NASEC 2014

**Robotics: New Solutions for a Changing World**

# Exploration of Extraterrestrial Planets Using Automated Intelligent Systems


Radhakrishnan Thiyagarajan - The Cooper Union
Colin Shea-Blymyer - Virginia Polytechnic Institute & State University
Tyler Stern - Emory University
Victor Amaral - Johns Hopkins University
Shelby Coppolino - University of Maryland, Baltimore County
Benedict Lee - The Ohio State University
Sarad Dhungel - Howard University
Shaun Hoyt - Clemson University
Ingrid Rumbaugh - Lafayette College

The United States Naval Academy


*Tuesday, November 18, 2014*

# Contents

# 1 Abstract

The Extraterrestrial Vehicular Explorer, EVE, is an automated intelligent system that can be utilized on extraterrestrial planets for the advancement of scientific exploration. A spaceship is deployed containing a Centralized Control Unit (CCU) and worker robots, which collaborate to collect data. Those worker robots send the data to the CCU which processes it. Reinforcement learning is applied in a feedback loop with humans on earth who communicate back to the CCU on whether this information is of interest. If a cluster of data is weighted higher than another, it means that it is of higher interest and the worker robots search for that pattern on the planet. The main sensor on the robots is a mass spectrometer which can identify atoms and molecules in a sample giving an idea of what molecules and elements are present on the planet. As time progresses, the automated system will learn what is relevant data and send appropriate information from time to time to Earth. Little to no human involvement is needed, thereby making the exploration process faster, more productive, and more efficient.

## 2 Introduction

Critical mission commands sent across the astronomical distances covered in space exploration are far too often hindered by the fact that the speed of light is constant, severely handicapping the ability to communicate data across those large amount of space between Earth and the planet of interest. However, this time difference can be mitigated by only transmitting essential information. By increasing the automation of the process of finding and sending relevant data and providing an effective platform, the challenges introduced by the lack of speed and communication in this endeavor may be reduced.

A solution to this problem is the Extraterrestrial Vehicular Explorer, otherwise known as EVE, which is an integrated platform. EVE consists of the Central Control Unit (CCU) and worker robots. The CCU acts as the brain of the operation and is in periodic communication with humans by sending data and receiving feedback on the relevance of the data. The CCU controls the roaming territory and mission of the worker robots while the worker robots collect the data to feed to the CCU. The human feedback to the CCU allows for the humans to guide the EVE system to search for specific areas of interest, which can be changed. This autonomous exploration can be used for exploration on earth as well as on extraterrestrial terrains.

Historically, robots have been used to search outer space. For example, NASA launched four robots to Mars. Curiosity is one of those robots. According to Ackerman in an article in *IEEE Spectrum* Curiosity used to be a remote controlled vehicle. As of August 2013, the Jet Propulsion Lab (JPL) allowed Curiosity to have more freedom than before by only designating a specific destination and allowing Curiosity to autonomously generate a route [1]. Curiosity is an advanced robot, yet it is limited by its lack of autonomy. With EVE, humans help the CCU determine what to look for, and the CCU controls the areas of exploration, allowing for a greater range of autonomy than previous robots, such as Curiosity, allowed for.

# 3 Platform

In order to achieve the above mentioned goals, a group of robots with a predetermined hierarchy will be deployed on the planet of interest. The robot at the top of the hierarchy is the Central Control Unit. The CCU serves as a base and a command center for the worker robots. Any information that is received, sent, or processed will have been reviewed by the Central Control Unit in order to gain relevant insight into the topic and matters at hand. In order for the CCU to accomplish this task, it needs some type of feedback over an ample amount of time. The feedback cycle will start before the CCU is launched to the planet of interest. It will be trained on Earth for a few months since time that it takes for information to travel astronomical distances is relatively long. After the training on Earth, it will be sent to the planet where it will continue to be trained for a couple of months. After constant training, the CCU will be able to operate autonomously without any human interaction sending the important data back to Earth. Furthermore, the CCU will have a sense of what type of data is important and thus, it will be able to coordinate the robots as well. Suppose a robot sends a help request to the CCU; the CCU can either accept that request and send other worker robots to that area or deny the request and command the robot to continue exploring on its own. This interactivity is something that has not been seen in unmanned missions to extraterrestrial planets.

