AIAA Infotech@Aerospace 2007 Conference and Exhibit
7 - 10 May 2007, Rohnert Park, California

AIAA 2007-2973

Preference-Based Trajectory Generation

Jamie A. Lennon*
Naval Research Laboratory, Washington, DC, 20735

Ella M. Atkins†
University of Michigan, Ann Arbor, MI, 48109

Numerous techniques exist to optimize aircraft and spacecraft trajectories over cost functions that include terms such as fuel, time, and separation from obstacles. Relative weighting factors can dramatically alter solution characteristics, and engineers often must manually adjust either cost weights or the trajectory itself to obtain desirable solutions. Further, when humans and robots work together, or when humans task robots, they may express their performance expectations in a "fuzzy" natural language fashion, or else as an uncertain range of more or less acceptable values. This work describes a software architecture which accepts both fuzzy linguistic and hard numeric constraints on trajectory performance and, using a trajectory generator provided by the user, automatically constructs trajectories to meet these specifications as closely as possible. The system respects hard constraints imposed by system dynamics or by the user, and will not let the user's preferences interfere with the system and user needs. The architecture's evaluation agent translates these requirements into cost functional weights expected to produce the desired motion characteristics. The quality of the resulting full-state trajectory is then evaluated based on a set of computed trajectory features compared to the specified constraints. If constraints are not met, the cost functional weights are adjusted according to precomputed heuristic equations. Heuristics are not generated in an ad hoc fashion, but are instead the result of a systematic testing of the simulated system under a range of simple conditions. The system is tested in a 2DOF linear and a 6DOF nonlinear domain with a variety of constraints and in the presence of obstacles. Results show that the system consistently meets all hard numeric constraints placed on the trajectory. Desired characteristics are often attainable or else, in those cases where they are discounted in favor of the hard constraints, failed by small margins. Results are discussed as a function of obstacles and of constraints.

Nomenclature

F_i = Trajectory feature vector for planning state s_i
$J(x,u,t,\{O\})$ = Domain-dependent multi-objective cost function with weights W_i
J_i = Integrated cost over trajectory $(\underline{t}_i, \underline{x}_i, \underline{u}_i)$
L_i = Feature vector limits (constraints) for planning state s_i (L_0 = initial/default limit set)
$\{O\}$ = Set of k obstacles $\{o_1,o_2,\ldots,o_k\}$
p_0 = Trajectory planning problem $<bc, W_0, L_0>$ with boundary conditions $bc = <t_0, \underline{x}_0, \underline{x}_f>$
$\underline{t}_i$ = Vector of trajectory time points $\{t_1, \ldots, t_m\}$ for planning state s_i
$\underline{u}_i$ = Control actuation vector over the m trajectory time points for planning state s_i
W_i = Cost function weighting factor vector used in planning state s_i
$\underline{x}_i$ = Position/velocity state vector over the m trajectory time points for planning state s_i
X = Solution $<J_n, L_n, \underline{t}_n, \underline{x}_n, \underline{u}_n>$ returned for planning problem p_0

*Aerospace Engineer, Naval Research Laboratory, Code 8231, 4555 Overlook Ave. SW, Washington, DC 20735, Member.
†Associate Professor, Aerospace Engineering, 3009 FXB Building, 1320 Beal Avenue, University of Michigan, Ann Arbor, MI 48109, Associate Fellow.

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

I. Introduction

INTELLIGENT robotic systems will play an important role in future space and planetary surface operations. Whether exploring on their own or accompanying and supporting human pioneers, they will need the capability to reason, plan ahead, and make decisions based on goals, the environment, and the desires of human or robotic teammates. Embodied robots must also translate mission goals into appropriate physical responses.

Balancing competing costs, while satisfying certain hard constraints, is an important component of "appropriateness." In space exploration problems, fuel and power conservation are dominant issues, whether the agent under discussion has a limited tank of fuel for positioning or, despite recharging capability, has a limited power budget constrained by battery weight. Timeliness is also a concern, as many scientists may wish to use a vehicle's capabilities for a variety of projects before its lifespan ends. Preserving vehicle health is another priority, and all of this must be done while respecting the dynamical constraints of the vehicle, and the dynamical properties of its environment.

Human users, whether on-site astronauts or ground-based controllers, might express their desired balance of these costs linguistically: "get this done quickly," "be very careful," "maximize your range." When communicating with other humans, this works very well – better, often, than attempting to fix numerical values to desired characteristics. As humans, we are very adept at giving and understanding these linguistic expressions even if we cannot precisely describe them in objective terms.

We can, for example, easily identify "aggressive driving" when we see it on the roads. An aggressive driver's behavior is marked by high traveling speeds, frequent lane changes, sudden accelerations, and the maintenance of slim safety margins to other vehicles. What is a "high traveling speed?" Even once the context is fixed (e.g., interstate vs. in-town), the linguistic term has some fuzziness to it. Certainly, it implies a speed higher than the legal, posted speed limit. It probably means a speed higher than the average speed of the other drivers. But is someone driving 5 m.p.h. faster than road speed an aggressive driver? And on the other extreme, is there a speed so high that we can say that we have gone past "aggressive driving" and are into a region of "reckless driving?" When, exactly, is that line crossed? The answers to these questions are easy for humans to intuit, but difficult to formalize.

These concerns follow us into the realm of trajectory optimization. Robot trajectories need not be optimal. In some domains, we may accept satisficing trajectories that are adequate but not optimal. But in the space domain in particular, we will always be concerned with conserving fuel and power. Even if we want an "aggressive, fast" trajectory, we will still want it to be the most fuel- efficient aggressive trajectory. We are concerned with fuel even if the result is not the fuel-optimal solution.

Trajectory optimization, at its most general, will have multiple objectives and constraints. Multiple objectives in particular give rise to multiple possible solutions. Consider a two objective case: we desire to save both time and fuel. These objectives compete. The most fuel-efficient trajectory is rarely the most time-efficient, and visa versa. There might be several solutions that take the same minimum time, and we would be interested in the most fuel-efficient one. Or, we might examine all of the minimum fuel solutions and pick the fastest of those. Or, we might want an intermediate solution– one that is neither the fastest nor the most fuel-efficient, but balances the two objectives. We may have an initial estimate of relative importance in saving time or fuel. How can we communicate such preferences to the numeric trajectory solver so that it will find the "right" optimal?

In addition to these fuzzy preference ideas, we may also have trajectory constraints, typically imposed to meet dynamic performance limits and to ensure safety (collision avoidance). Beyond that, we might further impose limits on either control inputs (e.g., actuator travel limits or thruster saturation) and on system state. Some of these limits could be "soft," like a posted speed limit that can be exceeded given appropriate circumstances (e.g., to avoid an erratic driver). Other limits are "hard", such as acceleration limits imposed to avoid pilot/astronaut blackout.

There are many approaches to solving the constrained multi-objective optimization problem,[1] including evolutionary algorithms, mixed integer linear programming, and optimal control theory. To a greater or lesser extent, each handles soft and hard constraints. All, however, include an iterative refinement loop to find solutions that best match user preferences: that is, to find the "optimal optimal" solution. As a common thread throughout the multi-objective optimization literature, it is tacitly understood that a human user interacts directly with the optimization algorithm, injecting preference information to focus the optimization on areas of interest to the user.

This work seeks to take the human user out of that loop as much as possible. Before optimization ever begins, classes of motion are typified with linguistic expressions: aggressive, curious, careful. Fuzzy logic[2] is an appropriate tool for approaching the problem of translating natural language utterances into numeric terms.[3] Words are correlated to fuzzy state values the system believes best represent user preferences for the resulting trajectory: the "numbers" we need to take the human supervisor out of the optimization loop. Some iteration may be required to ensure that the user's expectation matches the fuzzy definition of the linguistic expressions. However, once that

process is complete, the autonomous "supervisor" can balance optimization objectives without user input. This work considers a planetary rover and an Earth-observing satellite as motivating examples. The planetary rover example is a very simplified case with two degrees of freedom and linear dynamics that provided initial insight into the algorithms required to inject preferences into the trajectory generation process. The satellite case has more realistic six degree-of-freedom (DOF) dynamics. The satellite's hypothetical job is to provide imaging to support ground-based decision making. It can execute fuel burns to change its orbit in response to user demands on the ground. These user demands may have varying levels of urgency. Some may be matters of curiosity, but no urgency at all, in which case the satellite is free to execute maneuvers whenever it is most fuel-efficient to do so. It may also need to maneuver around other space objects (perhaps other satellites that require observation).

