

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) 29-06-2019	2. REPORT TYPE Final Report	3. DATES COVERED (From - To) 1-May-2013 - 30-Aug-2016
---	--------------------------------	--

4. TITLE AND SUBTITLE Final Report: Distributed Learning for Complex Policies	5a. CONTRACT NUMBER W911NF-13-1-0092
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102

6. AUTHORS	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213 -3589	8. PERFORMING ORGANIZATION REPORT NUMBER
--	--

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSOR/MONITOR'S ACRONYM(S) ARO
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) 59786-NS.4

12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.
--

13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:	17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Howard Choset
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU	19b. TELEPHONE NUMBER 412-268-2495

RPPR Final Report

as of 29-Oct-2019

Agency Code:

Proposal Number: 59786NS

Agreement Number: W911NF-13-1-0092

INVESTIGATOR(S):

Name: Howard M Choset
Email: choset@andrew.cmu.edu
Phone Number: 4122682495
Principal: Y

Name: Jeffrey Schneider
Email: schniede@cs.cmu.edu
Phone Number: 4122682339
Principal: N

Organization: **Carnegie Mellon University**

Address: 5000 Forbes Avenue, Pittsburgh, PA 152133589

Country: USA

DUNS Number: 052184116

EIN: 250969449

Report Date: 31-Oct-2016

Date Received: 29-Jun-2019

Final Report for Period Beginning 01-May-2013 and Ending 30-Aug-2016

Title: Distributed Learning for Complex Policies

Begin Performance Period: 01-May-2013

End Performance Period: 30-Aug-2016

Report Term: 0-Other

Submitted By: Howard Choset

Email: choset@andrew.cmu.edu

Phone: (412) 268-2495

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 1

STEM Participants: 0

Major Goals: The modular snake robots in Carnegie Mellon's Biorobotics lab provide an intriguing platform for research: they have already been shown to excel at a variety of locomotive tasks and have incredible potential for navigating complex terrains, but much of that potential remains untapped.

The motivation to expand the capabilities of these robots stems from experiencing several failures and limitations in real world tests. For example, this robot was able to navigate a narrow passage underneath a rubble pile at a disaster response training site, but was unable to pass over a four inch high piece of wood which lay across its path once the passage widened. In an archaeological expedition near the Red Sea, the robot was able to move further than a human could into a collapsed cave containing four-millenia-old ship timbers. However, a gradual sandy slope prevented the robot from moving further and potentially making an archaeological discovery.

To extend the capabilities and learn improved controllers for these robots, a model must be found on which to evaluate the controller. Unfortunately, analytic snake robot models have proven overly complex due to the robots' complex, multi-modal locomotion dynamics, and after pages of formulae these models only provide qualitatively accurate results.

1

(a) (b) (c)

Figure 1.1: This work develops learning algorithms for robotics applications; this includes (a) optimization of motion in steady-state settings, (b) learning from demonstration, and (c) overcoming obstacles with limited feedback.

Numerical simulation models evaluated through computer-based simulation engines provide a simpler approach. However, the accuracy of these models begins to break down as the frequency of collisions increases, and constraint violations due to numerical issues can cause meaningless results.

The only course remaining is to learn through evaluations on the physical robot itself. This is an expensive optimization problem due to the relatively limited experimental budget when working with a physical system as compared to a computer model. In this work, the

RPPR Final Report

as of 29-Oct-2019

CMU team developed machine learning techniques which address the challenges associated with such optimization problems. This work extends research from other fields, simultaneously advancing the robotics literature as well as those fields from which the methods originate.

This work also seeks to add low-level autonomy to improve robot capabilities and simplify the task for the operator, rather than providing complete autonomous control. In the desired applications for the snake robot system which motivate this work (archaeology, search and rescue), a human is crucial for analyzing the scene for objects of interest or potential victims and providing high level directional control. By lowering the cognitive workload necessary

2
to control the robot and understand the scene, I hope to allow the operator to place more focus on the task rather than the tool, improving mission performance. However, elements of this work could still be used for full autonomy in other scenarios.

