REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
29-10-2019	Final	03/01/2016 to 08/31/2019

4. TITLE AND SUBTITLE
Designing Collaborator Robots for Highly-Dynamic Multi-Human, Multi-Robot Teams

5a. CONTRACT NUMBER
N00014-16-1-2312

5b. GRANT NUMBER
GR00335

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Monica Nicolescu
Mircea Nicolescu
David Feil-Seifer

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Board of Regents, NSHE obo University
of Nevada, Reno
1664 N. Virginia St Mailstop 325
Reno, NV 89557-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 N. Randolph Street Suite 1425
Arlington VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
During this period we designed a unified architecture for human-robot teamwork. This includes: 1) architectural constructs for collaborative human-robot task execution and dialog, 2) active-vision perceptual system for object detection/tracking in cluttered environments, 3) perceptual skills for teammate intent recognition, 4) algorithms for task learning from language instructions, 5) architectural constructs for failure handling through human-robot interaction. We validated these systems in numerous experiments demonstrating collaborative task execution by multiple robots and humans, using two heterogeneous humanoid robot platforms (PR2, Baxter), in the context of a wide range of household tasks.

15. SUBJECT TERMS
Distributed control architecture, active-vision perception, task learning from natural language

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	UU	12	Thomas McKenna
					19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Grant or Contract Number: N00014-16-1-2312
Date Prepared: October 29, 2019
Project Title: Designing Collaborator Robots for Highly-Dynamic Multi-Human, Multi-Robot Teams
Annual Summary Report: CY2018
Principle Investigator:

Monica Nicolescu, (775) 784-1687, monica@unr.edu, University of Nevada, Reno
Mircea Nicolescu, (775) 784-4356, mircea@cse.unr.edu, University of Nevada, Reno
David Feil-Seifer, (775) 784-6469, dave@cse.unr.edu, University of Nevada, Reno

Section I: Project Summary

1. Overview of Project

Abstract:
In this project we propose to address the problem of designing collaborator robots that can operate effectively as members of multi-human, multi-robot teams. The solutions designed for this purpose would be applicable to multiple application domains: military (groups of human soldiers and robots), manufacturing (mixed human-robot assembly environments) or service robotics (humans and robots co-existing in physical worlds). The heterogeneous, multi-agent team poses different challenges from human-robot collaboration in single human-robot domains. In particular, successful interaction and collaboration among the agents of a group that work in close proximity requires several key capabilities: i) awareness, ii) flexibility in control, iii) sociability, and iv) adaptability.

Objective:
The research goal of this project is to design a unified architecture for robot teamwork, which incorporates novel perceptual, control, social behavior and learning systems that will provide robots with awareness, flexibility, sociability and adaptability during long-term interactions in multi-human, multi-robot teams.
In particular we will work toward the following objectives:

- Develop perceptual skills for teammate intent recognition.
- Develop robot control mechanisms for: 1) fast switching between different tasks and roles during task execution and 2) decision making for task allocation.
- Develop social capabilities that allow: 1) detecting initiation of interaction, 2) converting human instructions into executable robot controllers (human-to-robot) and 3) conveying critical information (task details, internal state, requests for assistance) from robots to humans
- Develop a learning algorithm that enables a robot to incorporate knowledge from past experiences with other teammates.
- Integrate the above capabilities in a unified robot control architecture.

Introduction: Please see Section 2 for project details.

Background:
Proposed solution. In our approach, we view that the main objective of a robot should not simply be to achieve its own goals, but rather to be an effective member of its team. Toward this end, the basic issues that this project will address focus on fundamental research on how a robot's perception, control, social behavior and learning systems need to be designed and interconnected

for multi-human multi-robot collaborative domains. Consequently, we will focus on developing and integrating in a unified system: i) a perceptual system that provides the robot with the awareness needed to engage and sustain dynamic, long-term interactions, ii) architectural control structures that enable fast task/role switching and task allocation, iii) natural language communication skills for sociable interactions, and iv) learning tools for adaptable robot teammates. We will validate our approach using two humanoid robot platforms (the Baxter research robot and the PR2), purchased through ONR DURIP Award N000141410776, through user studies that emphasize collaborative task achievement in multi-agent, heterogeneous teams.