This paper proposes an architecture to compute preference-optimal trajectories. A cognitively-inspired expert system moderates the trajectory generation and optimization process. At initialization, a solution technique is selected based on problem characteristics. If an initial trajectory estimate is required for the solution technique, one is generated, again with consideration for the problem characteristics. Expressed user preferences are transformed via fuzzy methods into an initial set of weights or other parameters, and the selected solution technique is run.

The expert system also considers the results generated by the chosen algorithm. Often in these problems, one or more user-defined constraints or preferences will not be met after the first iteration. Making changes to the weight vector or to other parameters may solve the problem; so may a different initial trajectory estimate or the use of a different solution technique (e.g., if the problem will not solve using the first technique). Given its knowledge base and the current history of repair attempts for this problem, the expert system continues to search for an appropriate trajectory. Below, related literature is introduced to motivate the posed problem and algorithm choices made. Next, the preference-optimal trajectory generator is introduced at the system architecture and algorithm level. Planar and 6-DOF systems are introduced, with results used to evaluate the fuzzy logic and iterative weight adjustment strategies over a series of simple and complex domain examples.

II. Related Work

Incorporating user preferences into a trajectory generation feedback loop requires assimilation of representations and techniques across multiple disciplines. Below, we motivate our work through connection to the literature in natural language, motion planning, multi-objective optimization, and fuzzy set theory. Emphasis is placed on defining the technological needs for insertion of preferences into trajectory generation processes and identification of gaps in existing techniques.

A. Natural Language

Verbal or written instructions are one possible mode of interaction between a human and a robotic vehicle.[4] With respect to motion words, descriptive verbs such as 'swagger,' 'slink,' 'slide,' and 'sway' effectively convey the nature of motion, but the specific choice of verb is open to interpretation.[5] Translated to our problem, the speaker has numerous choices to describe how a route from a location *A* to *B* is traversed. What verbs the speaker selects indicate to some extent the "manner of motion" required for the robot.

There is no appreciable literature dealing with the transformation of verbs to numbers. There is, however, a literature on assigning numeric values to spatial expressions such as 'near' or 'in front of'.[6,7,8] Researchers use, among other techniques, a potential field (first developed for robotic path planning[9]) as a membership function in the fuzzy set theory sense;, fuzzy terms like "crisp" and "scruffy" appear frequently. Essentially, one point or line is selected (by the researchers) to represent the ideal of "near," "along," or "in front of" some object in the space. This becomes the minimum for the potential field, which can be visualized as a bowl. The minimum of the potential field 'bowl' is located at the ideal value set for "near," "along," etc. This paper extends this idea to motion words. First, we define a "state feature space" composed of state features such as average forward velocity and maximum acceleration. A collection of points in state feature space is taken as the ideal representation of a verb or adverb/verb pair, like "jog" or "move stealthily." Fuzzy membership functions are then defined around these areas, so that similar but not identical kinds of motion can still be included in these classifications. This provides flexibility when trying to satisfy the multiple constraints and maximize the combination of objectives such terms imply while maintaining the user's preference for motion type.

B. Path and Kinodynamic Planning

Path planning focuses on finding a path through free space from an initial to a final location.[10] The robot's dynamics are not generally considered for holonomic robots. For nonholonomic robots, dynamic constraints that directly affect path, such as a turning radius, are used to reject infeasible paths. When following the path, the robot

is typically pre-programmed with a simple trapezoidal velocity profile that ramps up to a constant velocity then decelerates to the final zero-velocity state. For slow, wheeled robots, especially those in a laboratory or office environment, this model usually suffices to move the robot around. *Behavior-based* motion control[11] is the next step, in which environmental cues trigger pre-programmed responses. The resulting actions can be sophisticated or even appear emotional.[12] Although research begun to make parameters that define these behaviors adaptive to environmental stimuli, they are still reactive, favoring simplicity and real-time response over optimization.[13,14]

While most behavior-based protocols can be represented with a Markov decision process (MDP), another branch of research looks at hybrid dynamical systems.[15,16] In such systems, discrete events trigger shifts between different continuous dynamics. In the cited work, a simulated mouse agent adjusts its trajectory in response to the environment by changing weights. The weights, however, correspond to repelling and attracting potential functions for local navigation, not useful for global or multi-objective optimization. "Programming by reward" is a technique that elicits different dynamic behaviors from a system.[17] Like our research, it uses preference information to create these differences. Unlike our research, it injects the preference information into machine learning algorithms for the development of motion behaviors. These behaviors can form an optimal policy, given preferences. However, the "interiors" of the behaviors are still black boxes. The number of lane changes in the authors' driving example can be optimized for a safe driver and for a reckless driver, but the dynamics of that lane change are unexamined. Our multi-objective trajectory optimization application requires control of the low-level inputs that result in the desired behaviors, rather than assembling pre-typed behaviors into a policy.

Trajectory, or kinodynamic, planning incorporates velocity (dynamics) into planning processes. Methods based on velocity obstacles[18,19] are conceptually similar to path planning "roadmap" methods, with infeasible velocities modeled as velocity space obstacles. Spline methods[20,21] decouple path and velocity planning; once a clear path through space is found, interpolating splines find smooth trajectories along them. Randomized kinodynamic planning[22] explores the state space in a random fashion, working forward from the start state and backward from the goal state until the search trees meet. All these methods search for dynamically feasible trajectories but have a single parameter or fixed function to define 'optimality'. Velocity obstacle and spline methods have been used to generate time optimal trajectories[23,24] and randomized methods are often used as starting points for linear programming methods to generate fuel-optimal trajectories.[25]

C. Multi-objective Optimization

Trajectory planning is a multi-objective problem with constraints, typically including fuel use, traversal time, and obstacle clearance as often competing parameters to balance. While different approaches have been developed, they have one common feature: a parameter set that can be adjusted to reflect user priorities. Genetic and evolutionary algorithms (GAs and EAs) have become popular search and optimization tools.26 A population of potential solutions is generated and encoded then evaluated, with the most promising solutions modified or maintained over multiple generations that ultimately yield one or more solutions. Since there is rarely a single point where all objectives are simultaneously maximized or minimized, evolutionary multi-objective optimization (EMO) frequently makes use of *Pareto optimality27*28 in which a set of nondominated solution vectors are identified. These Pareto-optimal solutions are defined as the *Pareto front*. Recent GA/EA research has concentrated on ways to encourage population diversity to enable exploration of the entire Pareto front, and the addition of elitism, which stores nondominated individuals so that they will not be lost. Most work to-date has studied two or three objectives, for which Pareto front graphs are examined to select a solution from the nondominated set. Open questions remain for "many-objective optimization". It has been shown that, as the number of objectives increases, Pareto dominance becomes nearly useless in ranking individuals.[29] Also, for deployed applications it is infeasible to evaluate a comprehensive nondominated solution set, thus incorporating user preference into EAs is an emerging area of study.

There are two additional considerations to note when deliberating the use of EAs for trajectory optimization. First, EAs do not explicitly calculate gradients along the solution set. A good fitness function and the judicious use of crossover and mutation ensures that the solutions will tend to follow the gradient down to the minimum, but this is accomplished by selecting ever more-fit individuals, not by taking advantage of trends. In some cases, this is a strength – EAs are robust to discontinuities in the solution space, particularly prevalent at constraint boundaries. However, when gradients are available, their use typically speeds convergence. Second, EAs are an unconstrained optimization technique. Sometimes, constraints can be recast as objectives. An assortment of penalty functions can be invoked, penalizing the fitness of solutions that do not meet the constraints, although this increases the risk of degenerating into a random walk. Another approach iteratively shrinks the search space to focus on zones where the constraints are met.[30] These approaches work, some faster than others, and most have a set of parameters (such as the "rate of shrink"[30]) that must be set then tuned by the user.

More deterministic methods have also been employed to develop a Pareto front. Isoperformance[31,32] sets a required performance level, indicated by a fixed value for the cost function. This can be the single output of a complex system, such as the displacement of space telescope subjected to disturbance forces. Design variable sets that give the desired performance level are recorded. From that set, a nondominated front is further selected. A user would then select a single solution from among the nondominated solutions. Adaptive weighted sum methods[33] can be used when the cost function is a weighted sum of terms. Traditional Pareto front exploration would sample weights at constant fixed intervals, potentially missing important front features. The adaptive weighted sum approach begins with a constant-interval weight mesh, then refines the mesh in areas with large gaps in cost. Inequality constraints are also added to restrict calculations to areas where the Pareto front is believed to exist.

These techniques are promising for multi-objective design optimization (MDO), or more generally in any application where a user would want to obtain an entire nondominated set of solutions for consideration. If we were interested in only planning a route between two specific points, the time required to form the Pareto front of trajectories using one of these techniques might be worthwhile. However, our interest is in calculating many trajectories in similar but not identical environments. Since the trajectory generation process is itself computationally intensive, we avoid computing a full Pareto front for any one set of boundary conditions.

