Adaptive UGV Navigation using Advocates and Critics for Tactical Behaviors

Talib Hussain, Gordon Vidaver, David Montana
BBN Technologies
Ph: 617-873-8000, Fax: 617-873-4328
{thussain, gvidaver, dmontana} @ bbn.com, www.bbn.com

ABSTRACT

We present a system for highly adaptive navigation and control of unmanned ground vehicles that produces explicit, appropriate and understandable tactical behaviors. The system is based upon the use of an evolutionary algorithm to explore tactical alternatives for continual path and action planning to achieve mission objectives within a dynamic environment that may contain unknown terrain characteristics and one or more active enemy ground units with unknown tactical objectives and capabilities. Our system evaluates the relative merits of a number of "small changes" to a plan (made with multiple different planning rationales) as the tactical situation changes over time. We present a summary of simulation-based work we have conducted for the Army and discuss efforts to incorporate learning capabilities into the system.

1. INTRODUCTION

Today's unmanned ground vehicles demonstrate mobility systems that are effective at navigating through known types of terrain to visit pre-determined locations while avoiding known classes of obstacles that are typically fixed or slow-moving. When they encounter unexpected situations (i.e., for which they have not been programmed), they often fail and/or require human intervention to overcome the problem. Military needs will require these systems to improve along two key dimensions. On the one hand, UGVs will need to be able to (a) navigate effectively through complex terrain that has not been encountered before, and (b) accommodate uncertain, incomplete, dynamic and/or incorrect knowledge. On the other hand, UGVs will need to be able to (c) develop and execute plans to carry out higher-level mission goals (such as reconnaissance, surveillance, and target acquisition) and (d) respond rapidly and effectively to threats posed by mobile and intelligent enemy units. [1]

BBN has developed the Advocates and Critics for Tactical Behaviors (ACTB) approach, a navigation and control technique based upon the incorporation of tactical goals and behaviors into an evolutionary-based planning framework [2]. Our goal is to explore methods for enabling UGVs that perform in meaningful tactical ways while accomplishing their mission goals under a variety of known and novel situations. In particular, ACTB focuses upon "adaptive planning", in which the system uses fixed and random planning capabilities to dynamically replan as time passes and/or as the tactical situation changes (or "Changing the decisions we make"). In an Army Research Lab (ARL) sponsored project, the performance of an ACTB system was tested in simulation and shown to produce adaptive responses to dynamic tactical situations [3].

We present the details of the ACTB approach, summarize the results of the performance tests and discuss avenues for extending ACTB with the capability to learn from experience to improve its planning capabilities over time (or "Changing the way we make decisions").

2. UGV NAVIGATION

In response to an environment and a set of mission requirements that are dynamically changing, a UGV navigation system must perform replanning of both the reactive (e.g., avoid an obstacle or turn to run away from an enemy) and deliberative (e.g., modify entire path to circumvent an enemy and remain hidden en route to next mission goal) varieties. While there has been some previous work done on deliberative planning for robots and UGVs, there has been much more practical work in the reactive planning area [4-8].

Our approach is to view the deliberative planning problem as an optimization problem to determine an operation plan for one or more UGVs that achieves multiple mission goals while satisfying multiple tactical criteria as best as possible based upon the most recent knowledge available. We view the reactive planning problem as multiple sub-problems, some of which are amenable to autonomic processes that require very little information or adaptation (and hence may be programmed in detail), and some of which require intelligent exploration of alternative actions. In particular, we apply an evolutionary algorithm to perform the deliberative and intelligent reactive planning based on explicit tactical behaviors, and provide a mechanism for triggering autonomic processes when appropriate.

Prior work has explored the use of behavior-based navigation for reactive planning [4,5] and the use of evolutionary algorithms for robot path planning and vehicle routing [7,8]. The work closest to our approach is the Distributed Architecture for Mobile Navigation (DAMN) [4], which provides a sophisticated reactive control component. DAMN contains behaviors, each of which represent some higher-level navigation goals, such as 'road following', 'seeking the next navigation goal', 'obstacle avoidance', 'avoid hazards'. Each behavior provides a vote on the next direction to take, and a command arbiter decides upon the best direction, which is then taken by the UGV. While most of the behaviors are reactive, there is one behavioral input from a deliberative planner called the global navigator [5]. The global navigator is capable of determining a full path to a goal position using a D* (dynamic A*) search algorithm. However, this approach does not incorporate as many criteria and as much information at the deliberative planning level as we believe are necessary for robust performance under varying conditions.