2. Activities and Accomplishments

Concise activities:

During this period (06/14/18 to 06/15/19) we performed the following major activities:

1. Developed architectural constructs that enable collaborative human-robot task execution and dialog
2. Developed an active vision-based approach for object detection and manipulation in cluttered environments
3. Began the development of a pose detection system for dynamic object manipulation
4. Extended our approach for language-based task learning & joint human-robot task execution

Expanded activities:

Activity 1. Architectural constructs that enable collaborative human-robot task execution and dialog.

This work brings the following main contributions:
1. Enables joint human-robot collaborative task execution
2. Provides robots the ability to handle faulty/unexpected situations by requesting human help,
3. Enables detection and handling of task execution conflicts between humans and robots,
4. Provides the architectural mechanisms for task interruption/resumption.

In this work we extended our architecture to incorporate a dialogue-based management system in order to handle human-robot coordination as well as unexpected task faults.

For *fault* handling, we designed a system that automatically detects multiple types of manipulation failures: dropping an object, grasping an object out of reach, missed grasps or object positioning issues. The process is shown in Figure 1: the robot continuously checks for possible failures. When failure is detected, the robot stops the execution of its current behavior and initiates a dialog with a human helper to ask for assistance. The robot will continue the dialog until it is asked to retry its task or until the issue has successfully resolved with human assistance.

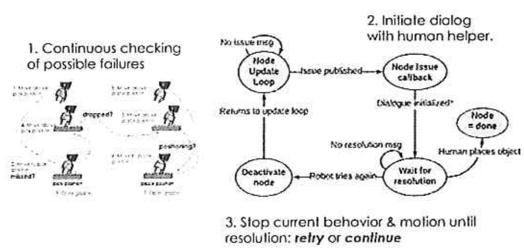


Figure 1. Fault handling through dialog.

To handle *human robot cooperation*, we extended our architecture to handle cases when both a human user and the robot decide to work on the same part of the task, which can occur when the human attempts a task shortly after the robot has already started. When such situations occur, the robot will initiate a dialogue in order to negotiate the conflict. The robot offers two options to its human helper: either to defer to the human or to continue and allow the human to work on another part of the task. Based on the result of the dialog, the architecture enables the robot to continue with the appropriate path.

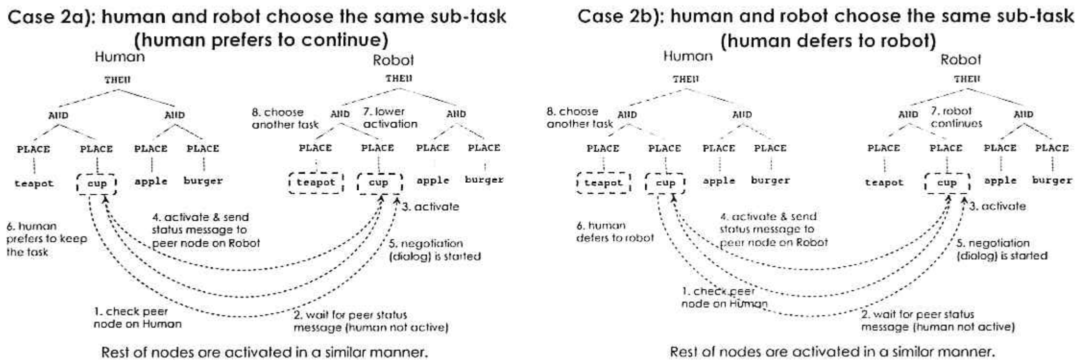


Figure 2. Handling coordination issues during human-robot collaboration.

Activity 2. Active vision-based approach for object detection and manipulation in cluttered environments

This work brings the following main contributions:

1. Uses multiple viewpoints (cameras) and reasons about missing data/occlusions for better object recognition
2. Dynamically plans for arm camera placement to improve object recognition in the presence of occlusions.

In this work, we combined our prior work on integrating multiple camera views with a real-time planner that automatically selects a new arm camera viewpoint, in order to handle occlusions and improve object recognition. The process is presented in Figure 3. The known location of the object of interest is used to define a camera angle that can keep the object in sight. Based on the known arm lengths, the planner computes an elbow and shoulder angle for the robot's arm.

