REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
10/29/2013

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
8/1/2009 to 7/31/2013

4. TITLE AND SUBTITLE
Context-Based Intent Understanding for Autonomous Systems in Naval and Collaborative Robot Applications

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-09-1-1121

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Monica Nicolescu, Mircea Nicolescu, Sushil Louis

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., MS 325
Reno, NV 89557-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

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

10. SPONSOR/MONITOR'S ACRONYM(S)

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

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The problem: Understanding intent is a critical aspect of communication among people and for many biological systems. While people are very good at recognizing intentions, endowing an autonomous system (robot or simulated agent) with similar skills is a more complex problem, which has not been sufficiently addressed in the field. The issue of intent recognition is particularly important in situations that involve collaboration among multiple agents or assessment of potential threats. In the former case collaboration can be greatly enhanced, while in the latter case dangerous situations can be detected before any harmful actions can be finalized. In this project, we propose to develop methodologies for intent understanding, with specific focus on autonomous systems for naval and collaborative robotics applications. The main research problems we will address in this project are to: 1) develop tools for understanding the high-level intentions of groups of agents, 2) develop algorithms for intent understanding based on contextual information, 3) develop vision-based techniques for learning of contextual information, and detection and identification of objects of interest.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		8	Monica Nicolescu
					19b. TELEPHONE NUMBER *(include area code)* 775 784 1687

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

1. Technical Report

a. Scientific and Technical Objectives

Understanding intent is a critical aspect of communication among people and for many biological systems. This is particularly important in situations that involve collaboration among multiple agents or assessment of potential threats. During the recent years, there has been an increased interest in using robotic technologies for security and defense applications, in order to reduce the danger for the people involved. In the context of these applications, being able to automatically detect any threatening situations is of critical importance. This reduces to the problem of understanding the intent of other agents, from their current actions, before any attack strategies are finalized.

The primary objective of this work is to design an effective and robust system for intent understanding that will provide reliable detection of intended activities for autonomous systems in both naval and service robotics applications. Specifically, we will work toward the following objectives:

- Develop tools for understanding intentions in large-scale systems

- Design algorithms that rely on extensive use of contextual information for intent understanding

- Develop vision-based techniques for learning of contextual information, and detection and identification of objects of interest

- Integrate the above capabilities into two prototype systems that will be tested under naval-type mission scenarios and a collaborative robot scenario.

b. Approach

Our approach to reaching our goals consists of the following main steps:

1) Develop a unified framework for intent understanding. The proposed approach relies on the use of extensive contextual information in order to identify the correct intentions of agents in naval and robot domains. This contextual information will be incorporated both at a low level (for detection of basic intentions) and at a high-level (for the detection of complex intentional activities). The main sources of contextual information we will consider are: object affordances, history, domain knowledge, general (space, time, etc.) and the actor's beliefs, perceptions, desires or personality. The framework includes four key components, described in detail below.

2) Develop techniques for the detection and tracking of relevant agents and or objects in the environment. Once detected, their 3-D positions, trajectories and speed are determined, in order to provide this information to the intent recognition module. We will leverage our current work in this area and extend our system's capabilities for a wide range of situations: different perceptual requirements depending on the particular scenarios, as well as different assumptions (moving vs. static cameras, moving vs. static objects of interest, generic detection of people/classes of objects vs. recognition of specific persons/objects, availability of

pre-learned models). We will also incorporate additional sensor data, such as 3D information from stereo cameras or laser rangefinders into our techniques for detection and tracking.

c. Concise Accomplishments

During the last reporting period we worked in the following research directions:

1) We refined a distributed architecture for intent recognition, based on activation spreading. Within this architecture, the hierarchical structure of activities and