3. ACTB ARCHITECTURE

The Advocates and Critics for Tactical Behavior architecture is illustrated in Figure 1. The figure shows planning and control for a simulated UGV that may be controlled via five basic commands: *move* to a given geographical location, *stop* movement, *look* in a given direction by rotating the camera, *aim* the gun in a given direction, and *fire* the gun. In general, planning may be performed for one or more vehicles [2]. ACTB incorporates three main components: a continual planning cycle based on *tactical advocates and critics*, a mechanism based on *attitudes* for influencing the nature of the planning performed, and *autonomic* control mechanisms.

The plan for each vehicle is a description for proposed UGV actions into the future, and specifies a sequence of move and/or "stop and look" commands. The plan does not specify any gun aiming or firing commands. During execution, a single plan is initially selected. A contiguous sequence of waypoints is passed directly to the simulator as a sequence of corresponding moves. A stop and look command is associated with the immediately preceding waypoint – when the vehicle arrives at that waypoint, it will stop and slew its camera to the indicated angle. When the camera is fully turned, the next action in the plan will be executed. If that is another stop and look command, the vehicle will remain at the same waypoint until that look is completed. Otherwise the next appropriate sequence of move commands will be issued.

The primary component of ACTB is a continual tactical planning cycle that is influenced by changes in the state of knowledge about the world. Initially, the planning process starts with a set of new plans that are generated randomly based on the initial world conditions. The new plans are evaluated by a set of *tactical critics* to determine their effectiveness. Critics examine each plan and evaluate its entire length based on that critic's tactical criteria (e.g., traversability, safety, mission success, exposure risk or situational awareness). The weighted sum of the critic evaluations produces a single score for the plan.