To this end, Mixed Integer Linear Programming (MILP) has become increasingly popular as a relatively fast way to generate and optimize trajectories.[34] Derived from the operations research field, recent increases in computing speed have allowed MILP to be considered for real-time applications. Equality and inequality constraints can all be handled robustly. MILP can approximate non-convex and logical constraints. Obstacle avoidance is achieved by placing inequality constraints directly on the path space, forcing the trajectory outside the region of the obstacles; penalty functions are not used. The cost functional could in principle include many terms, although applications typically focus on either fuel or time optimal trajectories. MILP has been applied to Aerospace applications such as spacecraft rendezvous and multi-satellite reconfiguration.[35,36] However, as its name suggests, it is suitable only for problems with linear constraints, including dynamic constraints. For space- based planning, researchers typically employ simplified models such as the linear Hill's/Clohessy-Wiltshire equations[35] and gravity-free (flat space) dynamic system formulations.[36] MILP does not handle nonlinear constraints, except by linearizing them. This is appropriate for some domains, but not for all. For example, trajectory planning for a satellite (or satellite formation) over a highly-elliptical orbit or over more complex gravity fields is not amenable to linearization. More fundamentally, as with EA, MILP does not support user preferences except through cost functions and constraints. As will be shown, although we adopt an optimal control trajectory planner to handle nonlinear dynamical systems, our architecture could also adopt MILP as its trajectory planner upon which our preference-based deliberation structure is placed.

D. Optimal Control Trajectory Planning

Optimal control methods[37,38] have been extensively used to solve complex trajectory planning problems, and are applicable to linear or nonlinear dynamical systems. The calculus of variations is used to frame the problem as a system of differential equations subject to conditions imposed at the initial and final time. Generally, a cost functional is of the form:

$$J = \int_{t_0}^{t_f} g(\boldsymbol{x}(t), \boldsymbol{x}'(t), t)\mathrm{d}t \tag{1}$$

where $\boldsymbol{x}(t)$ is the state vector and $\boldsymbol{x}'(t)$ is its derivative. A variation in the functional, δJ, can be defined for small changes of $g(\boldsymbol{x}(t),\boldsymbol{x}'(t),t)$. If a relative minimum for J exists, it is necessary that δJ be zero at that point, yielding the Euler equation. The problem is then to find an admissible input (or control) vector $\boldsymbol{u}^*(t)$ that causes a system described by the differential equations in Eq. (2) to follow an admissible trajectory $\boldsymbol{x}^*(t)$ that minimizes Eq. (3) cost.

$$\boldsymbol{x}'(t) = a(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \tag{2}$$

$$J(\boldsymbol{u}) = \int_{t_0}^{t_f} g(\boldsymbol{x}(t), \boldsymbol{u}(t), t)\mathrm{d}t \tag{3}$$

At all points along an admissible trajectory, Eq. (2) holds and can be rewritten:

$$a(\boldsymbol{x}(t), \boldsymbol{u}(t), t) - \boldsymbol{x}'(t) = 0$$

This dynamic constraint set with Lagrange multipliers $\boldsymbol{\lambda}$ forms an augmented cost functional:

$$\begin{aligned} J_a(\boldsymbol{u}) = \int_{t_0}^{t_f} & g(\boldsymbol{x}(t), \boldsymbol{u}(t), t) \\ & + \boldsymbol{\lambda}^T \left[a(\boldsymbol{x}(t), \boldsymbol{u}(t), t) - \boldsymbol{x}'(t) \right] \mathrm{d}t \end{aligned} \tag{4}$$

The extremals of the functional are where δJ_a is zero. Finding δJ_a and setting it to zero results in three necessary equations. They are most commonly expressed in terms of the Hamiltonian, defined as:

$$H(\mathbf{x}(t),\mathbf{u}(t),\boldsymbol{\lambda}(t),t) = g(\mathbf{x}(t),\mathbf{u}(t),t) + \boldsymbol{\lambda}^T[a(\mathbf{x}(t),\mathbf{u}(t),t)] \tag{5}$$

The necessary conditions are then:

$$\begin{aligned} \boldsymbol{x}'^*(t) &= \frac{\partial H}{\partial \boldsymbol{\lambda}}(\boldsymbol{x}^*(t),\boldsymbol{u}^*(t),\boldsymbol{\lambda}^*(t),t) \\ \boldsymbol{\lambda}'^*(t) &= -\frac{\partial H}{\partial x}(\boldsymbol{x}^*(t),\boldsymbol{u}^*(t),\boldsymbol{\lambda}^*(t),t) \\ 0 &= \frac{\partial H}{\partial \boldsymbol{u}}(\boldsymbol{x}^*(t),\boldsymbol{u}^*(t),\boldsymbol{\lambda}^*(t),t) \end{aligned} \tag{6}$$

for all $t \in [t_0, t_f]$. For a fixed final time and a fixed final state, we have boundary conditions $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, $\boldsymbol{x}(t_f) = \boldsymbol{x}_f$ which enables constants of integration to be determined. Solving the system of equations returns the function (trajectory) that minimizes the cost functional.

If the final time and final state are free, a new boundary condition called the transversality condition is produced:

$$g_a(\boldsymbol{x}(t_f), \boldsymbol{u}^*(t_f), \boldsymbol{\lambda}(t_f), t_f)\delta t_f - \boldsymbol{\lambda}^T(t_f)\delta \boldsymbol{x}_f = 0 \tag{7}$$

When $\boldsymbol{x}_f$ is fixed, as in this work, $\delta \boldsymbol{x}_f = 0$, so $g_a(\boldsymbol{x}(t_f), \boldsymbol{u}^*(t_f), \boldsymbol{\lambda}(t_f), t_f) = 0$. Except for certain special cases, there is no way to analytically solve the optimal control problem. A variety of numeric methods have been employed, including shooting methods[39] and collocation. The shooting method uses initial value problem (IVP) solutions as a starting point to "shoot" towards the solution of the optimal control boundary value problem (BVP). A stable BVP may require the integration of unstable IVPs (ones highly sensitive to changes in boundary values), a drawback that led to the development of the collocation approach. In collocation, the actual solution to differential equations (6) is approximated over a mesh, defined by "knot points." The approximation is made to satisfy the boundary constraints at t_0 and t_f, and further to satisfy Eq. (6) at each knot point and at the midpoint of each interval between them. An initial guess for the solution must be provided; the solution technique will alter the current solution estimate to bring its residual (a measure of error) to within acceptable bounds.

There are many ways to solve a collocation problem. Solution methods fall in general into two classes: direct and indirect. Direct methods[40] model the approximate solution as composed of basis functions; the solution is improved by altering a vector containing the coefficients for the basis functions. This allows vector optimization algorithms such as sequential quadratic programming, Newton-Gauss, or Levenberg-Marquardt to be applied.[41] Direct methods are considered faster and more robust than the indirect methods. Indirect methods[42] link knot points with continuous approximating functions (e.g., splines) over each subinterval. The coefficients of each of these functions must then be solved. This makes indirect methods more computationally intensive than direct methods. Their advantage is in flexibility; the basis functions in the direct methods must be chosen such that every function could be a feasible trajectory. The indirect method has no such constraint.

The optimal control problem's solution is governed by a single cost functional. Multiple objectives can only be optimized via an aggregation method. Since some constraints are likely to be non-convex, this means that certain solutions along the Pareto optimal front may be missed. Typically, a sufficient number of other solutions that also satisfy the user's preferences also exist where they can be detected. Optimal controls problems can incorporate constraints and discontinuities. Equality constraints on the state (such as satisfying system dynamics) are adjoined to the cost functional via Lagrangian multipliers, as discussed above. Constraints on the control inputs can be handled via Pontryagin's Minimum Principle and the resulting switching curves. Inequality constraints can be handled by the introduction of a function of a dummy variable, x_{n+1}, whose derivative is defined as:

$$x'_{n+1}(t) \equiv [f_1(\boldsymbol{x}(t),t)]^2 \Uparrow(-f_1) + [f_2(\boldsymbol{x}(t),t)]^2 \Uparrow(-f_2) + \ldots + [f_l(\boldsymbol{x}(t),t)]^2 \Uparrow(-f_l) \tag{8}$$

where $\Uparrow(-f_i)$ is a unit Heavyside function defined by:

$$\Uparrow(-f_i) = \begin{cases} 0, & f_i(\boldsymbol{x}(t),t) \geq 0 \\ 1, & f_i(\boldsymbol{x}(t),t) < 0 \end{cases} \tag{9}$$

for $i = 1, 2, \ldots, l$ (where $l \leq m$, the size of the control vector). x_{n+1} can then be defined as:

$$x_{n+1}(t) = \int_{t_0}^{t_f} \dot{x}_{n+1}(t)dt + x_{n+1}(t_0) \tag{10}$$

We require boundary conditions $x_{n+1}(t_0) = 0$ and $x_{n+1}(t_f) = 0$. Since the derivative is never less than zero, $x_{n+1}(t)$ must be zero for all t. This is a constraint of the form $f(\boldsymbol{x}(t),t) = 0$ treated by the method of Lagrange multipliers.