The next line of command is the worker robots. With the CCU in charge, worker robots will be more efficient. The worker robots can either be sent one by one, searching for different types of data, or they can be sent to the same place of interest all at once. Either way, it is possible to expand the current system, taking into account the changing need for different types of data. With an expandable system, it is possible to send robots that travel through different mediums such as the ground and air. Another advantage of this system is that robots that have different tasks can assist each other. For example, if a ground robot needs to be transferred to another part of the celestial object, it is possible to do so by sending an aerial robot to carry the ground robot to the place designated by the CCU. Thus, it saves valuable time spent by ground travel and speeds up the exploration process. The use of such a system would virtually involve absolutely no humans and can minimize the time difference between Earth and celestial objects in terms of controlling a robot from Earth.

# 4 Algorithms and Software

The physical and logical complexity of the task on hand will require a suite of novel and innovative algorithms to automate the majority of EVEs operation. These algorithms can be generally classified into five categories: identifying landmarks of scientific interest, interaction between the CCU and its workers, movement across the terrain, data gathering and analysis, and scheduling and tasking.

## 4.1 Reinforcement Learning

Scenario: The robots travel around and collect data, drilling through materials and analyzing its chemical properties. It transmits this back to the CCU along with its location. The CCU uses the data collected by all of the worker robots at their various locations to create clusters of elements they found. To reduce complexity of storage, new data of an element already located is stored only if it was found at a distance of radius r away from the previous data point. Otherwise, a frequency counter is kept to show how many times that element appeared in that specified radius, r. After enough data is received by the CCU as to take advantage of the full bandwidth of communication back to scientist, the information is transmitted. Meanwhile, the robots continue collecting and transmitting data. When the scientist receives the information, they can reward the system with a high rank for elements that they want to receive more information about and low rank or zero for those that are less interesting or those they want to completely ignore. The CCU receives these rewards and adjusts a reward function for each robot. Each robot is trying to maximize a reward function which represents collecting the data that the scientists denoted as most interesting. As they navigate and collect data, the learning algorithm decides, through trial-and-error, what is worthy to send to the CCU. To describe this trial-and-error learning, the topic of reinforcement learning needs to be explored by giving an overview along with relevant techniques that can be tested to achieve both precision and accuracy.

Reinforcement learning enables robots to autonomously discover an optimal behavior through trial-and-error interactions with its environment [Kober, Bagnell, Peters]. To do so, the robot tries to maximize a reward function using a mapping called the policy that maps states to actions. A state contains information about the current situation of the robot and actions change the state. For every state, the robot is given a reward. Thus, the robot needs to discover a relation between states, actions, and rewards. In order to do so, a policy that maximizes a cumulative function of rewards called the return is introduced. The policy can either be implemented as a look-up table with $(s_i,a_j)$ corresponding to the $i^{th}$ state and $j^{th}$ action, or as a function. Additionally it can be updated after every time step - stochastic - or it can be deterministic.

Typical reinforcement learning procedures are based on a Markov Decision Process (MDP) modeled with Markov Chains containing all the states, actions, rewards and a transition probability. The transition probability denotes the probability of the next state given the current state and action. MDPs carry with them the Markov Property which states that the state at time t does not depend on the states preceding it.

Some issues in reinforcement learning in robotics that needs to be dealt with are as follows

1. The algorithm must cope with delays from sensing and execution

2. As the amount of data collected by the robot increases, it will take a long time to cover all the computation in the state-action space, due to the phenomenon called the curse of dimensionality.

3. As the robot is exploring, its mechanical parts need to be intact, so itll require safe exploration

4. Exploration/Exploitation tradeoff: the robot needs to decide whether to risk taking an unknown action to try to discover higher rewards or play it safe and continue with actions known to give a decent reward

To continue this argument, consider on-policy methods. This means that the algorithm will collect samples from the environment using the current policy. Dynamic programming iterates between improving the policy and updating the policy until the policy does not change or until it converges. The value function can either be updated at every step (value iteration) until it converges to give optimal policy, or the policy can directly be updated (policy iteration). Value functions try to find a policy that maximizes the expected value of the return given estimates of any initial state. Dynamic programming can either be model-based or model-free. Model-based methods require a model of the environment. Since the environment being explored by the robot is unknown, it is necessary to choose a model-free approach which models the value functions directly. This replaces the policy evaluation step of dynamic programming with model-free methods that do not need an explicit transition function or the return function. Below are some algorithms can experiment with for the purpose of the EVE system.