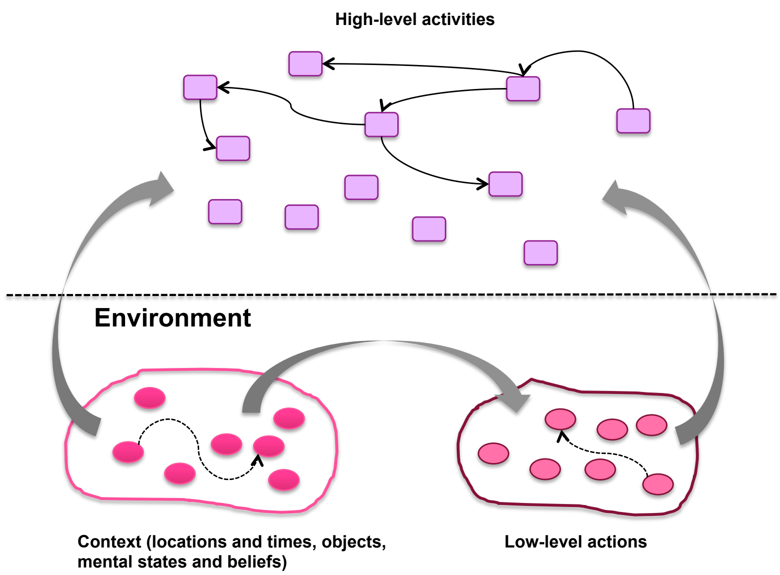

Figure 1. Intent recognition using activation spreading.

contextual information is represented in an interconnected network of nodes passing messages to each other (Figure 1).

2) We refined our infrastructure for the naval simulation domain to enable detection of threats posed by coordinated groups of boats. Our work consists of: i) updated models for detection of low-level intentions, ii) new classifiers for detecting an "intercept" behavior, iii) integration of the distributed architecture with the naval simulator, and iv) quantitative evaluation of the system's performance in various experimental scenarios.

3) We applied our work to a naval simulation domain. In this research, we demonstrated the ability to recognize coordinated attacks by multiple boats, using the distributed activation spreading architecture.

d. Expanded Accomplishments

During this reporting period we made progress in the following directions:

1) Intent Recognition using an activation spreading architecture.

As a part of this work we refined our previous distributed architecture prototype that uses the principle of activation spreading in interconnected networks. Similar to Anderson's spreading activation theory of memory, we assume that information regarding activities (such as their temporal or hierarchical structure) as well as related contextual information (such as locations, time, objects present in the visual field, mental states, beliefs) is represented as an interconnected network. The observation of certain states or basic actions in the environment increases the strength of corresponding nodes in the network, which begin to send activation to nodes that represent related activities. Activities that accumulate the highest level of activation are considered most likely to be those actually performed by the agent. Using the known temporal structure of the activities we can predict potential future actions of the agent before they are achieved.

During this period, we redesigned the basic structure of the architecture using the latest version of Scala, a functional and object-oriented language. Scala provides actor concurrency for distributed and asynchronous message passing, as needed for our network. We used a graph language to represent the structure of our network: each node in the graph is an actor and an edge from A to B indicates that A sends activation messages to B. The messages sent are activation messages that contain a single real number (the strength of activation passed on to neighbors and a type. The type can be "input" for low-level intentions and context or "internal" for high-level intentions.

2) Refined naval simulator infrastructure.

We previously developed a 3D, physics-based simulation engine, which provides the following features and capabilities:

- Large set of boat models, ranging from small cigarette boats and fishing boats, to aircraft carriers and destroyers
- Ability to run scenarios with a large number of boats (80 to 100)
- Ability to create individual controllers for each boat using a GUI, based on a set of basic boat behaviors
- Ability to generate and store multi-boat scenarios, with each boat automatically running its own controller

We extended our infrastructure for the naval simulation domain to enable detection of threats posed by groups of coordinated boats. Our extensions consist of:

i) updated models for detection of low-level intentions. We have re-structured and retrained models for all the low-level intentions that we have previously developed: *approach* (one boat gets closer to another), *follow* (one boat keeps a constant distance and bearing with respect to another), *overtake* (pass in front of another boat, coming from behind, going in the same direction), and *pass* (go by another boat, coming from opposite direction).

ii) new classifiers for detecting an "intercept" behavior. In order to implement the new scenarios, we modeled and trained a classifier for an "intercept" behavior. A boat is considered to be intercepting another, if it follows a course that will intersect with the course of the other boat at some point in the future.

iii) integration of the distributed intent recognition architecture with the naval simulator. The distributed architecture has been previously used solely in the robotic domain. During this period, we integrated it with the naval simulator, which allows us to test its performance in the naval domain.

iv) quantitative evaluation of the system's performance in various experimental scenarios. To test the baseline accuracy of the HMM-based approach to low-level intent recognitions, we trained models for 5 different intentions: approach, pass, overtake, follow, and intercept. We then generated 200 two-agent scenarios, resulting in 40 test scenarios for each of the trained intentions. All of our statistics represent the average performance of the intent recognition system over the 40 relevant scenarios. For a quantitative analysis of the intent recognition system, we used three standard measures for evaluating HMMs:

- *Accuracy rate*: the proportion of test scenarios for which the final recognized intention was correct
- *Average early detection*: $\frac{1}{N}\sum_{i=1}^{N}\frac{t_i^*}{T_i}$, where N is the number of test scenarios, T_i is the total runtime of test scenario i, and t_i^* is the earliest time at which the correct intention was recognized consistently until the end of scenario i.
- *Average correct duration*: $\frac{1}{N}\sum_{i=1}^{N}\frac{C_i}{T_i}$, where C_i is the total time during which the correct intention was recognized for scenario i.