Additionally, although there is no known simple analytic motion model for the snake robots, this work created a physics simulation of the robot using open source software. Although this model is of questionable fidelity compared to the real robot, we believe that using it as a relatively cheap alternative to evaluation of an action on the robot. We were motivated by how to best incorporate the knowledge generated by the simulator to improve locomotion optimization methods. Although the methods developed herein are primarily designed for improving snake robot teleoperative performance, they are applicable to a wider variety of problems with similar underlying challenges. By addressing problems faced by the current robot with algorithmic and learning advances rather than hardware development, this work addresses the following more general theoretic problems (Fig. 1.1):

1. Optimization of noisy, expensive, black-box functions, with extensions for multiple objectives and learning multiple related tasks.
2. Collection and generalization of demonstrated input for underactuated systems.
3. Optimization of expensive functions with stochastic binary outcomes.

Accomplishments: Improving Gaits for Locomoting Systems

Control of snake robots is difficult in part because of the large number of degrees of freedom in the system. One method of reducing the dimensionality to a more manageable level is through coupling the control of these individual actuators by using gaits. A gait is a cyclic motion in the robot's shape space which produces some (possibly trivial) motion of

3
the robot in the world. Gaits are the foundation of snake robot locomotion; their effective overall motion is easily characterized, and controllability proofs show that cyclic motion is optimal for locomotion over large distances.

Empirically, gaits, including those demonstrated on our robots, have produced impressive results when moving over uniform terrain, including pipes, or terrain with low amplitude disturbances. However, we have no reason to believe that our current set of gaits is optimal or complete; in addition, we know these existing gaits cannot handle all of the challenges that the robot has faced.

The gait optimization work in this work uses a surrogate function to create a model of gait performance as a function of gait parameters. Applying this method enables the robots to move more robustly over a set of small obstacles, and to climb steeper slopes than previously possible. We improved existing surrogate function optimization approaches to enable the optimization to run robustly, without finicky tuning parameters. Also, as stochasticity exists in physical systems being run in the real world, we also developed several approaches to explicitly take into account this noise and show the effects of these methods. As robotic systems often have multiple conflicting notions of optimality, we introduced a multi-objective formulation of the optimization setting above to generate Pareto optimal sets of gait parameters; these sets contain optimal parameter choices for multiple simultaneous (and potentially conflicting) objectives, such as speed and robustness.

We also extended the basic optimization setting to include parameters describing the environment or task, such as slope or crevice width. This extension seeks a control parameter policy over these environment parameters, as opposed to a single optimal gait parameter setting. The result has been demonstrated in a simple autonomous behavior by incorporating state estimation that allows the system to measure and react to its environment. In future work, we would like to more rigorously compare this work to other methods, both similarly

RPPR Final Report

as of 29-Oct-2019

grounded and based on different principles.

4

Learning and Optimizing Non-Gait Shape Changes

Although gaits are the motion of choice for traversing long distances over uniform terrain, more interesting environments will rarely be uniform for significant lengths. Instead, complex non-steady state motions must be learned that enable navigation over complex terrain and large obstacles. Rather than a simple gait optimization problem, this becomes a planning problem using a complex dynamical system in a 16-dimensional trajectory space. Unfortunately, planners require a model that can be evaluated many times in order to generate useful results – a requirement at odds with the expensive nature of physical systems. We used the simulator to provide such a model for the planner; albeit not useful for fine-tuned optimization, it produces reasonable estimation of gross motion.

We also developed an approach to record, simplify, and parameterize demonstrated trajectories from expert and novice users. As the settings where these motions are needed usually measure the result of the motion in terms of success and failure rather than a numerical score, we derived extensions to the optimization framework used for improving gaits to handle stochastic binary functions, and use this to optimize robustness of trajectories for moving over obstacles.

Transitions

Finally, a set of good motions are useless unless it is possible to move between them. Although the space of possible transitions is large, fairly simple methods can be employed to select reasonable transitions. Also, it is critical to be aware of the environment during a transition – even knowing the robot's orientation with respect to gravity has been shown to ensure that the new gait performs as desired. I propose the creation of a metric that will score a potential transition on its quality, given factors such as the environment and stability of the current robot location, using the simulator to help evaluate this metric.