**Adaptive Heuristic Critic**

This consists of maximizing a heuristic value rather than a reward. This value is provided by the critic. In the case of EVE, the critic is the CCU which provides the values sent to the CCU from the scientists to the worker robots. For every value function the critic provides, the robots will learn a new policy. The policy is learned using Suttons TD(0) algorithm. Whenever a state is visited, its estimated value is updated. This update requires notion of the next step ahead.

**Q-Learning**

In addition to value functions, action value functions can be maximized. Action value functions maximize the expected value of the return given that an action $a$ was taken at state $s$ and then followed policy $p$ afterwards. This uses a similar TD(0) algorithm but calculates a temporal error based on the estimate of the value function at time $t+1$ subtracted from the estimate at time $t$. When convergence occurs, the algorithm chooses the action with the highest value at each state.

### R-Learning

This algorithm uses an average reward criterion. The algorithm requires the robot to take actions to maximize its long run average reward. This is defined as the return function divided by the number of steps taken. In R-learning, a Monte Carlo method is utilized. The policy evaluation step can be replaced with a Monte Carlo estimate in case of an episodic problem with finite MDPs. An episodic problem is one where a task runs for $h$ steps and then is started over perhaps by human intervention. In the case of EVE, as a worker robot is exploring, the scientist may change what the robots are supposed to try to find. In this case, the action value function for all state action pairs $(s,a)$ is computed by averaging all the returns that were received from $(s,a)$ over time.

## 4.2 Swarming

Given that EVEs platform makes use of distributed and centralized automation, communication will need to be carried out both between the CCU and its workers and among the workers themselves. Centralized communication will largely be focused on scheduling, as the CCU will be responsible for analyzing what landmarks are of scientific interest. The CCUs primary focus on analysis and communication with the home base makes it the most qualified component in the system to schedule, command, and organize responses to the worker rovers.

Though the concept for a multi-rover exploration suite has been researched before, (see *An Integrated System* Estlin et al.) EVE differs in that it is not a purely peer based system. By rooting the CCU, it automatically becomes more reliable at energy gathering, data retrieval, and mission-base communication. Furthermore, it provides a known location for rover repair if given the capability to fix its companion rovers. The distributed worker rovers would then be tasked with direct goals to be carried out and the CCU would manage their performance. Workers could, if need be, form an ad-hoc communication extension from the CCU to any other worker if its data is deemed sufficiently worthy; a fact that it would communicate to a worker peer, which would then begin moving into a more favorable position. Using the same metric, as long as the CCU is within a normal range, the workers would collaborate using the same parameters to coalesce towards points of potentially greater scientific merit. By assigning mission critical planning to be performed by the CCU and allowing the workers some freedom to change that planning based on the reports of nearby rovers EVE creates a network of exploring agents that naturally seek out and report on the most interesting scientific phenomenon their environment has to offer.

One of the greatest benefits of this multi-rover architecture is the potential to cover more ground more quickly. Through the use of ground covering algorithms, already in use today for search and rescue missions, these rovers could make many more discoveries of scientific merit than anyone rover could while requiring far less oversight from an Earth-bound mission control. Furthermore, by including the CCUs low maintenance requirement, a possibility for rover repair or piecewise replacement of rovers then EVE could operate for far longer periods of time than contemporary rovers.

## 4.3 Transversability & Navigation

One of the difficult problems that need to be taken into consideration is how the robots will be capable of traversing the extra-terrestrial landscape. With robots on Earth, most of the environments that they will be moving across can be observed by humans and if they encounter any problems, a human in the loop will be able to correct the issues. With robots off planet and away from access of humans, humans cannot evaluate the decisions the makes while moving before it has already completed the motion, so the robots will need the ability to process its path to accurately avoid obstacles that will impede or cause harm to the robot and cannot be recovered from. The Mars rover Opportunity uses triangulation from an image taken through the camera to generate the terrain of the planet and calculates whether it can traverse the path and generates all of the paths it could possibly take at that location. The robot uses these paths and judges which one will result in the path that prevents any damage occurring to the robot and able to continue along on its mission.

Besides this factor, it is evident that the surface of these destinations as well as atmospheres will be very different compared to that of Earth; the robot needs to be able to adapt to the new environment in order to explore the new environment. The robots designed to travel on the ground will need wheels or treads that allows it to travel over rocky surfaces, since that is mostly what the robot will encounter. It also needs the ability to judge what it can travel over and it cant. For the aerial robot, it has to be determined whether the planet has the atmosphere to allow for a quad copter design to be able to move in the environment, or possibly consider other solutions to generate lift for it to be able to travel the planet.