The switching curves and Heavyside function are discontinuous, making them problematic for many numeric solvers. They can be approximated by a series of increasingly steep polynomials. However, the unchanging nature of x_{n+1} presents a further problem. Collocation solvers require gradient information to reduce the error between the current approximate solution and the true solution. Since x_{n+1} is identically zero for the entire trajectory, it provides no gradient information. In our particular case, the collocation solver required a Jacobian matrix which, when the Heaviside approximation was added, contained a full row of zeroes and so would not solve. Adjusting the Heaviside approximation further to provide gradient information in the allowable solution region amounted to instituting a penalty function, which is another way to treat state inequality constraints.

Penalty functions are often used in path and trajectory planning for obstacle avoidance. Often cubic in form, these penalty functions are centered over an obstacle and monotonically decrease as they move away from its center. Typically, they go to zero at some influence limit away from the obstacle, but this is not required. For our work, penalty functions assume a fixed value at an obstacle's center, at the edge of the object, and at a fixed distance from the edge of the object. These constant values are then connected through smooth cubic functions, the coefficients of which can be varied to achieve these conditions for obstacles of different sizes. With an optimal control approach, the value of the penalty function is added to the cost functional. As cost is minimized, the trajectory will move away from obstacles. However, if other costs are sufficiently great, it may be numerically less expensive to accept the penalty – which means planning a path through the obstacle. Penalty functions do not offer guarantees on constraint satisfaction, which means solutions generated via optimal control methods must be validated for obstacle clearance.

E. Fuzzy Set Theory

Above we reviewed available techniques for trajectory optimization with multiple objectives and nonlinear dynamics. The primary element missing from these methods is the ability to inject user preferences initially and at intermediate points during trajectory optimization processes. Because a human user will typically express preferences through natural language adverb/verb expressions as discussed above, we require a connection between these words and the mathematical cost and constraint formulations common to all trajectory optimization protocols. We employ fuzzy set theory to perform this translation. The main idea behind fuzzy set theory is that a member of a set may belong only partly to that set.[2] Classically, individuals either are or are not contained in a set. An individual may be 50% hot and 50% not hot, or 30% hot and 50% warm, for example. Complements, like "hot" and "not hot" must sum to 100% but non-complementary attributes may not. For example, the vertical line in Figure 1 indicates generic feature value F is about 45% "low," about 60% "medium," and 0% "high." The triangles in Figure 1 are membership functions. They correlate "crisp" numeric values, as measured in the real world, to these fuzzy levels. A fuzzy rules set then acts on these "fuzzified" inputs. For example, "If air temperature is LOW, turn heater fan to "HIGH" and "If air temperature is MEDIUM, turn heater fan to LOW." The fans speeds will have similar fuzzy membership function correlating speeds like "HIGH" and "LOW" to revolutions per minute. These outputs are scaled by the membership function of the inputs.

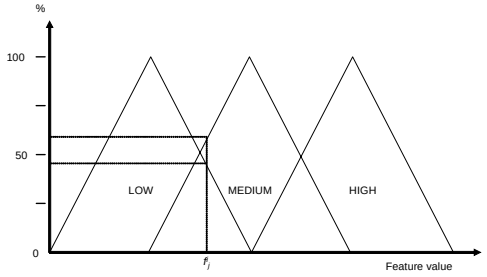

Figure 1: A fuzzy membership function

Natural language, while a desirable input modality, is inherently ambiguous. From interpreting sounds into words to parsing the words into sentences to interpreting the possible shades of meaning of a sentence, there are ambiguities. Classical mathematics does not manage ambiguities well. Fuzzy techniques, on the other hand, deal with them substantially better.[3] Fuzzy optimization applies fuzzy set theory to optimization problems. Fuzzy techniques are not themselves used to solve the problem, but are rather applied to candidate solutions to rank them. They are often used in conjunction with EAs, where the EAs generate candidate solutions and the fuzzy methods rank them before selection and breeding occur. Recent work[6] has shown that an expanded and fuzzified notion of Pareto dominance seems to perform more in accord with common sense than strict Pareto dominance, and should not have the same problem as Pareto dominance (e.g., that all solutions become equally good) as the number of objective functions increases to infinity. We combine fuzzy sets with optimal control trajectory planning to gain the benefits of gradient-based numerical optimization while employing fuzzification to translate "common sense" user preferences into numerical objective weights and constraints.

III. Architecture

Figure 2 shows an outline of the agent's processes[43]. At the center sits the evaluation model, overseeing all activities. The human user interacts with this module, monitoring events rather than directly participating in trajectory generation processes. The evaluation module, *EVAL*, accepts a planning problem, $\mathbf{P_0}$, which can be posed by the user or by any suitable high-level planner that builds task-level actions to achieve its goals, some of which may require vehicle motions.

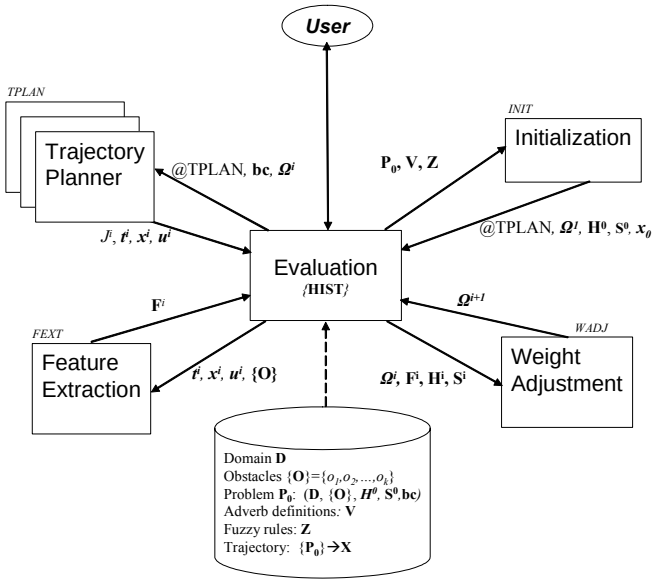

Figure 1: System architecture

A trajectory planning problem $\mathbf{P_0}$ is defined as $\{\mathbf{D}, \{\mathbf{O}\}, \mathbf{H_0}, \mathbf{S_0}, \mathbf{bc}\}$. Domain **D** describes system dynamics and the parameterized cost functional *J* to be minimized. **{O}** represents the set of obstacles in the environment. $\mathbf{H^0}$ describes the hard constraints (limits on state space values) to be met, whereas $\mathbf{S_0}$ is a set of soft constraints that indicate user preference, but are ultimately flexible. $\mathbf{H^0}$ are numeric; $\mathbf{S^0}$ may be numeric or fuzzy linguistic terms. Fuzzy terms must eventually be converted into soft numeric limits; **L**, the set of all limits, includes $\mathbf{H^0}$ and the extended $\mathbf{S^0}$. Members of **L** may be upper limits, lower limits, or range limits (when we want the state feature to be within an upper and a lower limit). The boundary conditions $\mathbf{bc} = \{t_0, \boldsymbol{x_0}, \boldsymbol{x_f}\}$ are split and can include all of the usual optimal controls cases (e.g., fixed or free final time or state, final state constrained to a fixed or moving surface). The goal is to return feasible and optimal solution $\mathbf{X} = \{J^n, \boldsymbol{L^n}, \boldsymbol{t^n}, \boldsymbol{x^n}, \boldsymbol{u^n}\}$, where J^n and $\boldsymbol{L^n}$ summarize solution cost and the feature limits/constraints, respectively, of the *n*th iteration and the set $\{\boldsymbol{t^n}, \boldsymbol{x^n}, \boldsymbol{u^n}\}$ specifies the full-state trajectory (i.e., time sequence $\boldsymbol{t_n}$, position/velocity vector sequence $\boldsymbol{x_n}$, control inputs $\boldsymbol{u_n}$) to be executed. This goal is achieved through intelligent selection of a trajectory planning function and selection and adjustment of a weight vector $\boldsymbol{\Omega^i}$ that influences the relative importance of terms in the cost functional *J*. *EVAL* incrementally builds a history of activities {**HIST**} = {$\mathbf{HIST^1}$, $\mathbf{HIST^2}$, …} with $\mathbf{HIST^i}$ including a record of the function used by *TPLAN*, the initial solution estimate, and the weight vector $\boldsymbol{\Omega^i}$ used for the *i*th iteration. *EVAL* can then use {**HIST**} to identify which weight adjustment strategies it has already employed, to avoid infinite loops. Figure 2 shows the possible paths through the architecture. Initialization combines the hard and soft constraints and selects initial weights which are expected to meet them[44]. The system loops until all $\mathbf{H^0}$ are met; if $\mathbf{S^0}$ are not met, a second