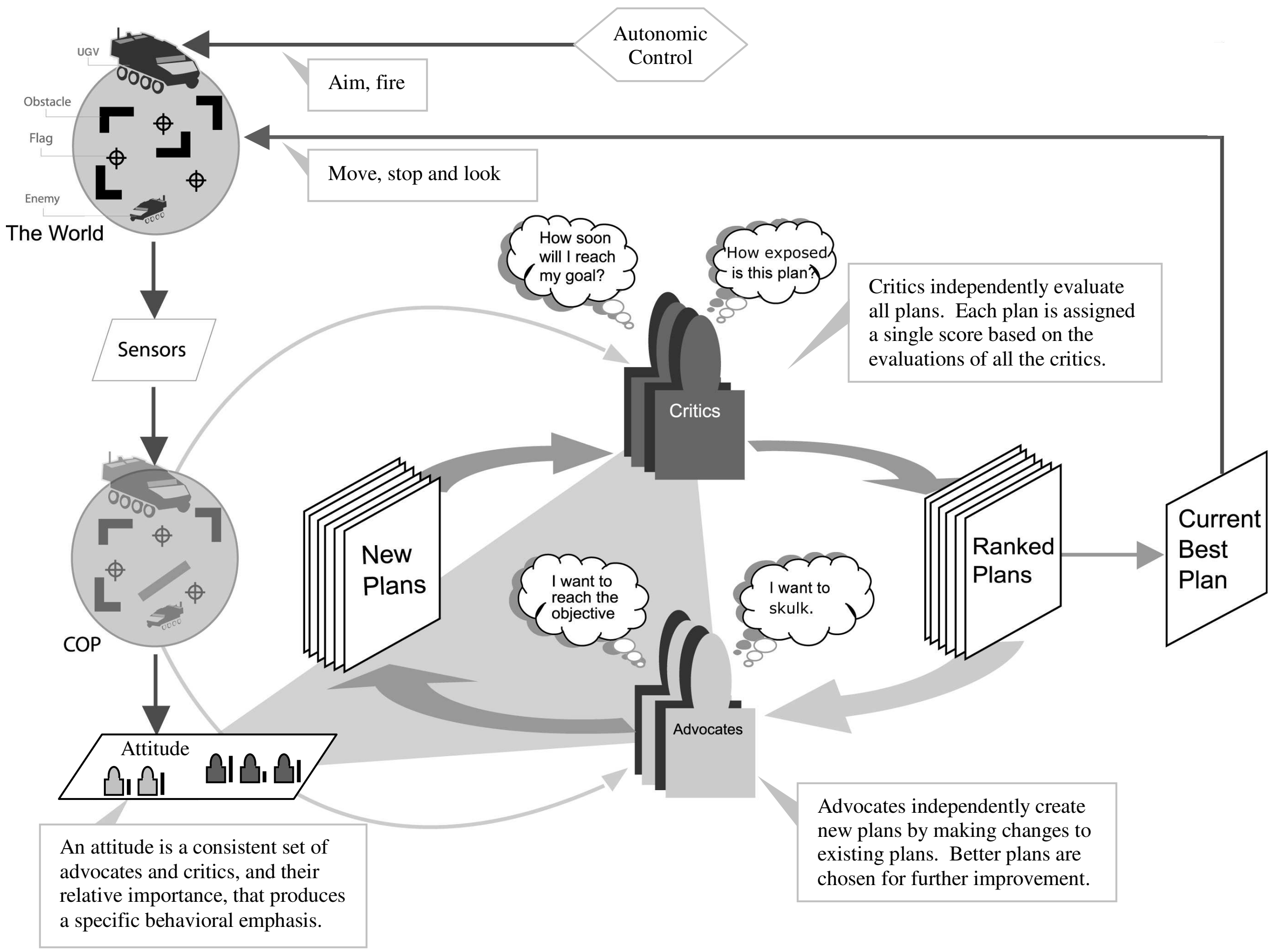


Figure 1: ACTB Architecture

Once all the new plans are scored, they are ranked with respect to each other and to plans from previous cycles. The highest ranked plan is always selected as the current plan for execution, and some of the lowest ranked plans are eliminated. As the planning cycle continues, some of the ranked plans are modified by a set of *tactical advocates*. Advocates examine an existing plan and, when appropriate, create one or more new plans that better meet that advocate's goal (e.g., get away from the enemy faster, surveil before entering new territory, maintain cover near walls, or head towards the mission objectives). Most advocates make a small, incremental change (e.g., by modifying a few move, stop or look commands). The change may occur anywhere in the plan.

Generally, planning continues indefinitely as the new plans generated by the advocates are in turn evaluated by the critics, and so on. During any given cycle, plans are adapted based on the most recent knowledge in the COP. Deliberative planning is achieved through the use of advocates that make changes to any part of a plan. Intelligent reactive planning is achieved through the use of advocates that focus upon the immediate path and actions of the vehicle. Adaptive planning is achieved through the continual planning cycle since new knowledge and events will be seamlessly incorporated. For example, a newly discovered obstacle may make a critic evaluate a previously good plan as a bad one since it tries to make the UGV go through that obstacle.

While the continual planning cycle enables ACTB to be adaptive to changes in the world, it is not always sufficient to enable ACTB to provide robust, timely response under tactically significant changes. While a large number of advocates and critics may enable a very rich search for new plans, that search will tend to be slow (i.e., trying to solve one difficult optimization problem). To speed up the search while maintaining effectiveness, the ACTB architecture incorporates high-level strategic ”rules of engagement” that influence the manner in which planning is performed by changing its *attitude*. An attitude reflects a bias in the tactical behaviors used in the planning process. Specifically, an attitude is a consistent set of advocates and critics, each with certain probabilities of selection and weights, respectively (i.e., ACTB uses attitudes to solve several simpler optimization problems).

The strategic rules define the tactical conditions under which the system will transition from one attitude to another. For example, during a mission it may be appropriate under different conditions for a UGV to become more cautious (e.g., it doesn't know where the enemy is so try to stay out of sight) or more aggressive (e.g., the enemy is in the way of achieving the mission, so close and attack).