For reliable intent recognition, we want *accuracy rate* and *average correct duration* to be close to 100%, and *average early detection* to be close to 0%. The results of our experiments are shown in Table 1.

Scenario	Accuracy Rate (%)	Avg. Early Detection (%)	Avg Correct Duration (%)
Approach	100	8.95	90.9
Pass	100	68.0	96.5
Overtake	100	56.8	64.6
Follow	100	1.92	99.3
Intercept	100	11.3	88.8

Table 1. Quantitative evaluation of low-level intent recognition module

As can be seen, the intent recognition system performs well in terms of early detection for the approach, intercept, and follow behaviors, recognizing them consistently within the first 12% of the completion of the action. These results are consistent with the current state of the art for single-agent intent recognition methods.

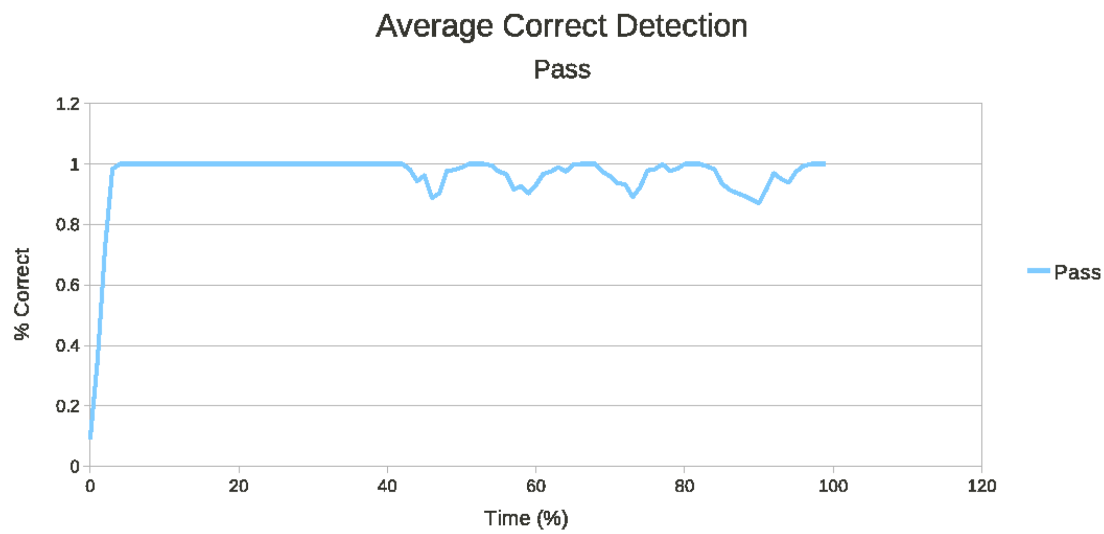

Figure 2. The average correct detection of *pass* over time.

