Y1 Progress Report
Award Number: 80NSSC20K1720
Project Name: RASPBERRY SI: Resource Adaptive Software Purpose-Built for Extraordinary Robotic Research Yields — Science Instruments (USC)
PI: Pooyan Jamshidi
Program Manager: Carolyn Mercer

Project Website: https://nasa-raspberry-si.github.io/raspberry-si/
Team Members: Bradley Schmerl (Co-I), David Garlan (Co-I), Javier Camara (Collaborator), Jianhai Su (Graduate Student), Abir Hossen (Graduate Student), Ellen Czaplinski (Science Consultant), Katherine Dzurilla (Science Consultant)

Y1 Progress:

In Y1, our primary focus has been on **Task 1 (Machine Learning Model Learning), Task 3 (Model Compression),** and **Task 5 (Quantitative Planning)**. We decided to focus on Task1 and Task 5 to develop/extend the main backbone of our autonomy module and infrastructure for evaluating the Autonomy. The end-to-end solution allowed us to integrate our Autonomy module with both virtual (OceanWATERS) and physical testbeds (OWLAT) and enabled us to run the integrated solution (Autonomy + Systems Under Test) in the most interesting OceaanWorld scenario--excavation.

In addition to testing and evaluating our solution in a realistic scenario (excavation), we were able to get some results that enabled us to come up with several research ideas and technical ideas to enhance the testbeds. It also enhances the possibility of **smooth integration with other autonomy modules** from other teams. Since our autonomy solutions involve machine learning (ML) predictions, this task enables us to compress large neural networks so that they can be deployed in resource-constrained hardware and consume less energy in the space lander. At the same time, they preserve their predictive power (accuracy). We are performing a large-scale empirical study involving multiple resource-constrained hardware and different deep neural networks. We consider two types of strategies (distributed neural network deployment and model quantization) to enable high throughput model inference with low-energy demand. Furthermore, we believe that the results will provide insights for deploying advanced neural architectures in space explorations.

We prepared a **demo** that explains our project in the context of the ARROW program, our progress so far, some research ideas, and plans for future directions. In addition, the demo disseminates our work and hopefully inspires future scientists. Here is a link to the demo: https://nasa-raspberry-si.github.io/raspberry-si/blog/2021/demo/

We extended our previous **probabilistic planner** for the RASPBERRY-SI project. This extension entailed analyzing the domain in cooperation with other team members and writing prototype specifications for the PRISM probabilistic model checker, which is used in the backend of the planner. These specifications capture an abstract version of an excavation scenario as a Markov decision process for which PRISM can synthesize a policy that selects, e.g., which locations should be excavated.

We designed the planning scenario and learned how to integrate it into the Rainbow framework in terms of what Monitoring is required, what models are needed, how to map to the planning scenario, etc. The Rainbow framework implements a MAPE-K-based autonomy against the virtual testbed OceanWATERS. The Rainbow framework is the backbone of our solution that integrate all components of our Autonomy into a cohesive and orchestrated framework for enabling **Monitoring** (of the system states), Analysis (of structural and behavioral constraints to trigger re-planning), **Planning** (synthesize a policy for generating actions), and **Execution** (that enacts the synthesized plan to the system under test).

We are implementing a **unified interface between the autonomy and both testbeds (virtual and physical)**. Specifically, the interface includes (i) a message loop that is currently implemented using ROS topics, adding a planner ROS node to communicate with ow_autonomy (the executor of PLEXIL plan), and enabling the support of the PLEXIL Update node in ow_autonomy node to retrieve the execution status of a PLEXIL plan. We have also implemented (ii) a translator between the planner and executor as part of this interface. In particular, our planner uses PRISM, a probabilistic model checking language. The planner builds a PRISM model based on the given list of excavation and dump locations, then exports a policy by querying it against a given property that encodes the multi-objective optimization problem and extracts a high-level plan from the policy. The translator translates the policy into a PLEXIL plan. The interface also includes (iii) a layer that translates testbed specific operations into an Intermediate Representation and allows the autonomy module to work seamlessly with both testbeds. This interface could be potentially useful for other teams involved in the ARROW program.

We implemented a time-based **fault injection** and removal for testing against the current Autonomy. A fault_tester ROS node is implemented to monitor the execution time of the current plan such that a specified arm fault will be injected when the corresponding time condition is triggered. The actual injection is made by utilizing the existing fault injection facility inside the OceanWATERS testbed. Since the OceanWATERS testbed can detect an arm fault and publish the fault status in a ROS topic, we enhanced our current Autonomy to periodically listen to the ROS topic and synthesize a new plan once the existence of a fault is notified via the ROS topic.