The ACTB planning cycle is rapid, especially with the use of attitudes. However, it is not appropriate for all vehicle control. Within our architecture, *autonomic behaviors* may be defined that are triggered based on certain tactical situations. The behaviors must operate independently of any planned actions. For instance, as illustrated in Figure 1, autonomic behaviors for gun control may be defined that operate independently of any planned move or stop-and-look actions.

4. TACTICAL PLANNING AND CONTROL

In an ARL-sponsored project, we developed a tactical navigation system based upon ACTB for controlling a single UGV. The system was tested in simulation. The simulator, developed by General Dynamics Robotics Systems (GDRS), defined a binary terrain and simulated UGVs. Figure 2 illustrates one terrain used in the test trials, where white represents clear areas and black represents obstacles. The mission objectives of the vehicle were to visit up to three target "flag" locations, and either kill any enemies discovered or survive for a fixed period of time (typically 10 to 20 minutes). The tactical navigation system was required to be robust to changes in the terrain, the number of enemies (one or two) and the location of the flags.

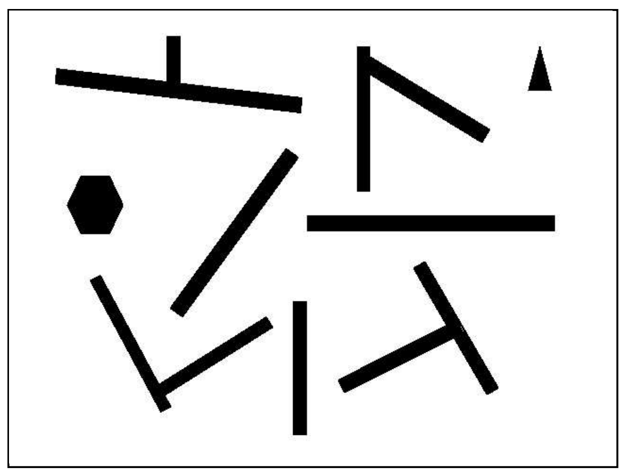


Figure 2: Simulated terrain

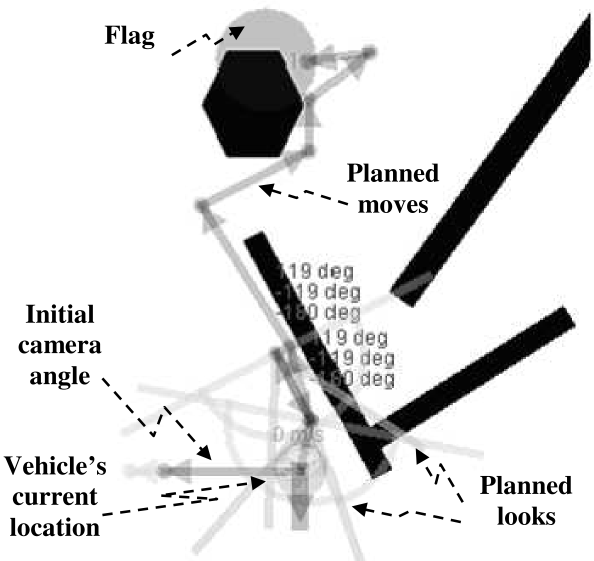
The simulated UGV platform had a gun turret and a camera, and could be controlled via five basic commands: *move* to a given geographical location, *stop* movement, *look* in a given direction by rotating the camera, *aim* the gun in a given direction, and *fire* the gun. Further, upon request, the simulator provided information about the world. The ACTB approach was used to develop a system with nine tactical critics, a number of tactical advocates, four attitudes and three autonomic behaviors.

