Naval Research Laboratory
Washington, DC 20375-5320

Final Report: Goal Deliberation for UUV Control

MARK A. WILSON
Email: mark.wilson@nrl.navy.mil
DAVID W. AHA
Email: david.aha@nrl.navy.mil

Navy Center for Applied Research in AI (NCARAI)
Information Technology Division
Naval Research Laboratory, Code 5510

Keywords: Underwater vehicles, unmanned vehicles, robots, goal reasoning, planning, planning and acting

12 September 2021

This **NCARAI Technical Note** is the final project report for the NRL IT Focus Area Base Program project *Goal Deliberation for UUV Control* (PI: David W. Aha).

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

1. Introduction

Goal reasoning (GR) is the ability for AI agents to dynamically reason about and change their goals to better align their activities with real-world conditions during execution. In previous work we have integrated existing GR models onboard unmanned underwater vehicles (UUVs) and demonstrated their ability to interact with UUV control systems and direct the UUVs to accomplish different goals based on sensor feedback from the environment. However, these GR models were based on discrete models of action, and providing them with dynamic, up-to-date cost estimation for UUV maneuvers was not straightforward. Further, these models used reactive, single-goal selection and did not consider the impact of changing goals on the long-term success of the mission. The objective of this project was to address these needs by developing a new GR model for use in embodied agents, and to demonstrate its efficacy in the UUV domain.

2. Technical Objective

Our objective was to design, develop, and evaluate novel strategies for controlling UUVs using an integration of goal reasoning and route planning techniques. This is, to our knowledge, the first time that a GR architecture will leverage information about maneuvering to select goals for UUV control. Such techniques will provide the Navy with the core components to advance the state-of-the-art on the deliberative control of UUVs.

3. Technical Approach

Our approach to goal deliberation for UUV control requires, among others, the following capabilities (which are detailed below):

1. A GR model of spatial domains for use in embodied agents
2. A goal selection function to select and order a sequence of potential mission goals to pursue
3. A method for tightly integrating GR with route planning
4. A route planner to generate vehicle trajectories and estimate their costs

3.1 Goal reasoning model

Goal reasoning (GR) provides agents with the ability to continuously deliberate about what goals to pursue. We use a GR model that directly represents spatial entities relevant to the GR process for embodied agents, including a designated operations area, regions of interest (e.g., for surveying the seabed), points of interest (e.g., mine-like objects on the seafloor or other vessels in the operations area), and an artificial multi-cost time-dependent road network for navigation. Rather than representing discrete actions that the agent may pursue in any order according to the output of a planner, we adopt a goal ontology appropriate for robotics applications, similar to that proposed by Balakirsky *et al*. (2017). In this model, the embodied agent accomplishes goals by completing *tasks* (e.g., surveying an area of the seabed). The agent completes tasks by executing a defined set or sequence of *behaviors* that encode reactive control instructions, possibly parameterized by the particulars of a given instantiation of the task (e.g., the behaviors may direct the robot to drive long, straight lines, with a sensor activated, over a specified location to complete a seabed survey). This model captures the nature of robotics applications more accurately and naturally than a discrete action-based planning model.

3.2 Goal selection

Wilson *et al*. (2013) introduced a variant of the ARTUE GR agent (Klenk *et al*., 2013) with a principled goal selection procedure; it uses a set of domain-independent *motivators* to rank goals by their estimated utility and selects a single goal to pursue. In this project, we extended this approach by computing the estimated costs of executing plans for achieving specified goals using tightly-integrated task and route planning. In particular, our goal selection algorithm: (1) identifies a set of candidate goals G for the agent to examine given the current world state s; (2) generates plan(s) P' for candidate subsets of goals, $G' \subseteq G$, using estimated execution cost as a direct cost for executing G' (e.g., energy, time); and (3) computes additional indirect costs for each P' (e.g., risk for navigating too close to other vessels). The algorithm optimizes results by searching for the plan with the highest reward without violating budgets for any direct or indirect cost.