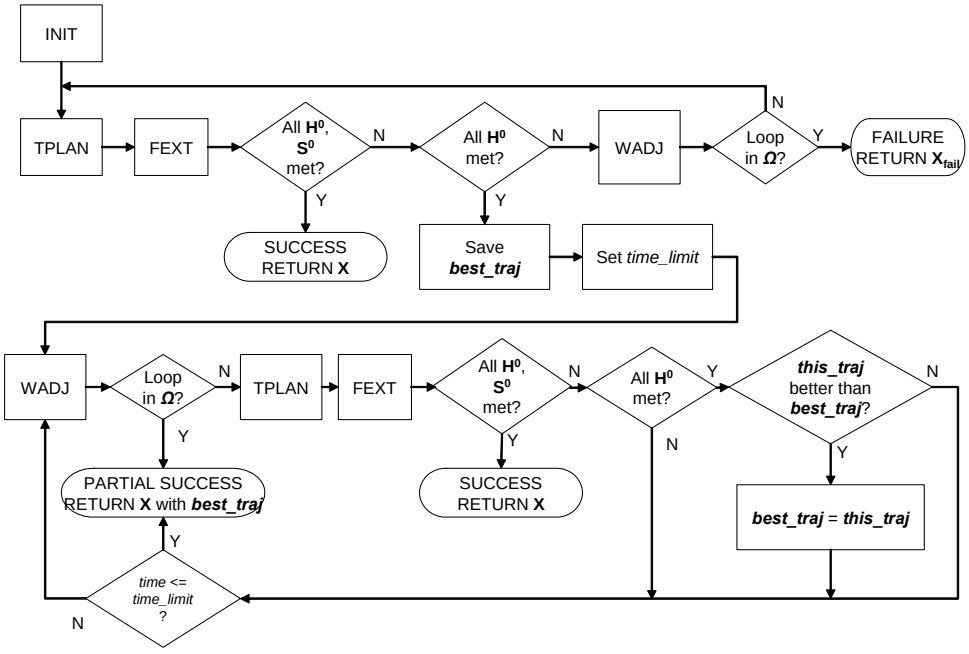

Figure 2: Paths through the architecture

loop is used to try and meet them. New trajectories are compared to the best found so far, using the 2-norm of the error vector as a measure of "goodness." The development of good weight adjustment (*WADJ*) heuristics was a key part of this work[43, 44, 45]. Our goal was to automate the process by which cost functional weights are tuned. This is typically done by hand, by a domain expert, until the desired results are achieved. We have attempted to encode these desired results into the limits $\mathbf{L}^0$, as functions of the features defined above. What remains is to extract domain expert knowledge and techniques and automate the adjustment process.

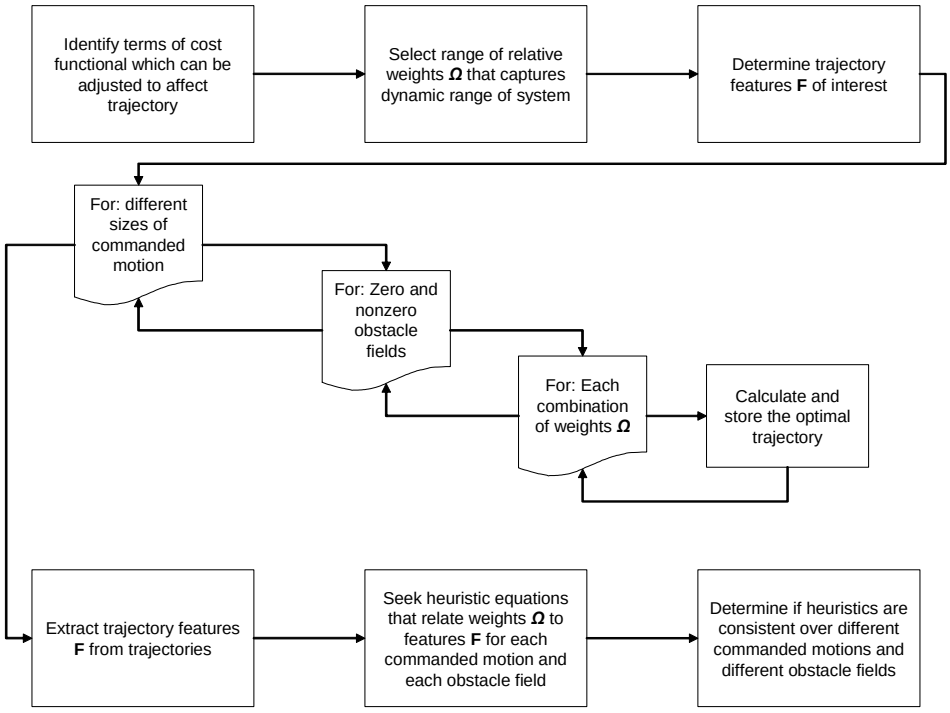

Figure 3: Developing *WADJ* rules

Early in this research we discovered that many of the features in our set could be expressed as functions of the weights used in the cost functional. Despite the different dimensionalities, cost functionals, and system dynamics, in the 2DOF and 6DOF systems we simulated, we were able to treat the generation of our *WADJ* heuristics in a similar fashion in both cases.

Each test matrix covered a combinatorial set of cost functional term weights, $\boldsymbol{\Omega}$. Since a cost functional can always be normalized, we knew that we would be looking at relative weights rather than be concerned with their absolute magnitudes. Early experimentation led us to conclude that a range of two orders of magnitude, from 0.1 to 10, would be sufficient to see a broad range of dynamic behavior in our systems. We varied the magnitude of the commanded motions and also the number of obstacles in the field. This was to ensure that the answers we were getting were not too specific to a single domain subcase. Once we had collected the trajectories, we used the feature extractor *FEXT* to compute the overall trajectory features in which we were interested. For features relating to time (e.g., velocity, acceleration, power), we found strong power relationships between the feature values and the *ratio* of the energy or fuel term weight and the time term weight (W_1/W_2):

$$time_feature = c_1\left(W_1 / W_2\right)^{-\alpha} \tag{11}$$

The exponent -α in these equations stayed fairly constant across field sizes, although the constant coefficient c_1 varied. Similarly, for path-based features like minimum separation from obstacles (*min-sep),* there was a linear relationship between the feature and the influence limit (*LIM*) in the obstacle penalty function:

$$path_feature = c_2 LIM \tag{12}$$

LIM is simply the distance over which the obstacle penalty function goes to zero, as measured from the edge of the obstacle. Rather than attempt to calculate tables for all possible constant coefficients c_1 and c_2 of these equations for all possible field sizes, they are computed online, using the current weight and feature values to back out the coefficient value. The coefficient, together with the *desired* feature value (e.g., the limit, if it was passed), are then used to recompute the weights. As more obstacles are added to the field, the rules' accuracy is affected. (Domain-specific examples are given in the relevant chapters.) They remain, however, useful rules of thumb for guiding weight adjustment, as our results will show.

The effects of changing the fuel/time weight ratio or the energy/time weight ratio and *LIM* were largely independent. This allowed us to decouple them, an important and useful assumption. They are not, however, entirely independent. As *LIM* decreases, for example, more direct paths which save both fuel and time can be found. The effect is not dramatic, but can mean the difference between a successful and unsuccessful solution. If the standard *WADJ* rules have failed to find a solution that mediates between competing time and fuel goals, a secondary *WADJ* rule will change *LIM* to take advantage of this secondary effect and find a successful solution.

We were concerned that the 6DOF spacecraft domain with nonlinear dynamics would not be amenable to this *WADJ* rule-generation process. Results for the 6DOF domain were in fact very similar to those for the 2DOF domain. A notable difference was the torque weight term, which is unsurprising given the coupled nature of the rotational and translational mechanics. Our torque heuristic is discussed below.