4.1. Critics

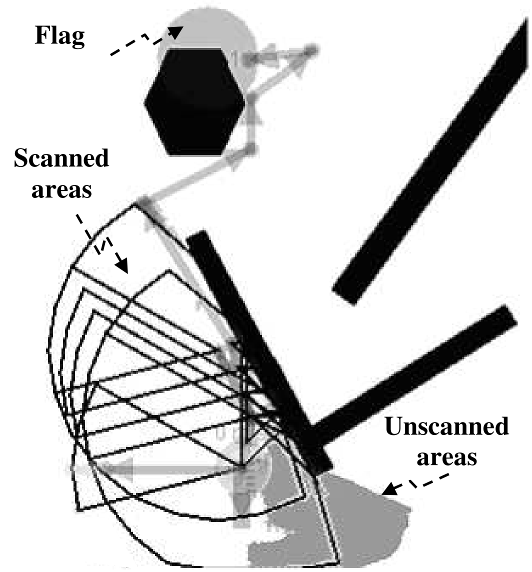
The critics used were:

- *Traversability*, which penalizes travel through obstacles over all portions of the path.
- *Safety*, which evaluates how long a plan puts the UGV within firing range of the enemy.
- *Proximity*, which rewards plans that quickly move the UGV out of the proximity of a currently observed enemy.
- *Time-to-Mission*, which rewards paths that take less time to visit the flags.
- *Mission-success*, which penalizes plans based on the number of flags that are not visited.
- *Exposure*, which penalizes a path based upon the sum of the exposure risk over all point in the plan. Exposure risk of a point is defined as the size of the area co-visible with that point.
- *Lingering*, which rewards plans that move the UGV to areas that have not been visited recently. Without such a critic, the best option might be to remain in one location since it is safe, scanned and enemy-free.
- *Complexity*, which penalizes plans with many short segments.
- *Awareness*, which rewards plans that make the UGV look towards areas not recently scanned prior to entering them (e.g., look around a corner first). Figure 3 illustrates the

awareness computation made for a given plan. The plan (Figure 3a) is comprised of both planned moves (dark gray arrows) and looks (light gray angles). The plan defines a path towards the known flag (gray circle). Figure 3b illustrates the viewable area that is scanned before entering (empty cones) and the viewable area that is not scanned (dark gray areas).



(a)



(b)

Figure 3: Illustration of the computation of the awareness critic

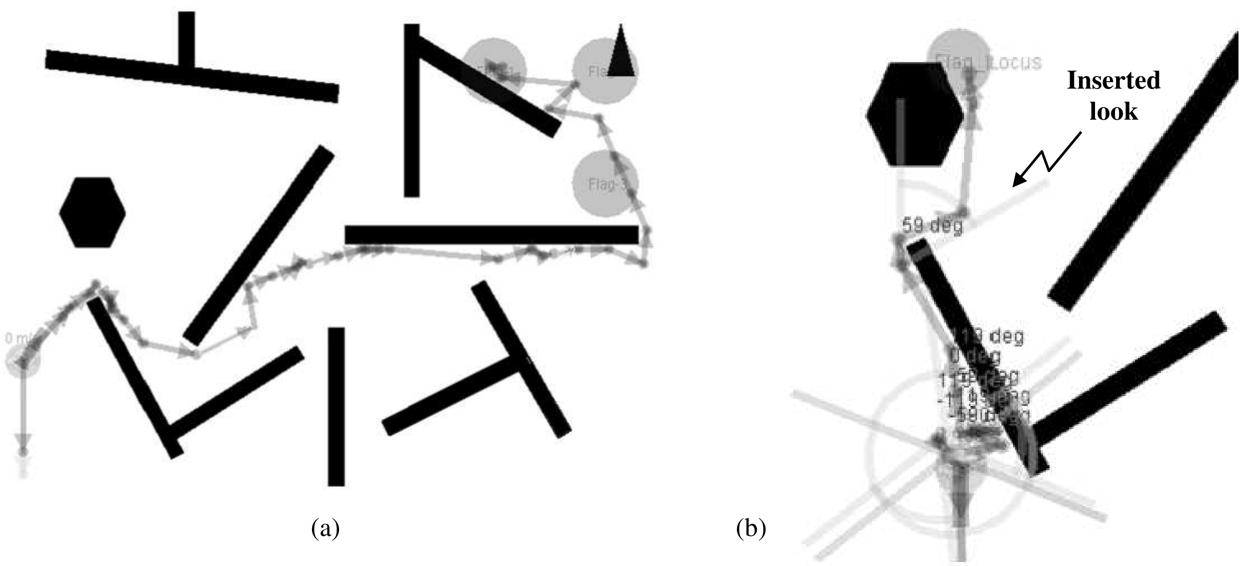
4.2. Advocates

The advocates used included:

- *Insert-Flag*, which exploits knowledge of the mission to improve plans that do not visit all remaining flags by insert appropriate waypoints to visit them into new plans.
- *Skulk Gradient*, which exploits knowledge about the exposure risk to identify a path segment that is highly exposed and move its endpoints in the direction of lower exposure based on the exposure gradient at those points. Figure 4a illustrates a plan in which the

skulk gradient advocate has played a significant role - the path tends to stay near cover (i.e., walls) and avoid open areas.

- *Peek*, which exploits knowledge about the terrain and exposure risk to identify a path segment where there is a large transition from low risk to high risk of exposure and insert a “stop and look” command at that transition. This will generally lead to an improved score on the awareness critic. Figure 4b illustrates the result of applying the peek advocate – a look is inserted on the path just prior to end of an obstacle, enabling the UGV to see around the “corner” before reaching it.



(a) (b)

Figure 4: Illustration of the effects of skulk gradient and peek advocates

- *Go-Around*, which replaces a section of the path that crosses an obstacle or passes close to an enemy with a new sequence of waypoints that form a circular path (either clockwise or counter-clockwise) around the obstacle/enemy.
- *Wall Trace*, which replaces a section of the path with a sequence of waypoints that moves the UGV along the contour of a nearby obstacle.

- To ensure a rich evolutionary search, several random advocates were also used, including: *insert-waypoint* (which inserts a new waypoint at a random position in the plan), *insert-look* (which inserts a stop-and-look action at a random position in the plan), *remove-section* (which randomly selects two actions in the plan and removes them and all actions between them), and *plan-crossover* (which performs variable-length one-point crossover between two plans).

4.3. Attitudes

The attitudes used were:

- *Mission-Oriented*, which places emphasis on achieving the mission objectives quickly by encouraging travel to new regions and allowing some risk.
- *Cautious*, which places emphasis on stealthy exploration.
- *Flee*, which places emphasis on getting away from the enemy rather than accomplishing the mission or planning looks.
- *Fight*, which places emphasis on achieving the mission while maintaining maneuverability, with safety as a low factor.

Critic	Attitude			
	Cautious	***Flee***	***Fight***	***Mission-Oriented***
Traversability				
Safety			Low	Low
Proximity			Low	Low
Exposure		n/a	n/a	
Awareness		n/a	n/a	Low
Time-To-Flag		Low	Low	High
Mission-Success		Low	Low	
Lingering	n/a	n/a	n/a	High
Complexity		Low		High

Table 1: Critics used in each attitude and their weights relative to default.

The relative difference in the critic weights between attitudes is summarized in Table 1. A critic could be used with the default weight (medium gray), a low weight (light gray), a high weight (dark gray) or not at all (n/a).

The attitudes were used to influence the ACTB planning cycle based on the tactical situation. Figure 5 illustrates the tactical situations that triggered transitions in the attitude. The strategy represented by Figure 5 reflects our expectation that the enemies would be highly aggressive and likely to guard the flags.