5

Applications to Other Systems

Although a primary goal of this work is to generate tools that allow improved teleoperation of a snake robot, the work would be shortsighted if these could not be applied to any other system. Future work could include applying the work in this proposal to other robotic systems and non-robotic problems. Especially given that many of these techniques have roots in a more general global optimization community, we believe this work will naturally extend to a wide variety of problems.

Training Opportunities: One PhD Student

Results Dissemination: Nothing to Report

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI

Participant: Howard Choset

Person Months Worked: 1.00

Project Contribution:

International Collaboration:

International Travel:

National Academy Member: N

Other Collaborators:

Funding Support:

Participant Type: Graduate Student (research assistant)

RPPR Final Report
as of 29-Oct-2019

Participant: Matt Tesch

Person Months Worked: 6.00

Project Contribution:

International Collaboration:

International Travel:

National Academy Member: N

Other Collaborators:

Funding Support:

DISSERTATIONS:

Publication Type: Thesis or Dissertation

Institution:

Date Received: 03-Aug-2014

Completion Date:

Title: Improving Robot Locomotion Through Learning Methods for Expensive Black-Box Systems

Authors:

Acknowledged Federal Support:

Distributed Learning for Complex Policies

Howie Choset, Carnegie Mellon University

Grant #59786-NS

Chapter 1

Introduction

The modular snake robots in Carnegie Mellon's Biorobotics lab provide an intriguing platform for research: they have already been shown to excel at a variety of locomotive tasks and have incredible potential for navigating complex terrains, but much of that potential remains untapped.

The motivation to expand the capabilities of these robots stems from experiencing several failures and limitations in real world tests. For example, this robot was able to navigate a narrow passage underneath a rubble pile at a disaster response training site, but was unable to pass over a four inch high piece of wood which lay across its path once the passage widened. In an archaeological expedition near the Red Sea, the robot was able to move further than a human could into a collapsed cave containing four-millenia-old ship timbers. However, a gradual sandy slope prevented the robot from moving further and potentially making an archaeological discovery.

To extend the capabilities and learn improved controllers for these robots, a model must be found on which to evaluate the controller. Unfortunately, analytic snake robot models have proven overly complex due to the robots' complex, multi-modal locomotion dynamics, and after pages of formulae these models only provide qualitatively accurate results.

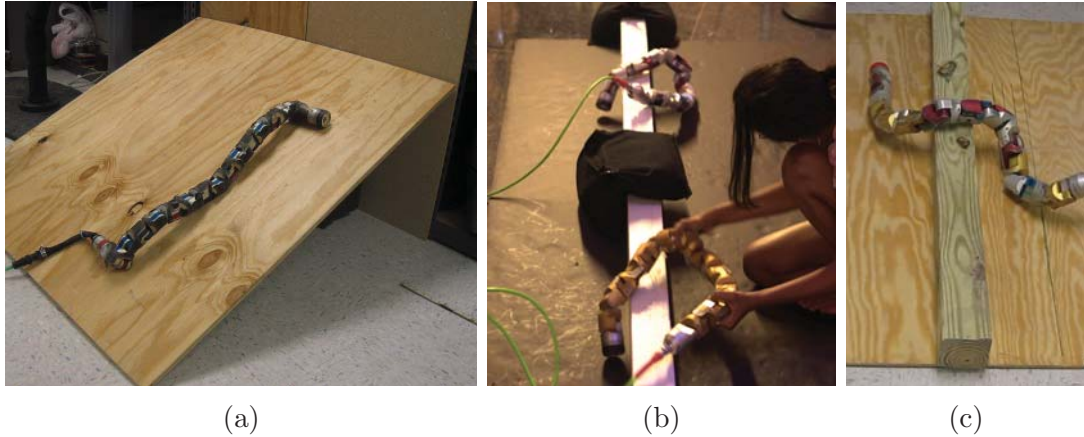


Figure 1.1: This work develops learning algorithms for robotics applications; this includes (a) optimization of motion in steady-state settings, (b) learning from demonstration, and (c) overcoming obstacles with limited feedback.