IV. Two Degree of Freedom Point Rover

A simplified 2-D domain model was developed as an intuitive baseline case for our architecture and as a method of developing initial modules to populate the Figure 1 architecture. We began our investigations with a 2DOF point-robot model, imagining a rover-like robot traveling in a plane, using electric motors for propulsion. We used this highly simplified domain to gain an intuition into the process of adjusting the cost functional weights and computing, then evaluating, the resulting trajectories. The model has simple linear dynamics:

$$\begin{bmatrix} \boldsymbol{x}'(t) \\ \boldsymbol{x}''(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -c_s / m \end{bmatrix} \begin{bmatrix} \boldsymbol{x}(t) \\ \boldsymbol{x}'(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{u}(t)/m \end{bmatrix} \tag{13}$$

where m is object mass and c_s is the coefficient of sliding friction. We assume an idealized system without motor saturation and perfect trajectory tracking. Our cost functional was an aggregated, weighted sum which penalized electrical energy use, time, and nearness to obstacles. To evaluate the performance of our system for the 2D robot domain, five different logical sets of constraints $\mathbf{L}^0$ of varied complexity were enforced on four different obstacle fields **{O}** for a total of twenty trials. There were, overall, 28 hard limits ($\mathbf{H}^0$) and 72 soft limits ($\mathbf{S}^0$). The simplest constraint set enforced one hard and two soft constraints; the most extensive had two hard constraints and six soft constraints.

Each of the twenty test cases was run from a default weight vector $\boldsymbol{\Omega}^1_{default}$ = [1, 1, 1, 1] and from a $\boldsymbol{\Omega}^1$ provided by *INIT*. Only the collocation *TPLAN* was used for trajectory generation. The results after one iteration (labeled *Default*^{1} and *INIT*^{1}) and after completion (*Default*^{n} and *INIT*^{n}) were examined for both starting weight vectors. Figure 4 shows the total number of failures for each of the four cases *Default*^{1}, *INIT*^{1}, *Default*^{n}, and *INIT*^{n}. *INIT*^{1} shows a clear advantage over *Default*^{1}, with both fewer failures to meet $\mathbf{H}^0$ and $\mathbf{S}^0$. We are not in this work formally working within an anytime planning framework, but this significant improvement in solution quality for the first iteration would be of benefit should we extend the work in that direction.

Final results are not nearly as dramatically different as initial results. Our different starting points in these cases did not, after repeated applications of *WADJ*, result in significant differences in final solution. (We believe that this demonstrates the robustness of our approach.) However, *INIT*^{n} converged to an acceptable solution in, on average, 5.25

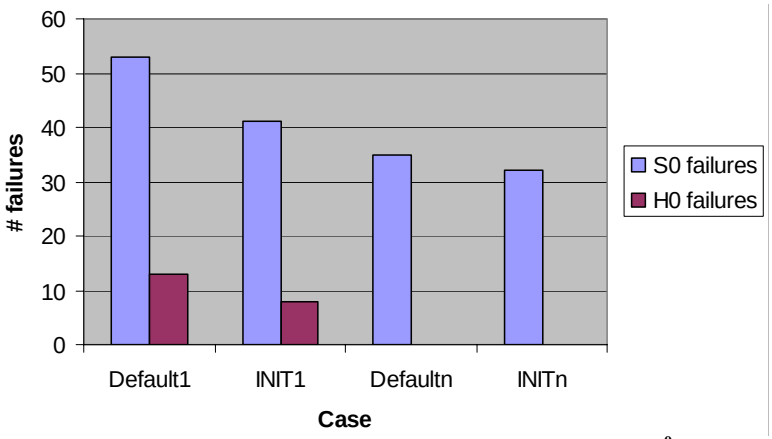

Figure 4: Failures for each 2DOF solution case out of 28 $\mathbf{H}^0$ and 72 $\mathbf{S}^0$

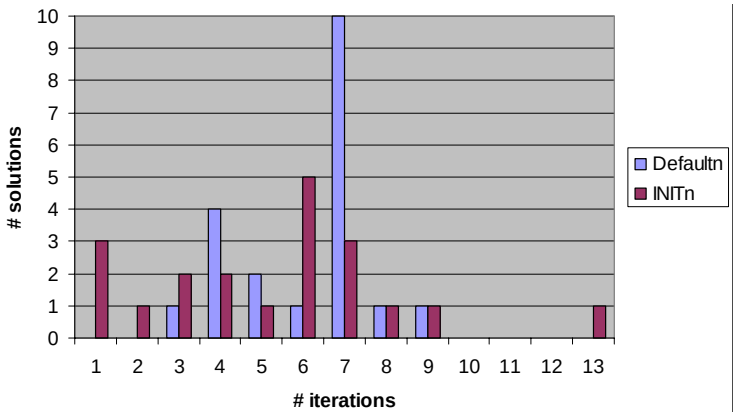

Figure 5: 2DOF iterations through architecture

iterations, whereas *Default*^{n} required on average 6.10 iterations, and it never found an acceptable trajectory on the first try. Figure 5 shows a histogram of number of iterations each solution required before returning. Overall, *INIT*^{n} has more valid solutions with fewer iterations than *Default*^{n}. The effects of obstacle arrangement did not have a strong affect on the number of $\mathbf{S^0}$ failures, as shown in Figure 6. Failure rates in $\mathbf{S^0}$ were impacted by the constraint set, as might be expected. Figure 7 gives the failure rates. Constraint Sets 1 and 4 had the overall highest rates. In Constraint Set 1, competing hard and soft constraints forced the desired *max-speed* into a 0.2 m/s window, a difficult value to achieve.

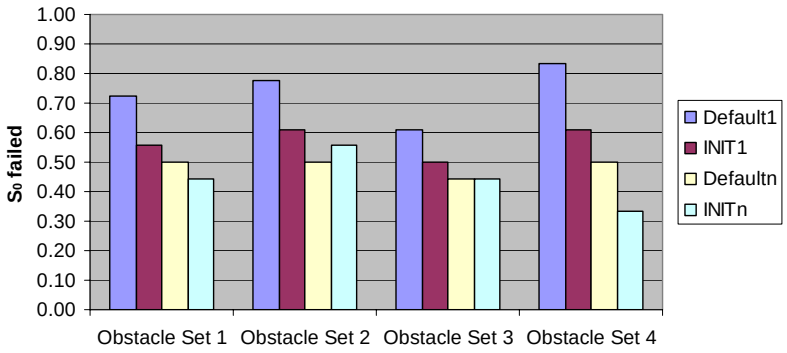

Figure 6: 2DOF $\mathbf{S^0}$ failure rates by obstacle set

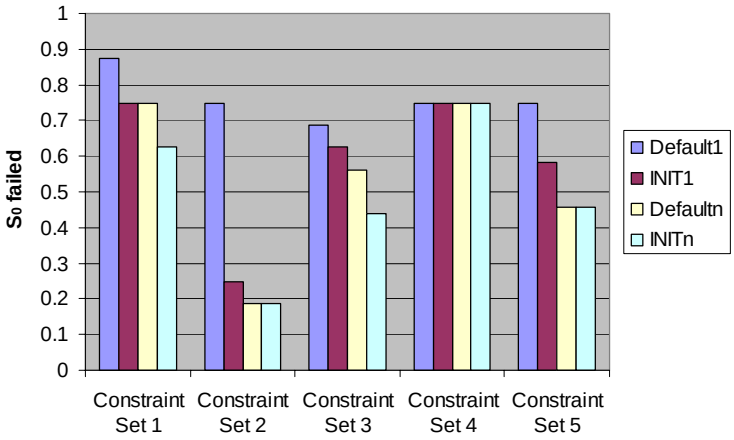

Figure 7: 2DOF $\mathbf{S^0}$ failure rates by constraint set

The Constraint Set 4 failure rates reflects a weakness in the *WADJ* rule generation method. Constraint Set 4 included two H_0 upper constraints on *max-acc* and *max-speed*, and then $\mathbf{S^0}$ numeric range constraints on *energy* and *avg-speed*. Except for Obstacle Set 4 (the easier obstacle set in Figure 6), the *energy* constraint was uniformly violated. Maneuvering around obstacles requires more energy than predicted by the fuzzy rules obtained from the *WADJ* curve data generated in empty space or in a field with one obstacle.

Out of all 40 test cases run to completion, 5 were able to meet all $\mathbf{H^0}$ and $\mathbf{S^0}$. In matters of solution quality, we found that initialization made little difference when meeting $\mathbf{H^0}$. By the final iteration, the margins by which the $\mathbf{H^0}$ had been met were similar, regardless of whether or not the system had started from initialized weights.

The margin by which $\mathbf{S^0}$ of the returned trajectory were made or failed also did not appear to depend on initialization. However, if this architecture were used to support anytime planning, the initialization had clear benefits. The first trajectory returned using initialized weights tended to meet more constraints overall and meet them by better margins. When it failed a constraint, the initialized solution tended to fail it by less than the solution returned by the default weight set.