# 5 Administrative Remarks

## 5.1 Testing

As efficient algorithms are a crucial part of any automated system prior to making a working prototype of EVE, testing the algorithm in a simulation is very important. Before it is deployed, EVE would undergo different levels of automated testing. The primary stage of testing will be computer-aided simulation, where 3D model imitating EVE consisting of the CCU and worker robots will be created. Then, a series of algorithms will be implemented in the CCU, which employs optimizing the trial and error learning process in which the CCU will execute commands to work on certain area of interest based on the feedback provided by scientists according to their level of interest.

After successfully testing different algorithms in computer aided simulation environment, a working prototype will be made to test on conditions similar to that of the planet of interest.

## 5.2 Power Management

It was decided that intelligent system described by the above criterion needs a power system that is self-sustainable and powerful enough so that the robot could perform maintenance operations dynamically. One of the essential aspects of any robotics system is its power model; a robot would not be able to operate without its power source. It was also decided that once the Extraterrestrial Vehicular Explore (EVE) is being deploy to space, the system will perform most of its task autonomously and through machine learning. Therefore, a stable, efficient, and renewable power source is needed in order for the system to operate for a long period of time. The operation of the robotics unit will include surveying and data collection. The location of the operation will be in the outer space on extraterrestrial planets. Based on these objectives, it is necessary to design power system compatible with the devices being used on board EVE.

EVE can be classified into the two types of unit being used:

1. Central Control Unit (CCU): The CCU is solar-powered, in which energy is absorbed through solar arrays of panels that are equipped on top of the CCU. Since the CCU is a central unit for the system, it is essential that the CCU is equipped with enough power for processing, data transmission, and charging other robotic vehicles. The CCU will serve as a power station for the robot worker to charge their battery. Any additional solar power will be stored in external battery storage in case of an emergency.

2. Worker robot: The worker robot will have its own small solar panel and battery storage. Only fully charged worker robot will be deployed from the CCU to conduct its exploration task. Whenever the worker robot returns to the CCU, its battery will be recharged again through the CCUs solar panels. The worker robots main power source is based on the battery storage. However, they are also equipped with solar panels which will serve as an alternative power source.

3. EVE Power Analysis

   The main issue that arises with this system is it having the ability to remain mobile, while going for extended periods of time without having to charge its batteries. Since EVE will be hugely based on solar power, it is necessary to use the batteries when solar power is not available. The easiest way to address this problem is to use methods to lengthen the battery life by efficiently delegate the use of power. This can be done by using a pulse style mode of operation. This is where the robotics unit enters an idle state while there is no signal being received, but once a signal is received the robotics unit will quickly switch itself into an operate mode where it is then able to receive and transmit data stream. After the device is no longer receiving a signal it will then return to an idle state as seen in modern transmission devices that often use stand-by mode to save power [2]. The stand-by mode uses discontinuous reception to control the deep sleep and wake up cycles. Discontinuous reception is defined by its DRX cycle, which is the periodicity of the discontinuous reception process. The longer the DRX cycle is, the better power saving quality the battery achieves; but the trade-off, however, is a longer wake up time, which is not something optimal in the network design of EVE. The following equation is how to determine the length of the DRX cycle [3].

$$\textit{DRX Cycle Length} = 2^k \textit{ where } k \textit{ is the DRX cycle length coefficient } [3]$$

This is the main method that will be used in order to conserve power among different transmission unit.

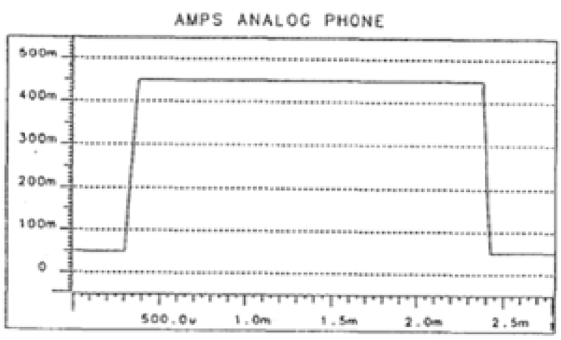


Figure 1: Pulse Mode Operation

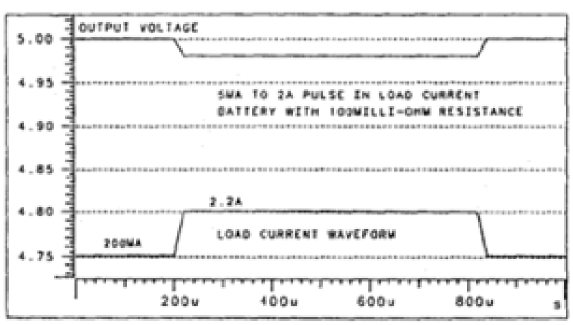


Figure 2: Load on battery of a pulse current

***Figures 1 and 2*** *are representations of a pulse mode style of operation. They came from Kevin Cawleys, Power Supply Transient Response Considerations for Testing Portable Wireless Devices, these pictures really help show how a pulse mode, or stand-by mode can help with saving power [2].*