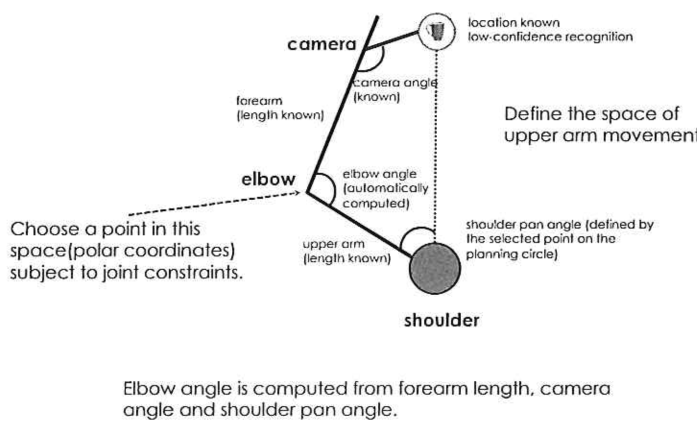
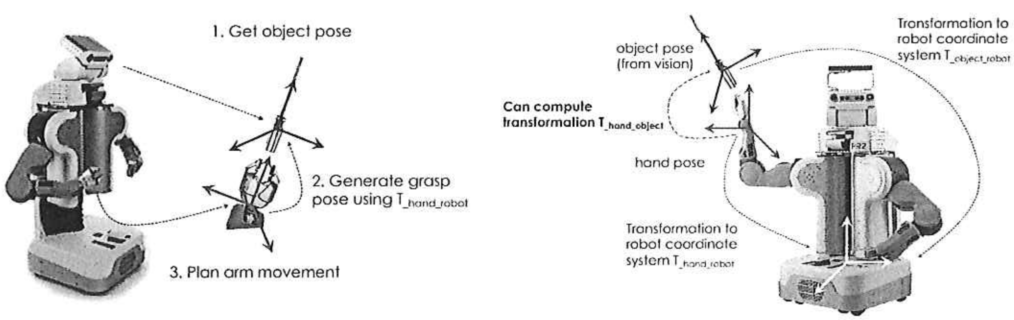


Figure 3. Automatic selection of secondary viewpoint.

Activity 3. Pose detection system for dynamic object manipulation

The contribution of this work is that it provides real-time pose detection of objects (household items, tools) to be used for dynamic object manipulation. Our motivation is that manipulation/tool use tasks require specific, precise grasping, which in turn requires precise knowledge of an object's pose, which changing in dynamic environments. Our approach (Figure 4) goes through a two-stage process. First, we define a set of pre-defined grasps which represent ways in which tools/objects need to be used for specialized purposes. In this training process we generate a transformation between the object and the robot's end-effector that encodes how the object should be grasped. Second, we detect the pose of the object using color and SIFT features, then use the learned grasp transformation to move the robot's hand in the location consistent with that grasp.



a) Define a set of task-specific grasps.

b) Detect pose in real-time, use detected pose to plan a grasp of object

Figure 4. Automatic grasping for specialized manipulation.

Activity 4. Extended approach for language-based task learning & joint human-robot task execution

The contribution of this work is that it provides ability to learn specialized instructions through the use of prepositions (to indicate specific locations) and adjectives (to indicate specific objects).

To indicate in more detail the full capabilities of the approach for parsing language instructions to controllers we experimented with additional, highly complex sentences. These tasks have not been validated on a robotic system, but show the representational power of the approach. The examples test two main capabilities: 1) conveying complex execution constraints in a single instruction and 2) use of multiple adjectives (descriptive, positive, comparative, superlative) and location prepositions to specify task-relevant information.

3. Findings and Conclusions

Findings 1. Architectural constructs that enable collaborative human-robot task execution and dialog.