V. Six Degree of Freedom Deep Space Satellite

The 2DOF experiments showed our architecture could be useful, but the 2DOF point rover problem was highly simplified and linear. As a next-step extension, we adopted a three-dimensional, six degree of freedom (space) domain example with flat space (gravity-free) dynamic properties that were nonlinear in attitude but less complex than orbital motion. We assumed a spacecraft with impulsive thrusters for translation and reaction wheels for torque generation, following the modeling described in Henshaw[46].The general state-space form of a rigid body in deep space is given by:

$$\boldsymbol{x}'(t) = \begin{bmatrix} \boldsymbol{v}' \\ \boldsymbol{p}' \\ \boldsymbol{\omega}' \\ \boldsymbol{\sigma}' \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{v} \\ -\mathrm{H}^{-1}\mathrm{S}(\mathrm{H}\boldsymbol{\omega})\boldsymbol{\omega} \\ G_{\sigma}(\boldsymbol{\sigma})\boldsymbol{\omega} \end{bmatrix} + \begin{bmatrix} \boldsymbol{u}(t)/m \\ 0 \\ \mathrm{H}^{-1}\boldsymbol{\tau}(t) \\ 0 \end{bmatrix} \tag{14}$$

where $\mathbf{p}$ is the spacecraft position vector in the inertial reference frame, $\mathbf{v}$ is its inertial velocity vector, $\boldsymbol{\omega}$ is the angular velocity vector expressed in the body frame, and $\boldsymbol{\sigma}$ is a representation of attitude (a modified Rodrigues vector[47]). H is a matrix of moments of inertia, S is the matrix representation of the cross product, and G_σ is an expression which, when multiplied by ω, gives the rate of change in σ:

$$\dot{\boldsymbol{\sigma}} = G_{\sigma}(\boldsymbol{\sigma})\boldsymbol{\omega} = \frac{1}{2}\left(\mathrm{I} - \mathrm{S}(\boldsymbol{\sigma}) + \boldsymbol{\sigma}\boldsymbol{\sigma}^{T} - \frac{1+\boldsymbol{\sigma}^{T}\boldsymbol{\sigma}}{2}\mathrm{I}\right)\boldsymbol{\omega} \tag{15}$$

where R is the rotation matrix from body to inertial coordinates, $\boldsymbol{u}$ is the translational control input (force) vector in the body frame, m is the spacecraft mass (presumed constant over each trajectory planning problem), and $\boldsymbol{\tau}$ is the vector of rotational control input (torque). Following Henshaw, the mass used was 1 kg and $I_{xx} = I_{yy} = I_{zz} = 1$ N-m/s^2, simulating a "micro-satellite" rather than a full-scale platform launched today. Maximum thruster output was ±30,000 N in each axis. Maximum torque output about each axis was smooth until a saturation value of ±1,000 N-m. We have the same concerns for 6DOF as for 2DOF: fuel and power must be conserved, goals must be accomplished in a timely fashion, and obstacles must be avoided. Our inputs are different in this case. Rather than a continuous electrically-powered motor, we have saturating thrusters for 3DOF translation and an electrically-powered reaction wheel for 3DOF attitude control. As a result, the cost functional has the form:

$$J = \int_{t_o}^{t_f}\left(W_1\|\boldsymbol{u}(t)\|_1 + \boldsymbol{\tau}(t)^T W_2 \boldsymbol{\tau}(t) + W_3 + W_4 \max_{i\in\{O\}}\left(o_i(r_i)\right)\right)\mathrm{d}t \tag{16}$$

The first term, the one-norm of the thruster force, results in a minimum-fuel control law. This control law is, however, discontinuous (which violates the assumptions of our numeric solver) thus is approximated with steep but continuous functions, such as a cubic spline made to increasingly approximate a step function[46]. The second term represents electrically-powered rotational actuators. This form of the cost functional is an energy-minimizing term, a standard cost representation for electrically-powered systems. The third term minimizes time over the integral. The fourth term is the obstacle penalty function. As in the 2DOF case, the obstacle penalty function contains a cubic spline term that penalizes proximity to obstacles. This function also includes a velocity-based component that penalizes speed near the obstacle. Since the cost functional is an integral over time, a penalty based purely on clearance to the obstacle can be minimized by being very close to the obstacle but going very fast, so that the sum over time is less – not a vehicle behavior we would typically wish to encourage.

A. Development of Weight Adjustment Heuristics and Fuzzy Rules

For the 6DOF weights, the *TPLAN* code BVP4C2 assumed that the force weight W_1 was normalized to 1, and all other weights were relative to this. As a result, it was more intuitive to work with the W_1 as the denominator in the weight ratios for our 6DOF spacecraft domain. All of the code written to implement these 6DOF heuristics uses torque/force and time/force weights, rather than the inverse as in the 2DOF case. A subset of the *WADJ* heuristics and fuzzy rules reflect this inversion and are labeled W_2/W_1 and W_3/W_1 as implemented. Our weight vector $\boldsymbol{\Omega}$ included W_1 the force weight, W_2 the torque weight, W_3 the time weight, and *LIM*. Since the obstacle penalty function weight W_4 is never adjusted, we do not include it in our weight vector representation.

To develop WADJ rules, we followed the general procedure outlined for the 2DOF example again in 6DOF. We did not test different field sizes for 6DOF, however, as we were confident from 2DOF that results would scale well. (This confidence was well-placed; our *WADJ* curves were generated at a scale of 50 m while our test cases were on the order of 10 to 20 m.) We tested pure translation, translation plus rotation, and a translation in the presence of obstacles. For the translation in the presence of obstacles, the final state orientation was identical to the initial state orientation; rotation was not required, but it was not forbidden, either. Following

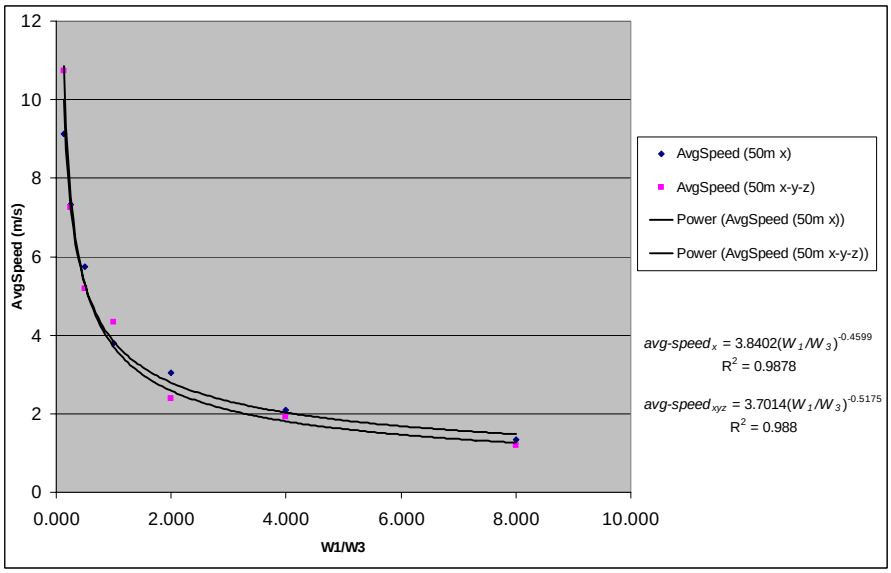

Figure 8: WADJ curve for *avg-speed* in 6DOF domain

the insights gained in the 2DOF domain, we plotted "per second" features versus the ratio of the weight of the translational inputs (here, thruster force) W_1 to the time weight W_3. The results for the feature *avg-speed* are shown in Figure 8. Once again, there is the power relationship between speed and the force/time weight ratio. We found this to be the case for the other force and time-based quantities as well. The path-based features (e.g., *min-sep*) were once again linear with the influence limit of the obstacle penalty function. Unlike the 2DOF case, the trajectories were much more likely to be plotted through obstacles. To handle this, we added an implicit hard constraint to every trajectory, *min-sep* > 0 m. If the path went inside an obstacle, the trajectory failed and *LIM* was adjusted to move the path out of the obstacle. This solved that problem.