We have implemented a **deep neural network for energy prediction** that incorporates factors that change energy consumption, including arm-dependent factors (such as velocity and weight), environment-dependent factors (such as surface properties). We performed training on the ground truth using an ML model to predict the energy. To generate the ground truth, we recorded the data (Motor log, motor current, voltage, motor position (angle), velocity, motor effort, depth information of the surface, and Torque-Power Conversion of the arm) at every time step.

We made several contributions to decide what trench location to select to maximize the science value and minimize operational costs. First, we needed to determine the brine from the icy surface and categorize the brine surface based on the scientific values. We, therefore, **curated a small dataset** for initial evaluation and integration of our machine learning models with our planner in the excavation scenario (we used this scenario for the development and extension of our infrastructure). We then built a solution that performs category labeling, instance spotting, and instance segmentation. Specifically, we developed an **ML model that predicted Brine classes** (e.g., High Brine, Medium Brine, Low Brine). The solution estimates the science values per location and provides these predicted values to the Autonomy to select the sampling locations.

We (the science consultants for our project) made several presentations explaining the **overall scientific value of sending a mission to an ocean world**. We summarized relevant science and engineering articles pertaining to project topics, such as the Mars Phoenix lander arm instruments and operations. Finally, we presented the main points from the Europa Lander Study 2016 Report (Hand et al., 2016) and how it relates to our project of a lander performing autonomous operations on Europa.

We designed an approach for **characterizing the "known unknowns" and the "unknown unknowns"** using our advancements in Causal Modeling and Inference. These situations demonstrate the capability of our Autonomy to deal with unexpected and unanticipated situations. In addition to characterizing these situations, we are developing a method to identify these situations on the fly and reason about their impacts on the mission objectives using the model we developed to characterize unexpected situations. The model-based reasoning allows the Autonomy to devise a reaction to avoid a possible fault or identify the root causes and repair the system if the fault has already occurred. This method can be potentially used for offline testing of autonomous systems.

We prepared a **website** for our RASPBERRY-SI project that contains all presentations and research outcomes. The website serves as a medium for **disseminating** the results of our project: https://nasa-raspberry-si.github.io/raspberry-si/

We **presented** the preliminary results of our project. In particular, we delivered a research presentation at CMU in 2021. We also have done an earlier presentation at Claflin University (an HBCU in SC) to promote computer science and to hire summer interns from **underrepresented and under-served** students in the STEM enterprise.

We have also done a couple of **media outreach**:

- **University of South Carolina:**
https://www.sc.edu/study/colleges_schools/engineering_and_computing/news_events/news/2020/jamshidi_ai_space.php
- **The Post and Courier newsletter:**
https://www.postandcourier.com/columbia/news/usc-researcher-wants-to-train-robots-for-nasa-deep-space-missions/article_93d9bb3c-3afa-11eb-bd4c-7700ac496485.html

Plans for Y2:

Task 2 (Transfer Learning): We plan to expand the curated dataset to involve multiple environments similar to the Europa moon in order to increase the generalizability of our predictive ML models to generalize to different environmental situations.

Task 4 (Online learning): We plan to develop a self-supervised approach based on contrastive learning to increase the accuracy of the deployed model on the lander. By pre-training the models using supervised learning and updating the model onboard using contrastive learning, the models not only become more accurate, but we expect to see improvements in model robustness to generalize better in unseen situations or when the distribution of input data changes (e.g., in extreme weather conditions).

We also plan to **incorporate scenarios beyond excavations**. We also plan to deploy and test the autonomy in the Physical testbed developed by NASA JPL.

Furthermore, we plan to develop an **evaluation infrastructure** to enable automated evaluation of the Autonomy based on the evaluation criteria defined by the NASA ARROW program. We will use our prior experience in a DARPA project for developing such an infrastructure.

Moreover, we also plan to do **comprehensive evaluations** of our solution in diverse test scenarios (different missions, environmental conditions, system states).

In Y2, we plan to write **three scientific papers** to disseminate the results of our project. The first paper will be about the characterizations of unknown unknowns and the second paper will be about using the causal model for testing and repair of autonomous systems. We also plan to write a paper about our end-to-end solutions.