3.3 Integrating goal reasoning with route planning

Most prior work on GR agents has used symbolic planners that reason about goals at a higher level of abstraction than motion goals. Thus, some method is needed to integrate GR with the intricacies of continuous robotic motion. We developed a new GR model that incorporates route planning into the task planning model, interacting directly with the model's representations of spatial components in the embodied GR problem and completing tasks as the central element of accomplishing goals. Every task is spatially situated and thus requires the agent to navigate to a starting spatial configuration (i.e., to survey a region, the agent must be present at that region). For the UUV domain, we adopt a principally 2D configuration space (i.e., x-y or latitude-longitude coordinates), but for future work or other domains, a higher-dimensional space may be appropriate. We incorporate a probabilistic roadmap (PRM) (Kavraki *et al*. 1996), with the requirement that it intersect every defined region of interest for the agent, and dynamically compute time-dependent costs for its edges during execution. For example, the risk cost of an edge will be high at times when other vessels are near it, and low otherwise; the navigation cost of an edge will be higher when ocean currents are counter to the direction of motion than when they are aligned with it. Using this time-dependent PRM, we are able to compute the cost of maneuvering from one task to the next, based on time of navigation, during plan construction.

3.4 Route planning

We previously integrated a hierarchical task network (HTN) planner, PHOBOS, with a reactive UUV controller to execute motion tasks. However, it cannot estimate the costs for executing motion tasks. Further, PHOBOS uses the PDDL planning language, which can represent symbolic abstract planning tasks but cannot easily represent spatial navigation in complex environments. To address these limitations, we employ the time-dependent PRM described above. Optimal routes can be computed across networks with time-dependent costs using an adapted A* planning algorithm, as described by Chabini and Lan (2002). To compute a route from one task to the next, we find the nearest nodes in our PRM and solve for an optimal path between them, using a user-specified weighting scheme to combine the multiple costs (e.g., risk and navigation time) for each edge. Departure time is computed based on navigation time to previous tasks and task models defining expected time cost for executing each task.

3.5 Evaluations

We have conducted evaluations in simulated environments using the uSimMarine utility included with MOOS-IvP (a robotics architecture we have used in prior studies), focusing on testing variants of our proposed GR agent in mine-countermeasures missions. We have also performed some testing of an early version of our agent on NRL's Bluefin-21 vehicle *Reliant* in Boston Harbor (in additions to earlier and more extensive in-water testing with smaller Oceanserver Iver2 UUVs at NRL's Chesapeake Bay Facility location), although technical difficulties with available UUVs and COVID-19 restrictions have limited our ability to perform at-sea trials since March 2020.

4. Technical Progress

4.1 Goal Reasoning Model

In FY18, we developed the first iteration of our spatial goal reasoning model for embodied agents. This model extended our prior work (Wilson et al., 2018), for which we received an NRL 2018 Alan Berman Publication Award. In that work, we developed and demonstrated methods that enable a UUV to respond dynamically to its environment by modifying its goals. However, these goals should be prioritized, in part, by estimating the costs for achieving them, which are continuous and spatially grounded. To address this need, we modeled goal accomplishment using predefined libraries of robotic behaviors to complete tasks. We also modeled the environment for the goal reasoning problem as a spatial region containing defined subregions and points relevant to tasks and to navigation. Finally, we developed a generic task model that permits authors to write domain-specific task descriptions expressing applicability, costs, and options for partially completing a task for use in planning, and wrote descriptions of common tasks in our test scenarios (i.e., survey, track, and rendezvous).

In FY19, we developed a new type of goal for the UUV to pursue. This goal permits the vehicle to loiter in one spot after completing another task. The ability to loiter allows the vehicle to wait for a threat to diminish or leave the area of operations before commencing a task that would otherwise be highly risky; by waiting, the vehicle may be able to provide an overall solution that either requires less time or puts the UUV at less risk than would otherwise be possible. Providing a loiter capability in the underlying route-planning system would result in untenable computational complexity, but implementing it as a goal permits better solutions with no change in the overall complexity of the problem.