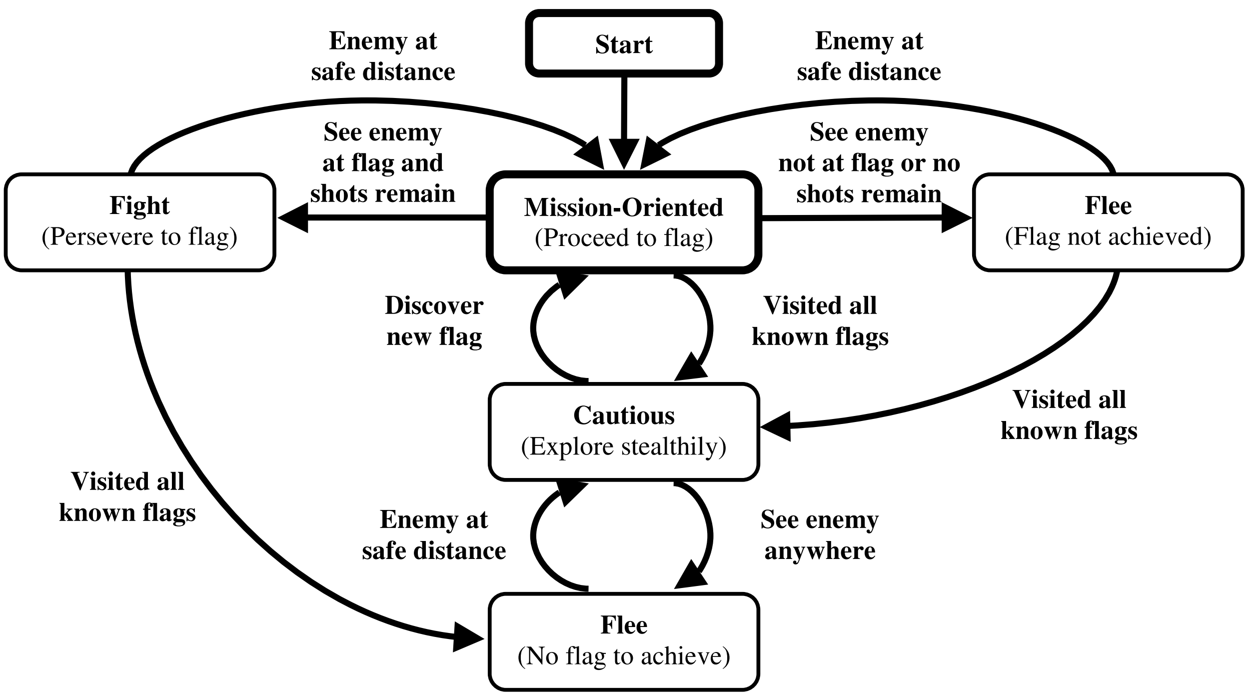


Figure 5: Rules of engagement select the attitude based on the current tactical situation

5. RESULTS

During several days in September 2004, our tactical behavior system was tested directly against the systems of two other competing project members. Two sets of trials were conducted. In the first set, 60 one-on-one trials were conducted, with each team's vehicle competing 20 times against each opponent. In the second set, 132 trials were conducted, with each system participating one-on-one against each competitor 36 times and against both competitors 24 times. Different trials used different terrain variations, such as the addition of obstacles, the removal of entire obstacles, and the partial removal of some obstacles (e.g., a "path" through a "wall").

The results of the trials are shown in Tables 2-4. Three types of aggregate data are summarized for each team: the number of flags visited over all trials, the number of enemy kills made, and the number of times the vehicle survived until end of the trial.

	# Flags Visited (out of 120)	# Kills (out of 40)	# Survivals (out of 40)
BBN	78	13	30
Competitor 1	27	7	26
Competitor 2	62	13	23

Table 2: Aggregate results of first set of testing trials

	# Flags Visited (out of 480)	# Kills (out of 120)	# Survivals (out of 96)
BBN	182	13	48
Competitor 1	32	5	71
Competitor 2	193	65	86

Table 3: Aggregate results of second set of testing trials

	# Flags Visited (out of 180)	# Kills (out of 36)	# Survivals (out of 36)
BBN	61	5	10
Competitor 2	56	26	31

Table 4: Head-to-head testing results between BBN and Competitor 2 in second set of trials

In the first set of trials, our system performed better than both competitors. We made a large number of kills, captured a large percentage of the flags (65%) and had a higher survival rate than our opponents (75%). In the second set of trials, our system performed very poorly relative to Competitor 2 in terms of our ability to target and shoot the enemy. However, we outperformed both opponents in terms of our ability to capture flags (e.g., in head-to-head trials against Competitor 2, we were able to capture more flags despite being killed 72% of the time).