The *fault handling* approach has been validated with a PR2 humanoid robot given the task to assemble an Ikea Eket base. Figure 5 shows the various stages of the experiment. In (d) the human steals the pink bar before it is placed, resulting in a *dropped* issue. The human then follows the dialogue to place it. In (g) The human steals the yellow bar after it is picked, resulting in a *dropped* issue. The human then follows the dialogue to allow the robot to pick and place it again. In (j) the robot encounters a *positioning* issue and asks for help from the human. In (l) the human steals the blue leg before it gets picked up, resulting in a *missed* issue. The human then follows the dialogue to allow the robot to try again. In (q) the robot encounters an *unreachable* error. The human then follows the dialogue to hand the robot to the object so it can be placed.

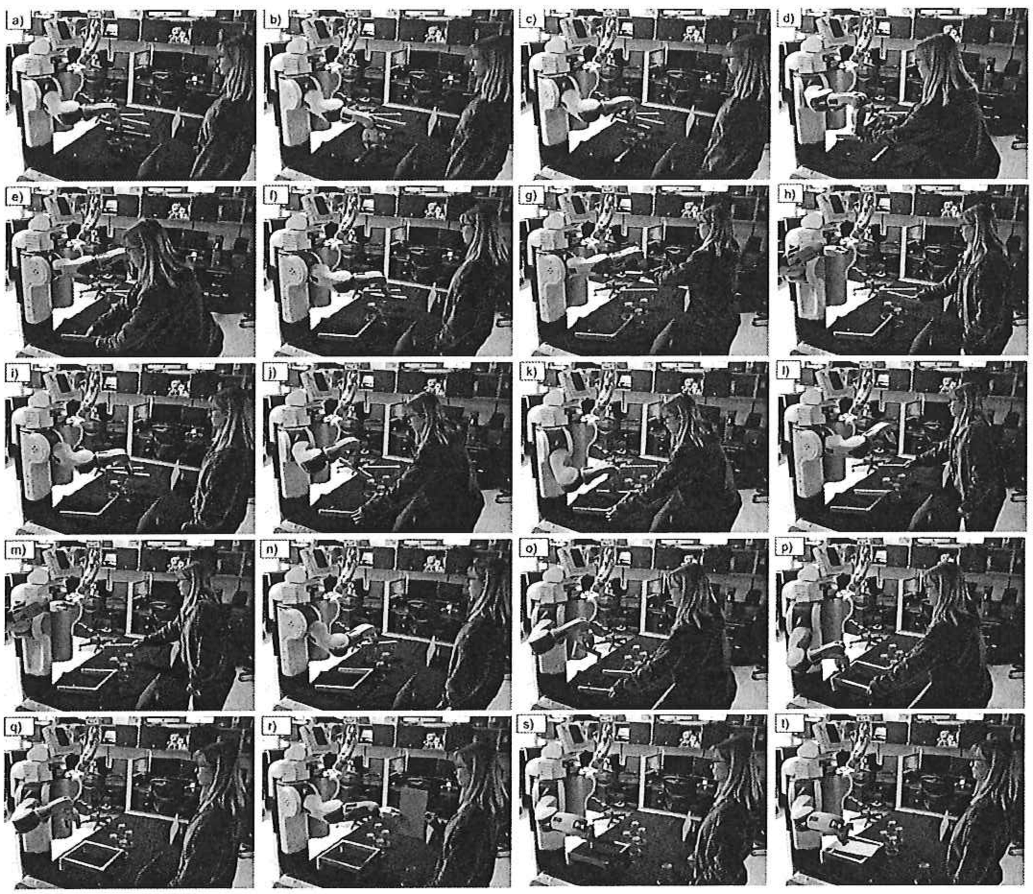


Figure 5. Experimental results for handling failures through dialog.

To demonstrate the handling of coordination during human-robot task execution, a distributive task between a human and a robot was designed. The task was performed in a lab environment with a human and a Baxter humanoid robot standing on opposite sides of a table containing the objects as shown in Figure 6. The 3D location of each object is provided by