In realistic UUV missions, there are frequently situations in which a particular goal ***must*** be accomplished as a prerequisite of another goal. For instance, in minesweeping operations, a region of interest must be surveyed to identify mine-like objects before those mine-like objects can be revisited and classified. In environmental observation scenarios, measurements of relevant features must be taken before the vehicle can surface and report its findings. Our initial development focused on integrating pathfinding and developing optimizers suitable for the time-dependent aspects of the problem. During FY20 we extended our goal-problem model to permit precedence constraints between goals and developed an incremental solution representation that tracks precedence satisfaction.

In FY21, we added two new types of goals for the UUV to pursue. In the first, we make use of topside communication capabilities on NRL vehicles by surfacing to exfiltrate a message, with future extension planned to exfiltrate data gathered during a mission. In the second, we return to a surveyed area to reacquire and classify objects of interest (e.g., mine-like objects). The reacquisition-and-classification task

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

uses the precedence-constraint capabilities added to the goal model in FY20. We implemented a formal state machine using the Boost C++ libraries to manage execution state, improving code complexity and stability in our goal reasoner. We also added a generalized ability for tasks to override certain discrepancies as a trigger for re-goaling while executing. This capability is useful for situations such as surveying in a mine-countermeasures mission, which produces numerous discrepancies as mine-like objects are discovered.

Finally, in FY21 we added support for multi-vehicle teams to the goal reasoner, including plan representation, problem representation, execution state, and the interface to the MOOS-IvP autonomy framework. We also added support for inter-vehicle communication through the MOOS-IvP framework and initial approaches for resolving conflicting data about world state between team members. Reasoning about multi-vehicle teams enhances the applicability of goal reasoning to real-world autonomy problems, and also supports the objectives of the related NRL UWFA Base Program, *Enabling Distributed Maritime Autonomy on Unmanned Platforms via Decentralized Multi-Vehicle Control* (PI: James McMahon).

4.2 Goal Selection

In FY18, we developed a preliminary model for goal selection based on required time-dependent motion plans and their associated costs. We model this as a graph of vertices, representing goals available to the agent, where edges represent time-dependent routes between two goals. The problem of choosing a high-quality sequence of goals can be modeled as choosing a subset of vertices to visit and an order in which to visit them that will maximize reward, without violating constraints on costs; this problem is generally known as the *Orienteering Problem* (OP) and is known to be NP-hard. (When an OP is subject to time-dependent costs or constraints it is known as the *Time-Dependent Orienteering Problem*.) We implemented a representation of the time-dependent goal selection problem; an efficient branch-and-bound optimizer for it (that provides interruptible, anytime generation of goal decisions in time-bounded situations); another optimizer for the goal problem based on the Greedy Randomized Adaptive Search Procedure (GRASP) algorithm, which locally optimizes a randomized but greedy initial solution; and a MOOS application that can use these tools to guide a vehicle's behavior.

In FY19, we developed another optimizer based on evolutionary computation algorithms that may provide better performance on scenarios in which tasks should NOT be included (vice the GRASP solver, which does not remove suboptimal tasks once added during its search procedure). By applying random task selection, both the GRASP and evolutionary solvers may achieve comparable results while requiring far less search time than the branch-and-bound method. Additionally, they are less susceptible to adversarial problems (i.e., environmental conditions designed to induce poor performance in deterministic algorithms such as branch-and-bound).

In FY20, we implemented an ant-colony optimizer for the goal-selection orienteering problem to provide a broad base of comparison among our proposed optimization methods. Ant-colony optimization has been employed as a general metaheuristic approach in variants of the orienteering problem and related problems such as the traveling salesman problem (e.g., Verbeeck et al., 2014; Dorigo & Gambardella, 1997). Ant-colony optimizers share some similarities with Q-learning; they explore by repeatedly finding a path (i.e., a sequence of goals) through the problem graph. High-value arcs receive "pheromone" to signal future iterations of the solver, which is biased to select arcs with higher pheromone levels.