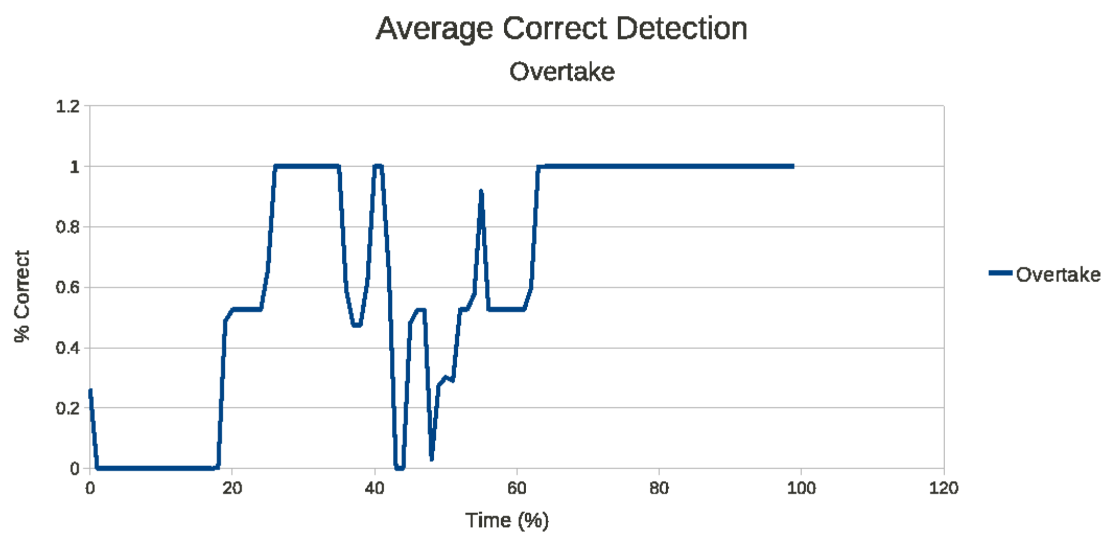

Figure 3. The average correct detection of *overtake* over time.

Figure 2 and Figure 3 provide an explanation for the poor performance of the pass and overtake behaviors. These figures show the percent accuracy of each intention over the duration of the run. For instance, if 20 out of the 40 runs correctly recognized *pass* 50% of the way through a scenario, then the value of the graph in Figure 1 at t=50 will be .5. From this analysis, we can see that both *pass* and *overtake* are correctly recognized for the majority of the duration of each scene (as borne out in Table 1), but consistently fail to be recognized when the agents have drawn abreast of each other. This is likely due to a lack of distinguishing evidence variables at this time. Given the evidence variables discussed above, the only difference between *pass* and *overtake* at this point would be "change in angle from target agent to acting agent," which is likely not enough to result in a distinct classification.

We also evaluated the **effectiveness of parallelizing the intent recognition** process. Toward this end, we implemented both serial and parallel versions of the intent recognition algorithm and ran them on scenes containing varying numbers of agents. We then recorded the average frame rate over each scene (with one frame defined as a single iteration of the intent recognition algorithm, from symbol generation to selection of most likely intent), with the results shown in Figure 4. We can see that while the performance of the serial implementation of the intent recognition process quickly drops below an acceptable frame rate for real-time systems, the parallel implementation maintains a speed of about 40 frames per second, which is definitely adequate for performing in real-time. The intent recognition problem as presented in this thesis has a computational complexity of $O(n^3)$ (intentions must be calculated for each pair of agents, from the perspective of each agent). Thus, we can expect that the intent recognition system will continue to perform in the neighborhood of 40 fps as long as $n < \sqrt[3]{m}$, where n is the number of agents and m is the maximum number of threads provided by the GPU. On our system (which uses the Tesla C2050), this means that we should be able to continue performing intent recognition in real-time as long as n < 30,000.

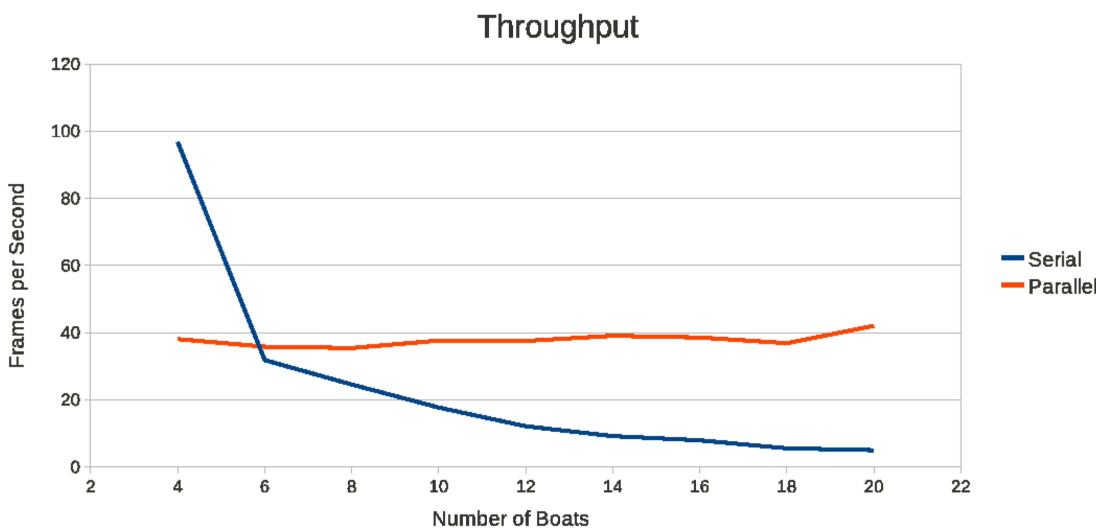

Figure 4. Performance of serial implementation of intent recognition vs. parallel implementation.