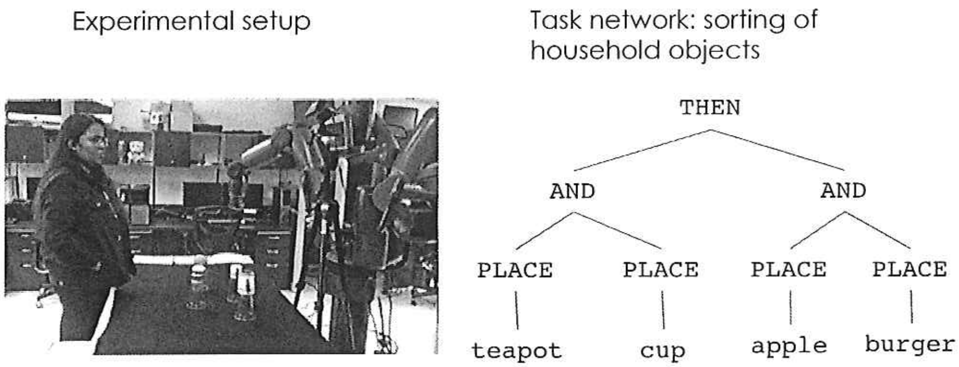


Figure 6. Experimental setup for human-robot collaboration.

our vision system. A Kinect v1 camera, next to the Baxter was used to observe human intent, and a Kinect v2 camera on top of the Baxter's head was used for the robot end of the architecture. A joint tea-making task was designed based on the task tree which encodes the constraints of both THEN and AND nodes and the scenario contained both overlapping and non-overlapping sub-tasks between human and robot. During the experiment, the robot and the human both went for the cup to pick and place, which resulted in a collision. The robot started to negotiate; the human told the robot to finish the current task. While the robot was performing the task, the human moved to the next object, which was picking and placing the teapot. A collision was again detected as the human and the robot were both going for the apple which started the dialogue between the robot and the human again. The human wanted to perform the current task and informed the robot. The robot stopped going for the apple and moved to the next task to pick and place the burger.

Findings 2. Active vision-based approach for object detection and manipulation in cluttered environments

Tables 1 and 2 below show the results of our active vision system for our two approaches for classification fusion (Dempster-Shafer and Bayesian fusion respectively). The results indicate that the active perception system with camera movement generated by our proposed planner provides better results than a static two camera system, in particular for the situations in which the objects are partially occluded.

Table 1. Performance results of the proposed active vision system.

Performance Measure	Single Camera		Active Perception (no camera motion)		Active Perception (with camera motion)	
	Non-occluded	Partially occluded	Non-occluded	Partially occluded	Non-occluded	Partially occluded
Macro-Averaging Precision	0.860	0.745	0.944	0.903	0.989	0.928
Macro-Averaging Recall	0.855	0.700	0.933	0.833	0.977	0.855
Accuracy	0.846	0.684	0.933	0.833	0.977	0.855
F_1 Score	0.857	0.721	0.938	0.866	0.982	0.890

Table 2. Performance results of the proposed active vision system with Bayesian fusion.

Performance Measure	Single Camera		Active Perception (no camera motion)		Active Perception (with camera motion)	
	Non-occluded	Partially occluded	Non-occluded	Partially occluded	Non-occluded	Partially occluded
Macro-Averaging Precision	0.884	0.765	0.922	0.878	0.967	0.907
Macro-Averaging Recall	0.877	0.722	0.911	0.811	0.955	0.833
Accuracy	0.868	0.706	0.911	0.811	0.955	0.833
F_1 Score	0.880	0.742	0.916	0.843	0.960	0.868

Findings 3. Pose detection system for dynamic object manipulation

We tested the pose detection and grasping system on two different tools: a wrench and a screwdriver. For each object, we defined two different types of grasps: a top grasp (in which the object is picked up from the top of the long end) and a side grasp (in which the object is grasped perpendicular to its long axis). We placed the objects on a tripod which allowed us to move them in various poses and had the robot plan a path to grasp them using both types of grasps. Figure 7 below shows several examples of the top and side grasps.