Also during FY20, we developed two optimizers based on related existing solutions, to solve our extended variant of the orienteering problem with precedence constraints. The first is based on unpublished work by James McMahon and collaborators for solving multi-vehicle traveling salesman problems with precedence constraints. This approach uses Khan's algorithm to arrange precedence-constrained tasks into "layered" directed acyclic graphs, where tasks within layers do not depend on each other or on tasks in subsequent layers. Each layer can then be solved as an independent assignment problem without respect to precedence.

Adapting this approach for use on the orienteering problem weakens the assumption of independence between layers: In contrast to the traveling salesman problem, the orienteering problem allows goals to be dropped if they are deemed too expensive, but doing so in a precedence-constrained problem may imperil the ability of the agent to accomplish high-reward goals that depend on the dropped goal. However, the approach has the values of relative simplicity and the ability to reuse our existing optimizers for solving individual layers, which in our adaptation are essentially smaller orienteering problems without precedence constraints. We conducted preliminary investigations into techniques for allocating portions of the problem budget to each layer and achieved reasonable performance on static test problems using a fuzzy allocation, wherein budget is allocated proportional to the possible reward for accomplishing all tasks in a layer, then the allocation is perturbed according to a normal distribution.

The second approach we developed was adapted from a bee colony metaheuristic approach to time-dependent team orienteering with time window constraints (Yu et al., 2019). Time window constraints stipulate that visits to certain goals are valid only during particular time windows. While this differs from the precedence constraints in our variant, the approach deals with the time window constraints by simply dropping goals that violate them and attempting to rectify the violation in subsequent iterations, which is adaptable to our precedence constraints. The bee colony approach operates by running parallel local searches in the solution space, using a simulated annealing acceptance rule. We also adapted the existing approach by replacing some methods for generating neighbor solutions in the local search with related methods that are suitable for our particular problem variant.

In FY21 we developed and implemented new optimizers for the precedence constrained goal selection problem, based on Khan's algorithm , which we used in FY20 for our GRASP optimizer. The new optimizers were based on the existing ant colony, evolutionary, and branch-and-bound optimizers for the unconstrained problem. These optimizers will provide a broad base of comparison for approaches to solving the precedence constrained time-dependent orienteering problem, along with approaches lacking GR capabilities (discussed below). Additionally, we updated all optimizers to support assignment of tasks across multiple vehicles, enhancing their utility for teams of autonomous UUVs.

4.3 Route planning

In FY18, we first developed a model for incorporating route planning into goal reasoning. To do so, we modeled these time-varying navigation costs using a probabilistic roadmap defined over an operations area, enabled human operators to express preferences for an agent's behavior, and defined a goal utility function that combines rewards and costs. We implemented a library for solving time-dependent route planning problems using this model.

In FY19, we integrated ocean current predictions into the cost models for plan selection. This required updating our agent to use time instead of distance as the primary cost, parsing the ocean current estimates from a given data file, and applying these estimates to edges in a road network in a geometrically appropriate fashion.

In FY20, we determined that NRL's Bluefin 21 inch vehicles (i.e., Reliant and Black Pearl) are more dynamically constrained than originally allowed for in the pathfinding model underlying our goal sequence selection algorithm. To address navigation issues encountered with Reliant and Black Pearl during testing in FY18-19, we revised the pathfinding model to permit constructing paths using only sufficiently-distant neighbor nodes, which allows the vehicle more space to turn while traversing a path in water. Additionally, we investigated the effect of this change on overall path length by conducting Monte-Carlo tests of navigation between points in representative road maps, and addressed a corner case in which path length was significantly adversely affected by bypassing the pathfinding algorithm for sufficiently close endpoints.