3) Intent recognition for the naval domain.

We created 6 different scenarios in our simulator in which naval vessels needed to recognize potentially hostile intentions (approach and intercept) as enemy ships maneuvered to attack.

In the *Straits of Hormuz* scenario, a convoy of naval vessels is attempting to traverse the straits. As they do so, a pair of other ships pass close by the convoy, creating a distraction. Shortly after this, more ships break free of a group of trawlers, and begin a suicide run towards the convoy in an attempt to damage it. The *San Diego* scenario is constructed similarly. Here, a group of naval vessels is attempting to exit the San Diego harbor. As they travel towards the harbor mouth, a ship that had been behaving like a fishing boat comes about and begins a run towards the navy vessels. In *hide*, the naval vessels are traveling through a channel, while passing some container ships. As this happens, a small boat accelerates to a position behind one of the container ships and hides there until it is abreast of the navy vessels. At this point, it breaks from hiding and attacks the navy vessels. *Blockade* and *Hammer and anvil* are examples of some scenarios in which more complex intentions (in which agents must cooperate to perform a task) may occur. In *blockade*, a naval vessel is attempting to pass through a channel when some other ships emerge from hiding behind nearby islands and intercept it, forming a blockade. *Hammer and anvil* begins similarly, but once the channel is blocked by the blockading ships, an additional pair of ships approaches from behind the naval vessel in order to attack and cut off escape. In order for these techniques to work, we must also be able to accurately recognize the low-level intentions (intercept and approach) which make up the overall attacks. In performing a quantitative analysis of the more complex scenarios, we first define *key intentions* as those intentions that make up actions, which are threatening to the naval vessels in the scene. For instance, in the *hide* scenario, the container ships may have the intention of *passing* the naval vessels, but this would not be a key intention. However, the aggressive ship must *overtake* a container ship in order to hide behind it, and must *approach* the navy vessels in order to attack them, and both of these would be considered key intentions. For the purposes of determining the performance of the intent recognition in the complex scenes, we will focus on the average early detection for key intentions in each scene, and the accuracy rate for those key intentions as well. The accuracy rate for our system is 100% for key intentions in the complex scenarios. In each of the 5 scenarios all of the key intentions were correctly identified. In addition, it can be seen in Table 2 that the early detection rate for the key intentions is below 13%. In every case, the key intentions were recognized almost as soon as they began.

Scenario	Early Detection (%)
Straits of Hormuz	1.50
San Diego	2.31
Hide	3.03
Blockade	0.0
Hammer and Anvil	5.04

Table 2. Intent recognition in complex scenarios.

e. Work Plan

This is the last year of the project. We plan to further extend our work as a part of a recently started ONR project.

f. Major Problems/Issues

N/A.

g. Technology Transfer

Our physics based 3D naval simulator is currently used every day at SWOS for training exercises with the Full Mission Bridge.

h. Foreign Collaborations and Supported Foreign Nationals

N/A.

2. Publications, Patents, Presentations and Awards

- R. Kelley, A. Tavakkoli, C. King, M. N. Nicolescu, M. Nicolescu, "Context-Based Bayesian Intent Recognition", in IEEE Transactions on Autonomous Mental Development, *, 4*(3), 215-225.
- Kelley, R., Tavakkoli, A., King, C., Ambardekar, A., Wigand, L., Nicolescu, M. N., Nicolescu, M. (2013). Intent Recognition for Human-Robot Interaction. *Plan, Activity, and Intent Recognition*. Elsevier, in press.
- Siming Liu, Sushil Louis and Monica Nicolescu, "Using CIGAR for Finding Effective Group Behaviors in RTS Game", in proceedings IEEE Conference on Computational Intelligence and Games, 2013.
- Siming Liu, Sushil Louis and Monica Nicolescu, "Comparing Heuristic Search Methods for Finding Effective Group Behaviors in RTS Game", IEEE Congress on Evolutionary Computation, 2013.

3. Documentation of award participants

- Monica Nicolescu (PI)
- Mircea Nicolescu (Co-PI)
- Sushil Louis (Co-PI)
- Richard Kelley (student)
- Daniel Bigelow (student)
- Liesl Wigand (student)