. For the spatial criterion (Figure 5.7b), the verification result is better than the LOS case (Figure 5.3) in the sense that all the verification errors occur within the uncertain region of the performance guarantee.

VI. CONCLUSION

In this paper, we have applied our approach to the probabilistic verification of behavior-based systems to obtain, and experimentally validate, formal performance guarantees for a heterogeneous robot mission in which a human operator pilots a quadrotor to search an area for a target, and once it is found, a ground robot is automatically directed to acquire the target. A novel approach to modeling the skill level of the operator is presented and integrated into VIPARS automatic verification. The method leverages a linear Gaussian model of skill, parameterized by experimental calibration. This approach allows us to model different skill levels directly in terms of the behavior of the lumped, robot/operator system.

We present a comparison of our predicted guarantees for two missions: one where the operator flies the quadrotor by line of sight and one where the operator flies the quadrotor using its video feed. The results in both cases are well aligned and we discuss the specific details of each result. However, lesser verification accuracy was observed for the FBV time criterion as compared to the LOS case (Figure 5.2) for the time criterion. We estimate that this can be attributed to the nonlinearity in the performance of the human-quadrotor system, which we approximated with a linear performance model, eq. (4-8), and with a single dependent variable, distance to target.

VII. REFERENCES

- R. Jhala and R. Majumdar, "Software Model Checking," ACM Computing Surveys, 41(4), 2009.
- [2] D. Lyons, R. Arkin, S. Jiang, T.-L. Liu and P. Nirmal, "Performance Verification for Behavior-based Robot Missions," *IEEE Trans. on Robotics*, 31(3), 2015.
- [3] D. Lyons, R. Arkin, S. Jiang, D. Harrington and T. Liu, "Verifying and Validating Multirobot Missions," in *IEEE/RSJ Int. Conf. on Robots and Systems*, Chicago, 2014.
- [4] D. Lyons, R. Arkin, S. Jiang, D. Harrington, F. Tang and P. Tang, "Probabilistic Verification of Multi-Robot Missions in Uncertain Environments," *IEEE Int. Conf. Tools with AI*, Vietro sul Mare, Italy, 2015.
- [5] D. Lyons, R. Arkin, S. Jiang, D. Harrington, F. Tang, P. Tang, "Formal Performance Guarantees for Behavior-based Localization Missions," *IEEE Int. Conf. Tools with AI*, San Jose CA, 2016.
- [6] M. O'Brien and R. Arkin, "An Analysis of Displays for Probabilistic Robotic Mission Verification Results," *Int. Conf. on App. Human Factors and Ergonomics*, Las Vegas NV, 2016.
- [7] M. Bolton, E. Bass and R. Siminiceanu, "Using Formal Verification to Evaluate Human-Automation Interaction: A Review," *IEEE Trans. SMC: Sys* 43(3), 2013.
- [8] M. Webster, C. Dixon, M. Fischer, M. Salem, J. Saunders, K. Koay and D. a. S.-P. J. Dautenbahn, "Towards Reliable Autonous Assistants through Formal verification: As Case Study," *IEEE Trans. HMC*, 46(2), 2016.
- [9] M. Bolton, R. Siminiceanu and E. Bass, "A systematic approach to model checking human-automation interaction using taskanalytic models," *IEEE Trans. SMC-A* 41(5) 2011.
- [10] K. Driggs-Campbell, V. Shia and R. Bajcsy, "Improved driver modeling for human-in-the-loop vehicular control," in *IEEE Int. Conf. on Rob. & Aut.*, 2015.
- [11] D. MacKenzie, R. Arkin and R. Cameron, "Multiagent Mission Specification and Execution," *Autonomous Robots*, 4(1), 1997.
- [12] M. O'Brien, R. Arkin, D. Harrington, D. Lyons and S. Jiang, "Automatic Verification of Autonomous Robot Missions," Int. Conf. Sim., Mod. & Prog. for Aut. Robots, Bergamo, Italy, 2014.
- [13] D. Lyons, R. Arkin, T.-L. Liu, S. Jiang and P. Nirmal, "Verifying Performance for Autonomous Robot Missions with Uncertainty," *IFAC Int. Veh. Sym.* Gold Coast Australia, 2013.