4.4 Evaluations

In FY18, we tested our first iteration of our new models for goal reasoning in simulation and integrated with UUV platforms for in-water testing (using OceanServer Iver2 UUVs belonging to NRL Code 7160 and NRL's Bluefin-21 vehicle Reliant) by developing an implementation that interfaces with the open-source MOOS-IvP autonomy framework. Preliminary in-water testing took place at NRL's Chesapeake Bay Detachment and in Boston Harbor.

In the process of developing our UUV agent framework and performing initial in-water tests during FY19, we created several scenarios that may be run in simulation or in at-sea trials. These scenarios provided a limited battery of tests that may be used to evaluate the agent's performance in different missions and under differing environmental conditions. To evaluate the performance of different approaches (e.g., constraint solvers or route-planning algorithms) to the problem of navigation-dependent goal selection, we developed testing tools that automatically simulated many test scenarios and evaluated the final results in terms of time to completion and maximal risk incurred by the UUV. We also developed several new test scenarios intended to explore the space of low-probability, high-risk challenges the UUV may encounter during execution (e.g., large numbers of threats, or threats located in unlikely but challenging configurations).

The MOOS-IvP autonomy framework employed on our test vehicles provides a vehicle-simulation application, but it simulates only one vehicle at a time and requires an external PID controller application. Thus, a separate instance of both applications is required for every simulated vehicle, and configuring simulated missions quickly becomes infeasible when any reasonable number of simulated vessels are needed. To address this, in FY20 we implemented a new MOOS application that simulates a large number of vessels from one instance. Simulated vessels can be configured to perform different subjective behaviors as observed during in-water experiments in Boston Harbor (i.e., following a constant heading for shipping and ferry traffic, loitering at random waypoints for pleasure and fishing boats). We also performed miscellaneous technical improvements throughout the year (e.g., gathering all 5514 MOOS applications and libraries under one repository, improving code reuse across MOOS applications, automating build scripts, incorporating well-tested geometric, string, and utility code from the widely-used and open-source Boost libraries).

In FY21, we implemented a new bearing-only sonar sensor simulator and multi-vessel tracker, based on but simplifying the approach employed by Calkins *et al*. (2021). This tracker simulates passive detection capabilities present on some NRL vehicles but is capable of running faster than realtime in simulation, improving our ability to run tests and experiments in scenarios of realistic duration (i.e., multiple hours). We integrated this tracker with existing NRL tooling for MOOS-IvP simulations, including our goal-reasoning software and behaviors developed by NRL Code 7130. We also implemented a simple sidescan-sonar simulator, permitting detection of simulated mine-like objects and the ability to incorporate them into the goal reasoner's world model. Further, we implemented an ability to execute plans generated by the goal reasoner either without replanning (i.e., executing a static plan) or with replanning only (i.e., no changes in goals) to serve as a baseline for evaluation.

Also in FY21, we refactored our mission configuration to support multi-vehicle teams by launching multiple MOOS communities. This configuration is capable of launching a simulated mission with global objects (e.g., other vessels, mine-like objects) simulated in a central community that also handles inter-vehicle communications, or launching multiple real-world vehicles in a team configuration with a shared seed for randomized optimizers and inter-vehicle communications supported by onboard micromodems.

We are currently running experiments involving simulated missions at our usual test location in Boston Harbor and are preparing to submit a paper (to the ACS-21 Workshop on Goal Reasoning) on our multi-vehicle, precedence-constrained model for GR on UUVs.

5. Conclusions

We have developed and implemented a new model of goal reasoning (GR) that incorporates more information about the requirements of plan execution for embodied agents, naturally supports reasoning about temporally extended and complex navigation, selects goal sequences for entire missions rather than reactive, short-term needs, and can support multi-vehicle teams. We have implemented UUV-specific task descriptions, mission scenarios, and integration with an autonomy framework. We have developed efficient solvers for navigation cost estimation and a broad base of optimizers to evaluate for the precedence-constrained, time-dependent orienteering problem representing the goal selection problem. We are currently evaluating our work in simulated scenarios and have tested our code onboard UUVs in at-sea trials, and will report our results in an appropriate venue.

