Online Plan Modification for Autonomous Underwater Vehicle Missions: Dissertation Abstract

Catherine Harris and Richard Dearden
School of Computer Science
University of Birmingham
Birmingham, United Kingdom B15 2TT
[C.A.Harris.1, R.W.Dearden]@cs.bham.ac.uk

In recent years, the use of autonomous underwater vehicles (AUVs) has become increasingly popular for a wide variety of applications, ranging from geological surveys and environmental monitoring to pipeline and hull inspection. This is in part due to their ability to act independently from the support ship and crew, allowing multiple tasks and experiments to be undertaken simultaneously, saving valuable ship time. As the cost of deploying a vehicle and the risk of loss or damage are often high, AUV missions typically consist of simple pre-scripted behaviours (Pebody 2007), such as 'lawnmower' surveys (where the vehicle surveys a section of the sea floor in a series of overlapping parallel passes). By analysing data from the vehicle between deployments, operators could tailor subsequent missions to further investigate areas of interest. The recent development of long-range vehicles, such as Autosub Long-Range (Furlong et al. 2012), designed to perform missions lasting many days or weeks, has the potential to revolutionise the collection of oceanographic data. However, as the progress of long-duration missions cannot be periodically reviewed by human operators, long-range vehicles will require an increased level of autonomy to fully capitalise on their increased capabilities. AUV missions can be thought of as oversubscribed planning problems, where finite amounts of battery power and data storage space limit the number and duration of data-collection tasks achievable by the vehicle. Designed to safeguard the vehicle whilst operating in inherently uncertain environments, pre-scripted missions are inevitably over-conservative, reserving a significant proportion of battery as a contingency should usage be higher than expected. This means that in the average case, the vehicle is not being used to its full potential.

We can formally represent the AUV planning domain as a Markov Decision Process (MDP) (Russell and Norvig 2003), in which at any given time the vehicle is considered to be in one of many states. The vehicle may transition between states by performing various actions, such as moving to a new location, collecting a data set, travelling to and from the surface, and transmitting data sets back to a support vessel or laboratory. Upon transitioning to certain states, such as following the successful transmission of a data set, the vehicle receives a reward. Data sets are only of value to scientists if they are successfully recovered from the vehicle, either by being present on the hard-drive at the point of vehicle collection or by being transmitted by the vehicle mid-mission. Without this stipulation, the potential value of the data and the cost of losing it (such as through the corruption of onboard storage or the total loss of the vehicle) would not be represented within the problem. The vehicle has two resources, battery power and data storage space, which are each represented as a continuous state variable. All actions will consume battery power; however, the exact amount consumed is not known in advance but modelled with a probability distribution specific to each action. For example, there is much less uncertainty in the power used while surfacing than in collecting data from a survey area. Similarly, the amount of memory a data-set will consume once compressed is also uncertain.

For large-scale planning domains with continuous state variables and a high degree of resource uncertainty it is infeasible to compute the optimal course of action for every possible eventuality in advance. This prevents the use of an MDP solver to generate a full policy for the AUV planning domain. Instead, we have focussed on the development of an online planning algorithm, which revises and refines the plan during execution in response to fluctuating vehicle resources and changing mission goals. An online approach allows us to reason about goals and reduce the uncertainty of future states using observations made during the execution of the plan. If a state is observed to have more resources available than previously expected, an additional goal may be included to increase the scientific data return and consequently the reward available to the vehicle. Conversely, if resource usage has been higher than expected, actions and goals may be removed to avert plan and mission failure.

We hypothesized that by allowing the vehicle to change its planned behaviour during execution in response to unexpected situations, the expected reward achieved by the vehicle for this 'switching-plan' would increase above and beyond that of any single straight-line plan. Following on from work by Bresina et al. (2002), in which the authors computed the optimal value function associated with the Mars rover planning problem as a function of continuous resource variables, we performed an investigation to compare the optimal value functions for a hand-constructed subset of both straight-line and branching plans within the AUV domain. For any combination of initial resources the optimal value function represents the expected reward obtained by following a particular plan. The optimal value function of the switching plan was found to completely dominate that of the straight-line plans, exceeding the expected reward for all resource combinations (Harris and Dearden 2012). This showed that changing the plan during execution is at least as good as the best straight-line plan. For some combinations of resources, the switching plan significantly outperformed the best straight-line plan as it was able to either utilise resources to achieve greater rewards or avert failure by branching to a simpler plan, depending on the observed resource usage prior to the branch point.

As stated earlier, the large state space in the AUV planning domain means that the use of an MDP solver would be computationally infeasible. Conversely, although the use of classical planning algorithms such as GraphPlan (Blum and Furst 1997) would be feasible, they do not represent uncertain resource usage. Instead, we theorised that a better solution might be to quickly generate candidate plans using a modified classical planner, before evaluating the quality of each plan using the MDP model. As the benefits of planning are often hard to quantify in advance, it is important that any online plan generation does not require the vehicle to invest significant time and resources mid-mission. Rather than using the classical planner to generate a new full plan during execution in response to an unexpected situation, we are researching the development of an algorithm which makes online modifications to an existing plan by interleaving existing actions with those of rapidly generated sub-plans, each solving a single additional goal. The initial plan would be generated offline (prior to the deployment of the vehicle) using GraphPlan. To minimise online computation, we hope to re-use information from GraphPlan's planning graph representation, such as causal relationships and possible orderings, to efficiently combine the initial plan with new sub-plans during execution. We are currently investigating using techniques from the field of partial-order planning, such as threat-resolution (Penberthy and Weld 1992), in combination with information contained within the planning graph representation to generate a partially-ordered plan, listing all valid orderings of the actions within both plans. Full orderings can then be extracted and evaluated using the MDP model to select an ordering which maximises reward without violating resource constraints. We are also investigating how to optimise the quality of the eventual solution; for example, by detecting additional causal links which might allow the removal of a redundant action. However, the computational complexity of the algorithm remains an important consideration as online planning algorithms ultimately have to successfully balance the inevitable trade-off between quality and run-time complexity.

References

Pebody, M. 2007. The Contribution of Scripted Command Sequences and Low Level Control Behaviours to Autonomous Underwater Vehicle Control Systems and Their Impact on Reliability and Mission Success. In *Proceedings of OCEANS – Europe*, 1–5.

Furlong, M. E.; Paxton, D.; Stevenson, P.; Pebody, M.; McPhail, S. D.; and Perrett, J. 2012. Autosub Long Range: A long range deep diving AUV for ocean monitoring. In *Proceedings of IEEE/OES Autonomous Underwater Vehicles (AUV)*, Southampton, UK.

Russell, S., and Norvig, P. 2003. Making Complex Decisions. In *Artificial Intelligence, A Modern Approach*. 2nd ed. Prentice Hall.

Bresina, J.; Dearden, R.; Meuleau, N.; Ramakrishnan, S.; Smith, D.; and Washington, R. 2002. Planning Under Continuous Time and Resource Uncertainty: A Challenge for AI. In *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*. 77-84. Morgan Kaufmann.

Harris, C., and Dearden, R. 2012. Contingency Planning for Long-Duration AUV Missions. In *Proceedings of IEEE/OES Autonomous Underwater Vehicles (AUV)*, Southampton, UK.

Blum, A., and Furst, M. 1997. Fast Planning Through Planning Graph Analysis. *Artificial Intelligence* 90:281-300.

Penberthy, J. S., and Weld, D. 1992. UCPOP: A Sound, Complete, Partial-Order Planner for ADL. In *Proceedings of the Third International Conference on Knowledge Representation and Reasoning (KR-92)*. Cambridge, MA.