An issue that arises from using this method for power conservation is that when the device enters into the stand-by mode there are then problems with current leakage through the transistors that are used in circuits of the devices. This leakage current is represented in the following equation, the first component is the active power and the second component is the leakage component [4].

$$P_{total} = CV_{DD}^2F + I_{leak}V_{DD}[4]$$

One way to deal with this is the use of dynamic voltage scaling which uses the lowest $V_{DD}$ to meet the voltage requirements. This helps with two things, the first being obviously the active power, and the second being leakage component, this can be seen in the equation above [4].

## 5.3 Benefits

The main advantage of the EVE system as compared to other space exploration system can be broken down into 2 parts:

1. Cost The proposed EVE system is definitely cheaper in a long run scenario for space exploration due to the systems ability to continuously and automatically collect data without human supervision. Thereby, it drastically reduces the cost of systems management.

2. Faster Discovery & Limited Human Supervision One of the main limitations of space exploration is the low data transmission rate from Earth to Mars and vice-versa. Based on the report from Jet Propulsion Laboratory, the data rate from Mars to Earth varies from 500 bits per second to 32,000 bits per second [5]. Besides data rate, space exploration vehicle can only transmit for a few hours a day due to power limitation. The lack of communication from Earth to rover may delay its mission due to unforeseen conflict.

However, if an autonomous and self-learning system is deployed, artificial intelligence can make logical decisions by itself without the need for human feedback. In addition, there is no need for human supervision since the EVE will perform maintenance operation dynamically.

## 5.4 Funding & Beneficiaries

Although the cost of the EVE project would no doubt be fairly hefty, it should be noted that this is largely contributed by the portion that actually launches the probes into space. While the ultimate goal is to see EVE in action on a distant world, it is evident that gaining funds to realize that dream is difficult.

Considering this financial constraint, it is pragmatic to find funding to develop the technology and systems on Earth thereby refining the EVE designs without the risk and added cost of sending it somewhere in the solar system. Hopefully, NASA, or another space agency with a much large budget, adopts this concept and implements it in a real mission.

After searching through various funding websites, the most likely sources of finances would come from NASAs Discovery Program, which offers a maximum grant of $425 million, albeit with several parameters [8]. For instance, this total budget must include the cost of design, development, the spacecraft, educational outreach, as well as the launch vehicle and mission operations, implying that a mission needs to be planned and carried. Furthermore, the development time of the project until mission launch must be less than 36 months.

If the route of researching the robotic behavior on earth and forgo the trip through space is chosen, the National Robotics Initiative, an organization that provides grants for robotics research pursuits, may be a potential source of funding The NRI is supported by NASA, DARPA, and the NSF, among other esteemed organizations [9]. The awards provided by NRI range anywhere from $20,000 to $1.2 million, which gives a wide range of potential funding for this project. Regardless of whatever path this project ends up taking, the implications of automated planetary exploration is both broad and far-reaching.

The most immediate beneficiaries of EVE are space-venturing companies and organizations such as NASA or SpaceX due to the expansion capability of EVE. However, the most compelling benefit of EVE is the fact that its goal can be manipulated to serve whatever purpose the user sees fit. For instance, In the future, if a mining company wanted to survey and ultimately mine a particular moon for resources, they could utilize a modified version of the EVE to do so.

The technology developed through this research and future space mission may touch many other lives rather than those closest to NASA. As humanity has witnessed in the past, space travel and research has led to the development of technology and science that is commonplace in today's society. This same logic applies to this research as well [10]. With NASAs current funding comprising .05% of the Federal GDP, it is imperative to help push space back to the forefront of the world's mind. It is unknown exactly what sort of future technology could potentially spin off from the EVE project, but it is certain that making the leap into space has greatly impacted lives back on earth in the past, and will continue to do so in the future.