Numerical simulation models evaluated through computer-based simulation engines provide a simpler approach. However, the accuracy of these models begins to break down as the frequency of collisions increases, and constraint violations due to numerical issues can cause meaningless results.

The only course remaining is to learn through evaluations on the physical robot itself. This is an *expensive* optimization problem due to the relatively limited experimental budget when working with a physical system as compared to a computer model. In this work, the CMU team developed machine learning techniques which address the challenges associated with such optimization problems. This work extends research from other fields, simultaneously advancing the robotics literature as well as those fields from which the methods originate.

This work also seeks to add low-level autonomy to improve robot capabilities and simplify the task for the operator, rather than providing complete autonomous control. In the desired applications for the snake robot system which motivate this work (archaeology, search and rescue), a human is crucial for analyzing the scene for objects of interest or potential victims and providing high level directional control. By lowering the cognitive workload necessary

to control the robot and understand the scene, I hope to allow the operator to place more focus on the task rather than the tool, improving mission performance. However, elements of this work could still be used for full autonomy in other scenarios.

Additionally, although there is no known simple analytic motion model for the snake robots, this work created a physics simulation of the robot using open source software. Although this model is of questionable fidelity compared to the real robot, we believe that using it as a relatively cheap alternative to evaluation of an action on the robot. We were motivated by how to best incorporate the knowledge generated by the simulator to improve locomotion optimization methods. Although the methods developed herein are primarily designed for improving snake robot teleoperative performance, they are applicable to a wider variety of problems with similar underlying challenges. By addressing problems faced by the current robot with algorithmic and learning advances rather than hardware development, this work addresses the following more general theoretic problems (Fig. 1.1):

1. Optimization of *noisy, expensive, black-box* functions, with extensions for *multiple objectives* and learning *multiple related tasks*.
2. Collection and generalization of *demonstrated input* for *underactuated systems*.
3. Optimization of expensive functions with *stochastic binary* outcomes.

The specific contributions of the proposed work include

Improving Gaits for Locomoting Systems

Control of snake robots is difficult in part because of the large number of degrees of freedom in the system. One method of reducing the dimensionality to a more manageable level is through coupling the control of these individual actuators by using *gaits*. A gait is a cyclic motion in the robot's shape space which produces some (possibly trivial) motion of

the robot in the world. Gaits are the foundation of snake robot locomotion; their effective overall motion is easily characterized, and controllability proofs show that cyclic motion is optimal for locomotion over large distances.

Empirically, gaits, including those demonstrated on our robots, have produced impressive results when moving over uniform terrain, including pipes, or terrain with low amplitude disturbances. However, we have no reason to believe that our current set of gaits is optimal or complete; in addition, we know these existing gaits cannot handle all of the challenges that the robot has faced.

The gait optimization work in this work uses a surrogate function to create a model of gait performance as a function of gait parameters. Applying this method enables the robots to move more robustly over a set of small obstacles, and to climb steeper slopes than previously possible. We improved existing surrogate function optimization approaches to enable the optimization to run robustly, without finicky tuning parameters. Also, as stochasticity exists in physical systems being run in the real world, we also developed several approaches to explicitly take into account this noise and show the effects of these methods.

As robotic systems often have multiple conflicting notions of optimality, we introduced a multi-objective formulation of the optimization setting above to generate Pareto optimal sets of gait parameters; these sets contain optimal parameter choices for multiple simultaneous (and potentially conflicting) objectives, such as speed and robustness.

We also extended the basic optimization setting to include parameters describing the environment or task, such as slope or crevice width. This extension seeks a control parameter policy over these environment parameters, as opposed to a single optimal gait parameter setting. The result has been demonstrated in a simple autonomous behavior by incorporating state estimation that allows the system to measure and react to its environment. In future work, we would like to more rigorously compare this work to other methods, both similarly grounded and based on different principles.

Learning and Optimizing Non-Gait Shape Changes

Although gaits are the motion of choice for traversing long distances over uniform terrain, more interesting environments will rarely be uniform for significant lengths. Instead, complex non-steady state motions must be learned that enable navigation over complex terrain and large obstacles. Rather than a simple gait optimization problem, this becomes a planning problem using a complex dynamical system in a 16-dimensional trajectory space. Unfortunately, planners require a model that can be evaluated many times in order to generate useful results – a requirement at odds with the expensive nature of physical systems. We used the simulator to provide such a model for the planner; albeit not useful for fine-tuned optimization, it produces reasonable estimation of gross motion.