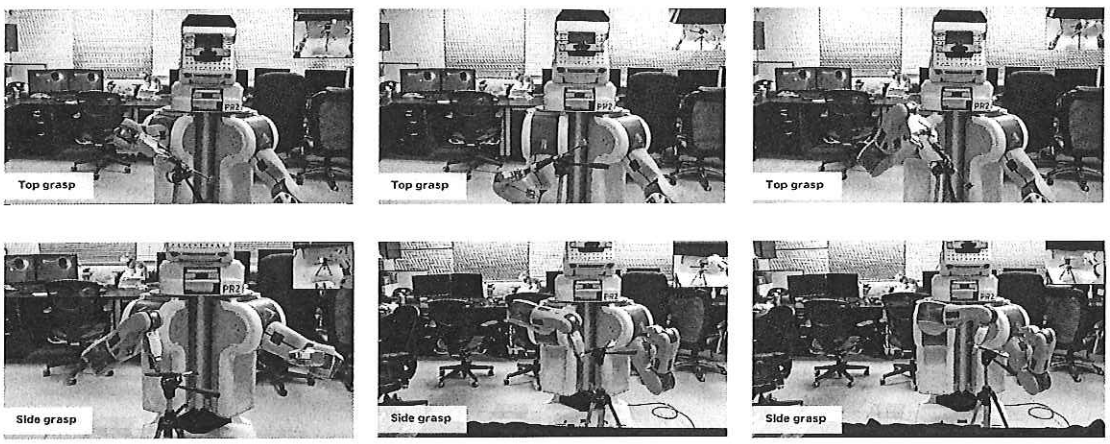


Figure 7. Grasping results for two objects and two different grasps.

Findings 4. Extended approach for language-based task learning & joint human-robot task execution

The following examples illustrate results from translating natural language instructions into executable controllers.

Push the large tall chair around the small pretty table
`(push tall_large_chair pretty_small_table around)`

Move the tool above the table
`(move tool table above)`

Put the small yellow book under the brown round table
`(put yellow_small_book brown_table under)`

Move the sharp tool near the tiny red box
`(move sharp_tool red_tiny_box near)`

Move the big purple ball from the tiny red table
`(move purple_big_ball red_tiny_table from)`

Chase the big man to the right door
`(chase big_man right_door to)`

Figure 8. Examples of multiple adjectives and location prepositions.

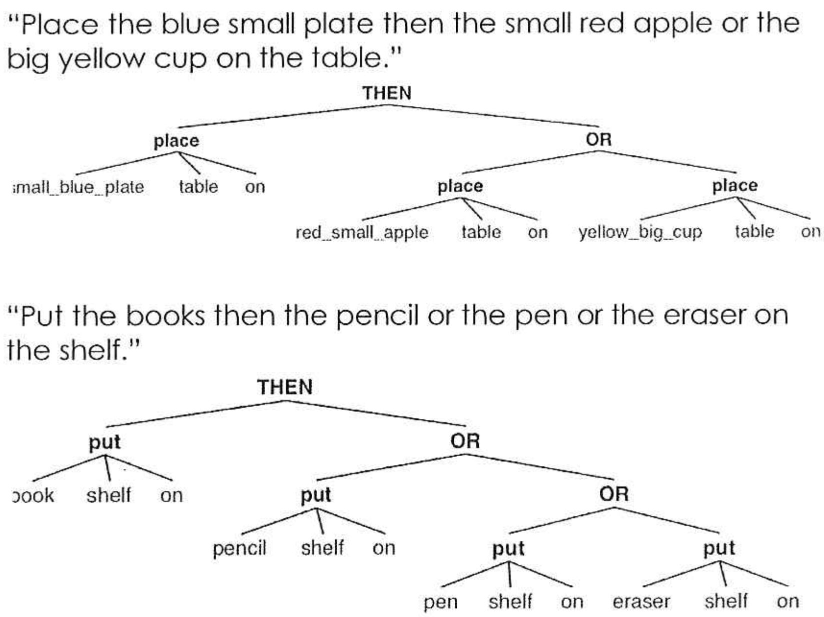


Figure 9. Example instructions with complex execution constraints.

[Appendix (Data and Charts)] Not applicable or no information to report.

4. Plans and Upcoming Events

Recommendations for Future Work: This report covers the final year of the award. For the future, new approaches for multi-robot systems (distributed control under communication constraints, learning) will be discussed with the ONR program manager.

5. Transitions and Impacts

Not applicable or no information to report.

6. Collaborations

Not applicable or no information to report.