In all trials, our system showed a rich set of timely responses and performed in noticeably tactical ways. It did not always make the same decisions when faced with a similar situation, but still demonstrated tactical behaviors appropriate to those situations. The performance was consistent with our high-level rules of engagement, and our autonomic behaviors and planned behaviors supported each other. For example, our UGV would often flee from an enemy, but continue to target and fire at it; a number of kills were obtained in this way.

6. CONCLUSIONS

The development of effective tactical behaviors for unmanned ground vehicles is a critical area of research. Our Advocates and Critics for Tactical Behaviors (ACTB) approach provides a rich planning and control architecture that enables us to incorporate a variety of competing tactical behaviors into a single planning system and explicitly define rules of engagement to achieve robust performance under different tactical situations. Our approach to adaptive navigation has been shown to be effective against two other state-of-the-art approaches.

A key avenue of our current research is the incorporation of learning capabilities into ACTB. Learning to handle novel tactical situations requires the ability to not only identify that situation, but also to determine what the best type of tactical response is for that situation and how to execute that response. For example, what does it mean to learn:

1) *Is this a dangerous situation?*
2) *What should I be looking for?*
3) *What will my adversary do next?*
4) *How should I avoid what I need to avoid?*
5) *What is the best strategy to achieve my goal*
6) *What is the best way to attack?*

Each key component of the original ACTB system may potentially be augmented with learning capabilities. As illustrated in Table 5, each level has different issues and results in different benefits from the use of learning.

ACTB Level	Key Benefit of Learning	Key Issue
Autonomic	Improved accuracy of immediate-term actions	Tightly scoped domain
Advocate	Improved tactical responses through experience	Frequent change, Explosion of components
Critic	Improved tactical reasoning in novel situations	Maintaining commander's intent
Attitude	Improved ability to handle multiple tactical situations	Impact on system stability

Table 5: Summary of the benefits and challenges for different levels of learning

The partitioning of the solution into advocates, critics and attitudes in ACTB has enabled us to quickly design effective solutions by decomposing the problem into manageable "bites" – it is easy to be an advocate for a specific behavior, it is easy to be a critic about what went wrong, and it is easy to associate broad priorities with certain key tactical situations. We believe that the same "bite-size" approach will be critical for achieving effective learning of tactical behaviors.[*]

REFERENCES

1. National Research Council Staff, *Technology Development for Army Unmanned Ground Vehicles*. National Academies Press, Washington, D.C., 2002
2. T. Hussain, D. Montana and G. Vidaver, "Evolution-Based Deliberative Planning for Cooperating Unmanned Ground Vehicles in a Dynamic Environment," *Genetic and Evolutionary Computation - GECCO 2004, Part II. Lecture Notes in Computer Science 3103*. K. Deb et al (Eds.), p. 1017-1029, Springer-Verlag, Berlin, 2004.
3. T. Hussain, G. Vidaver and J. Berliner, "Advocates and critics for tactical behaviors in UGV navigation," *Unmanned Ground Vehicles VII Conference*, 2005.
4. J. Rosenblatt, "The distributed architecture for mobile navigation," *Journal of Experimental and Theoretical Artificial Intelligence*, *9*, p. 339-360, 1997.
5. B. Brumitt and A. Stentz,, "Dynamic mission planning for multiple mobile robots," *IEEE International Conference on Robotics and Automation*, p. 2396-2401, 1996.
6. B. Gerkey and M. Mataric, "Sold!: Auction methods for multi-robot coordination," *IEEE Transactions on Robotics and Automation*, *18*, p. 758-786, 2002.
7. I. Ashiru, C. Czarnecki and T. Routen, "Characteristics of a genetic based approach to path planning for mobile robots," *Network and Computer Applications*, *19*, p. 149-169, 1996.
8. B. Baker and M. Ayechew, "A genetic algorithm for the vehicle routing problem," *Computers and Operations Research*, *30*, p. 787-800, 2003.

[*] Some of the work presented in this paper was conducted under subcontract to General Dynamics Robotics Systems (GDRS) under contract #DAAD19-01-C-0066. We would like to acknowledge the contributions of Jeffrey Berliner, Richard Estrada, Richard Lazarus and Stephen Milligan to the ideas presented in this paper.