We also developed an approach to record, simplify, and parameterize demonstrated trajectories from expert and novice users. As the settings where these motions are needed usually measure the result of the motion in terms of success and failure rather than a numerical score, we derived extensions to the optimization framework used for improving gaits to handle stochastic binary functions, and use this to optimize robustness of trajectories for moving over obstacles.

Transitions

Finally, a set of good motions are useless unless it is possible to move between them. Although the space of possible transitions is large, fairly simple methods can be employed to select reasonable transitions. Also, it is critical to be aware of the environment during a transition – even knowing the robot’s orientation with respect to gravity has been shown to ensure that the new gait performs as desired. I propose the creation of a metric that will score a potential transition on its quality, given factors such as the environment and stability of the current robot location, using the simulator to help evaluate this metric.

Applications to Other Systems

Although a primary goal of this work is to generate tools that allow improved teleoperation of a snake robot, the work would be shortsighted if these could not be applied to any other system. Future work could include applying the work in this proposal to other robotic systems and non-robotic problems. Especially given that many of these techniques have roots in a more general global optimization community, we believe this work will naturally extend to a wide variety of problems.

1.1 Relevance to Army

One of the key features of the developed algorithms is that they apply to a variety of Army-relevant applications occurring over different scales. Here, scale means the size of the area inside of which a task, such as search, must occur. So, the different scales range from large (city-size), to medium (block-size), to small (building size). For each scale, one can envision a different robot having applicability, say a team of blimps searching a city, mobile ground vehicles searching the city block, and a snake robot looking through a building. Many applications have components of all three scales; the proposed work is inspired by the following Army-relevant applications:

Reconnaissance and Surveillance in Urban Engagements: One of the largest concerns of our defense personnel centers on knowing the location of enemies in their immediate surroundings in urban situations. If the enemy does not cooperate, then the military has to blindly maneuver until they blunder onto the enemy at which point the battle begins. Here, the enemy decides when to fight and puts our military on the defensive. Simply locating the enemy and confirming regions are enemy-free makes the military's job significantly easier, safer, and more effective. Imagine if we can quietly deploy a team of blimps, most likely at night, to fly low over a several city blocks. They can collectively search for enemy locations

and serve as sensors in the sky. Mobile robots can drive up and down city blocks deploying sensors for ad-hoc sensor networks. Snake robots can exploit their internal degrees of freedom to peek through holes, maneuver through wall-spaces, and climb through ducts, to furnish a complete view of a room on the other side of a wall. The robot can also place a remote sensor in a room through a hole thereby supplying the war-fighters with telepresence capabilities. With this minimally invasive capability, war-fighters can gain information about the enemy without informing the enemy of our defenses position and intentions.

Search and Rescue: An accident or terrorist attack in an urban center with bio-chemical agents can kill instantly. Assessing the situation, as quickly as possible, is of paramount importance. A bio-chemical technician, wearing a protective suit, can enter the scene, but at extreme risk and at a snail's pace because of the heavy protective suit. Moreover, infrastructure may be unusable and a team of blimps can be deployed to locate regions in a city containing high densities of survivors. Mobile robot may be able to clear city blocks and map out the regions that are safe to traverse. In a collapsed or bombed building, with mobility severely limited and the environment difficult to evaluate, snake robots are well suited to extend the reach of rescue workers to speed the retrieval of trapped victims. With a camera and microphone installed on the robot, rescue workers can more quickly locate survivors and diagnose problems.

Perimeter Security: Wherever the Army has significant assets it would like to protect, it needs to set up a perimeter around those assets that functions as a monitoring system and a first line of defense. Ideally, the perimeter will be a significant distance from the assets in order to afford time for a response to intrusions. This distance means that the length of the perimeter will be significant. Similarly, the ideal perimeter will have some depth that also increases the total area and functionality to be covered. These requirements point to a team of robots with multiple abilities requiring both learning and coordination.