7. Personnel

Principal investigator: Monica Nicolescu (1.5 months)
Co-investigators: Mircea Nicolescu (1.2 months), David Feil-Seifer (1.2 months)

8. Students

Graduate students: 5; Undergraduate students: 1.

9. Technology Transfer

Not applicable or no information to report.

10. Products, Publications, Patents, License Agreements, etc.

Publications resulting from this project during the current reporting period:

Journal articles:

- Seyed Pourya Hoseini Alinodehi, Janelle Blankenburg, Mircea Nicolescu, Monica Nicolescu, David Feil-Seifer, "Active Eye-in-Hand Data Management to Improve the Robotic Object Detection Performance", to appear in *Computers - Special Issue on Vision, Image and Signal Processing*, 2019.
- Banafsheh Rekabdar, Luke Fraser, Monica Nicolescu, Mircea Nicolescu, "A Real-Time Spike- Timing Classifier of Spatio-Temporal Patterns," *Neurocomputing*, vol. 311, pages 183-196, October 2018.

Conference publications:

- Monica Nicolescu, Natalie Arnold, Janelle Blankenburg, David Feil-Seifer, Santosh Balajee Banisetty, Mircea Nicolescu, Andrew Palmer, Thor Monteverde, "Learning of Complex-Structured Tasks from Verbal Instruction", in *Proceedings of IEEE-RAS International Conference on Humanoid Robots,* 2019.
- Bashira Akter Anima, Mariya Zagainova, S. Pourya Hoseini Alinodehi, Muhammed Tawfiq Chowdhury, Janelle Blankenburg, Monica Nicolescu, David Feil-Seifer, Mircea Nicolescu, "Collaborative Human-Robot Hierarchical Task Execution with an Activation Spreading Architecture", in *Proceedings of the International Conference on Social Robotics*, 2019.
- Seyed Pourya Hoseini Alinodehi, Janelle Blankenburg, Mircea Nicolescu, Monica Nicolescu, David Feil-Seifer, "A Dual-Camera Robotic Vision System Based on the Concept of Active Perception", in *Proceedings of the International Symposium on Visual Computing*, 2019.

11. Point of Contact in Navy

Not applicable or no information to report.

12. Acknowledgement/Disclaimer

This work was sponsored by the Office of Naval Research (ONR), under grant (or contract) number N00014-16-1-2312. The views and conclusions contained herein are those of the authors only and should not be interpreted as representing those of ONR, the U.S. Navy or the U.S. Government.

Section II: Project Metrics

Grant or Contract Number: N00014-16-1-2312
Date Prepared: October 29, 2019
Project Title: Designing Collaborator Robots for Highly-Dynamic Multi-Human, Multi-Robot Teams
Annual Summary Report: FY2018
PI:

Monica Nicolescu, (775) 784-1687, monica@unr.edu, University of Nevada, Reno
Mircea Nicolescu, (775) 784-4356, mircea@cse.unr.edu, University of Nevada, Reno
David Feil-Seifer, (775) 784-6469, dave@cse.unr.edu, University of Nevada, Reno

Metrics

Number of faculty supported under this project during this reporting period: 3

Number of post-doctoral researchers supported under this project during this period: 0

Number of graduate students supported under this project during this reporting period: 4

Number of undergraduate students supported under this project during this period: 1

Number of scientists / engineers / technicians supported under this project during this reporting period: 0

Number of refereed publications during this reporting period for which at least 1/3 of the work was done under this effort: 5

Number of publications (all) during this reporting period: 5

Number of patents during this reporting period: 0

Number of M.S. students graduated during this reporting period: 0

Number of Ph.D. students graduated during this reporting period: 0

Awards received during this reporting period: 0

Invited talks given: 0

Conferences at which presentations were given (not including invited talks above): 3

1. Financial information

This has been the final year of the project, the financial information is reported for the period from 06/14/18 to 08/31/19.

FY 2018	Total Budget	Obligated This Period	Obligated Cumulative	Expended This Period	Expended Cumulative	Grant/ Contract Period of Performance
6.1 (Basic Research Funding)	$656,511	$307,115.60	$656,511	$307,115.60	$656,511	04/01/2016 08/31/2019

2. Administrative notes and other items of interest

Not applicable or no information to report.