Acknowledgements

The authors gratefully acknowledge the collaboration of James McMahon (NRL Code 7130), Jennifer Nelson (SSEP student employee), Gregory Hyde (NREIP student intern), Ari Goodman (NAVAIR Lakehurst), Jon Decker (NRL Code 5581), Phil Baldoni (NRL Code 7130), Luke Calkins (NRL Code 7130), and Artur Wolek (ASEE postdoctoral fellow) during the course of this project (FY18-FY21)

References and Publications

References

1. Balakirsky, S., Schlenoff, C., Fiorini, S.R., Redfield, S., Barreto, M., Nakawala, H., Carbonera, J.L., Soldatova, L., Bermejo-Alonso, J., Maikore, F., Goncalves, P.J.S., De Momi, E., Sampath Kumar, V.R., & Haidegger, T. (2017). Towards a Robot Task Ontology Standard. In *Proceedings of the ASME 12th International Manufacturing Science and Engineering Conference.* Los Angeles, CA: American Society of Mechanical Engineers.

2. Calkins, L., Baldoni, P., McMahon, J., Wilhelmi, C., & Zavlanos, M. M. (2021). Bearing-Only Active Sensing Under Merged Measurements. *IEEE Robotics and Automation Letters*, 6(3), 4544-4551.

3. Chabini, I., & Lan, S. (2002). Adaptations of the A* algorithm for the computation of fastest paths in deterministic discrete-time dynamic networks. *IEEE Transactions on Intelligent Transportation Systems*, *3*(1), 60-74.

4. Dorigo, M., & Gambardella, L.M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation, *1*(1), 53-66.

5. Kavraki, L.E., Svestka, P., Latombe, J.-C., & Overmars, M.H. (1996). Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*,*12*(4), pp. 566-580.

6. Klenk, M., Molineaux, M., & Aha, D.W. (2013). Goal-driven autonomy for responding to unexpected events in strategy simulations. *Computational Intelligence*, *29*(2), 187-206.

7. Verbeeck, C., Soerensen, K., Aghezzaf, E.-H., & Vansteenwegen, P. (2014). A fast solution method for the time-dependent orienteering problem. *European Journal of Operational Research*, *236*(2), 419-432.

8. Wilson, M., Molineaux, M., & Aha, D.W. (2013). Domain-independent heuristics for goal formulation. In *Proceedings of the Twenty-Sixth Florida Artificial Intelligence Research Society Conference*. St. Pete Beach, FL: AAAI Press.

9. Wilson, M.A., McMahon, J., Wolek, A., Aha, D.W., & Houston, B.H. (2018). Goal reasoning for autonomous underwater vehicles: Responding to unexpected agents. *AI Communications*, *31*(2), 151-166.10. Yu, V., Jewpanya, P., Lin, S.-W., Perwira Redi, A.A.N. (2019). Team orienteering problem with time windows and time-dependent scores. *Computers & Industrial Engineering*, *127*, 213-224.

Publications

10. Wilson, M.A., Nelson, J., Hyde, G., and Aha, D.W. (2019). A Time-Dependent Utility Model for Long-Term Spatiotemporal Goal Reasoning. NCARAI Technical Note.

11. Wilson, M.A., McMahon, J., and Aha, D.W. (2019). Goal Sequence Selection for Underwater Vehicles with MOOS-IvP. MOOS Development and Applications Working Group Meeting.

Presentations

We have presented aspects of this work to several NRL visitors and others within the research and defense communities, including: ONR's Science of Autonomy review meeting; ad the AI Strategic Challenge TTCP; Denise Crimmins (SSTM, USW Rapid Prototyping; DASN RDT&E and NUWC Newport) during a visit to NRL; LCDR Kyle Woerner (DARPA TTO Program Officer); VADM Merz (Deputy CNO), VADM Radakin (UK Second Sea Lord), RADML David Hahn (CNR); a group from NDU; the Senate Arms Service Commiteee (SASC); and many other VIP visitors to NRL.

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.