## 5.5 Alternative Applications

As stated before, it is evident that EVE may lead to various other technologies and developments related or unrelated to autonomous machines, but there are several immediate fields that utilize robots and could benefit from the EVE project.

For instance, teams of search and rescue robots could expand their functionality tremendously if the swarming concept was combined with the machine learning algorithms currently being explored by certain individuals in the robotics field. It is imaginable that robots, such as those developed and used by Dr. Robin Murphy, can be deployed into a disaster area, comb different sectors, and identify survivors lost in the rubble of a disastrous event [11].

Another use of EVEs technology could be for deep-sea exploration. Marine biologists, oceanographers, or any type of marine scientist would greatly benefit from the increased accessibility of the ocean that a team of automated machines derived from the EVE could provide.

# 6 Conclusion

EVE will help to automate and simplify space exploration to other celestial bodies, saving time, money and even human lives in the future. EVE will be taught using various machine learning techniques, eliminating the need for constant communication between the CCU and Earth. This means that scientists can focus their attention on relevant data and findings from the celestial body.

EVEs worker robots will be able to traverse multiple types of terrain, such as rocky surfaces, specific to the celestial body being explored, which makes this system very versatile and nimble. These robots will be fit for space exploration than the human body, making EVE very useful and potentially saving human life. While the process of teaching these robots may take some time, it will save time by eliminating the need for scientists to sift through repetitive data.

Estimates show that the cost of this endeavor will not require much more money than is currently being spent on space exploration. The most likely sources of finance would come from NASAs Discovery Program, DARPA, or the NSF. While acquiring funding from large organizations, it is also important to educate the economy about current space projects so that they too, can fund this research.

Finally, the EVE system has far-reaching applications outside of the realm of space exploration. For example, these ideas and concepts can also be applied to areas such as deep sea exploration and other places where conditions do not easily allow human access. The hope is that after some conditioning on Earth before the probe launches, as well as some supervision during the mission, these robots will become completely autonomous, enabling scientific exploration on distant worldś without the presence of humans thus reducing cost, time, and risk.

# 7 Acknowledgements

This report was prepared with the assistance of Radhakrishnan Thiyagarajan, Colin Shea-Blymyer, Tyler Stern, Victor Amaral, Shelby Coppolino, Benedict Lee, Sarad Dhungel, Shaun Hoyt and Ingrid Rumbaugh. Jackie Lasky, Zachary Ha and Colin Heinzmann also helped in coming up with the idea of the EVE system. Further appreciation is given to the United States Naval Academy for the invitation to the Naval Academy Science & Engineering Conference in 2014 through which this report was possible. Deborah Mullen of the United States Naval Academy also helped in directing the conversation during the conference and her help was appreciated. The texts consulted are cited below for further reading.

# 8 References


[1] Ackerman, E. (2013). NASA Lets Curiosity Rover Loose on Mars in Autonomous Driving Mode. *IEEE Spectrum*. Retrieved from http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/nasa-mars-curiosity-rover-autonomous-driving-mode

[2] Cawley, Kevin. *Power Supply Transient Response Considerations For Testing Portable Wireless Devices*. Online. IEEE Explore. 10 Nov. 2014.

[3] Richardson, Anderson. *WCDMA Design Handbook*. (2005, Nov. 11). Retrieved from https://books.google.com/

[4] Clark, Lawrence T., Patel, Rakesh., Beatty, Timothy S. *Managing Standby and Active Mode Leakage Power in Deep Sub-micron Design*. Online. IEEE Explore. 7 March, 2014.

[5] *Data Rates/Returns - Mars Science Laboratory*. Jet Propulsion Laboratory, n.d. Web. 10 Nov. 2014.

[6] Kober, J., J. A. Bagnell, and J. Peters. "Reinforcement Learning in Robotics: A Survey." *The International Journal of Robotics Research* 32.11 (2013): 1238-274. Web.

[7] Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey." *arXiv preprint cs/9605103* (1996).

[8] Research Areas. (n.d.). Retrieved November 16, 2014, from http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503641&org=CISE/

[9] The Planetary Society. (n.d.). Retrieved from http://www.planetary.org/blogs/casey-dreier/20120809-curiosity-comes-cheap.html

[10] National Aeronautics and Space Administration - Marshall Space Flight Center. (n.d.). Retrieved from http://discovery.nasa.gov/p_mission.cfml

[11] Robin Murphy-Department of Computer Science and Engineering. (n.d.). Retrieved from http://faculty.cs.tamu.edu/murphy/