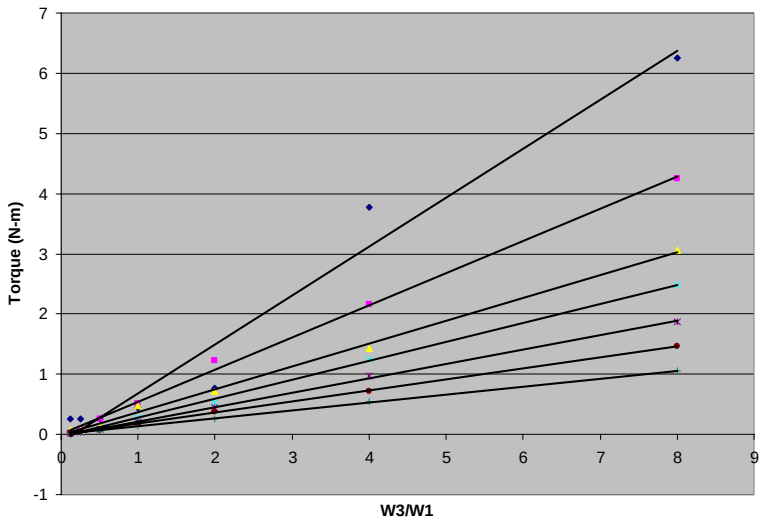

Figure 9: First stage *WADJ* heuristic to determine torque in 6DOF

Torque presented us with a challenge. Our *WADJ* test for torque included a translation and a rotation, so that we would see the effects of dynamic coupling. Following the intuition we had from the translational features, we tried plotting the ratio of the torque weight, W_2, versus the time weight W_3. However, since torque and the resulting rotational motions are the source of nonlinearity in the system, initial results indicated no power law for *WADJ* thus we at first were concerned this heuristic may not be applicable. Upon further examination, however, we identified a more promising heuristic. Figure 9 shows the torque data grouped by time and force. First, the data were grouped by their torque/force weight ratio (W_2/W_1) but plotted versus the time/force ratio (W_3/W_1) as shown in Figure 9. Each line in Figure 9 represents a fixed W_2/W_1 ratio. Even though that ratio is fixed, the amount of torque applied can be increased or decreased by adjusting the W_3/W_1 ratio. Conversely, if the W_3/W_1 were known and fixed, changing the W_2/W_1 ratio could jump the torque up or down that family of linear curves. Was there a predictable relationship between the slopes of the lines in Figure 9 and W_2/W_1 ? Figure 10 shows that there was. Our torque heuristic was implemented as follows: First, all non-torque features were checked for limit failures and, if there were failures, the weights were adjusted. Then the torque feature was checked. If it failed, the desired torque value was divided by the current W_3/W_1 value to get the slope of the line we would like to be traverse in Figure 9. Then the power relationship shown in Figure 10 was used to calculate the necessary W_2/W_1 ratio from the desired slope. The fuzzy rules were generated as they had been for the 2DOF case. The *WADJ* data was reverse engineered so that "very high," "high," etc., weight values were correlated back to the trajectory features they elicited. We again found that the range from 2^{-3}, 2^{-2}, …, 2^2, 2^3 was sufficient to describe the *WADJ* relationship.

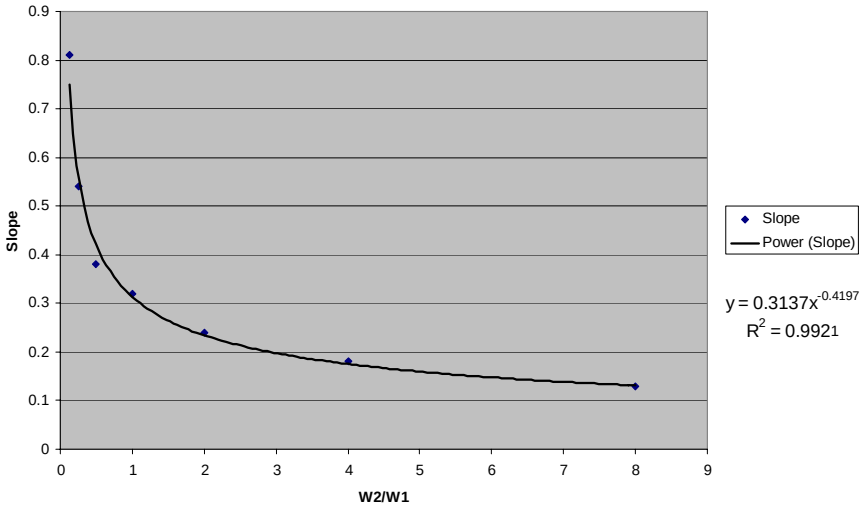

Figure 10: Second stage in torque heuristic in 6DOF

B. Results

Six different sets of constraints $\mathbf{L}^0$ were enforced on four different obstacle fields {**O**} for a total of twenty-four trials. However, one constraint/obstacle pairing proved to be intractable and BVP4C2 could not converge on a solution. This trial (Constraint Set 4, Obstacle Set 4) is omitted from the following results. There were, overall, 20 hard limits ($\mathbf{H}^0$) and 60 soft limits ($\mathbf{S}^0$). The simplest constraint set enforced one soft constraint; the most extensive had two hard constraints and six soft constraints. Each of the twenty-three successful test cases was run from a default weight vector $\boldsymbol{\Omega}^I_{default} = [1, 1, 1, 1]$ and from a $\boldsymbol{\Omega}^I$ provided by *INIT*. Again, only the collocation *TPLAN* BVP4C2 was used for trajectory generation. The results after one iteration (labeled $Default^1$ and $INIT^1$) and after program completion ($Default^n$ and $INIT^n$) were examined for both starting weight vectors.

Figure 11 shows the total number of failures for each of our solution cases. As in the 2DOF case, these are $INIT^1$, the solution generated using the weights suggested by *INIT*, $INIT^n$, the solution generated by running the $INIT^1$ solution to conclusion through the architecture, $Default^1$, the solution generated using a default weight vector with all weights equal and *LIM* = 1, and $Default^n$, the $Default^1$ solution run to completion.

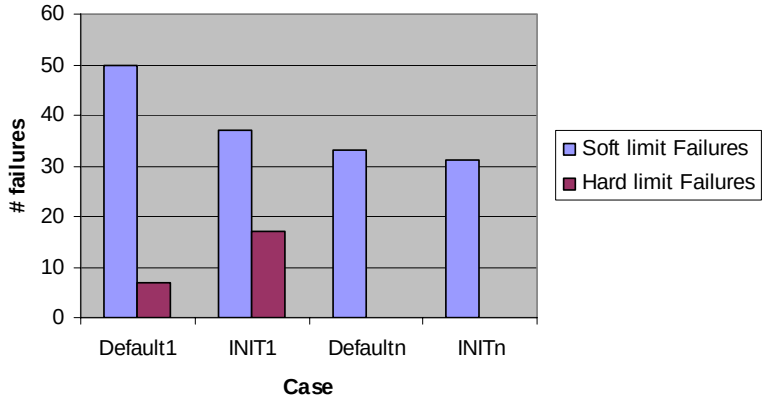

Figure 11: Failures for each solution case out of 20 $\mathbf{H}^0$ and 60 $\mathbf{S}^0$

The results for soft limit failures $\mathbf{S}^0$ are similar to the 2DOF case. The *INIT* procedure results in noticeably fewer soft constraint failures after only one iteration. There are more hard limit $\mathbf{H}^0$ failures with *INIT* due to our examples with competing constraints, but again no hard limit failures were present in the final solutions. Margins of success and of failure were largely comparable between the methods, and similar average numbers of iterations through the architecture were needed to complete the trajectory. However, using INIT did have one advantage, as shown in Figure 12: a majority of the runs started with *INIT* finish in one or two runs, while those starting from default weights need a minimum of three runs. Analyzing the $\mathbf{S}^0$ failures by obstacle set shows trends similar to the 2DOF case, although the median failure rate is higher. In the 2DOF case, the median $\mathbf{S}^0$ failure rate was 50% ±6% and no clear relationship to obstacle set. For this 6DOF case, the median $\mathbf{S}^0$ failure rate was 65% with a spread of 12% - greater variability, but no clear trends by obstacle set.

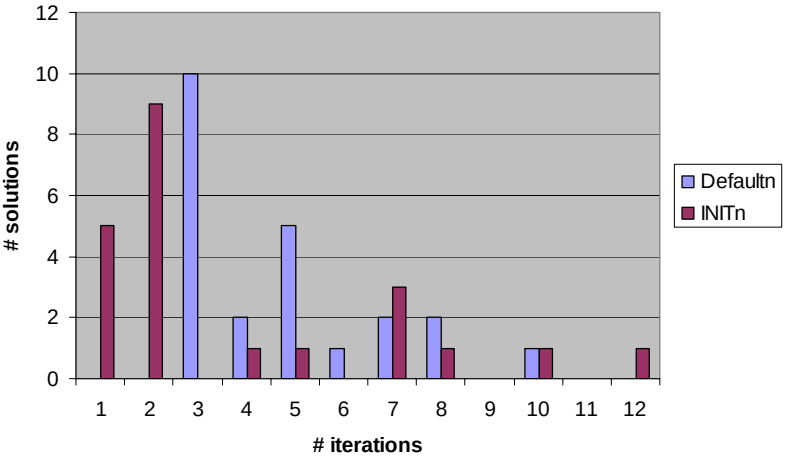

Figure 12: Number of iterations through architecture

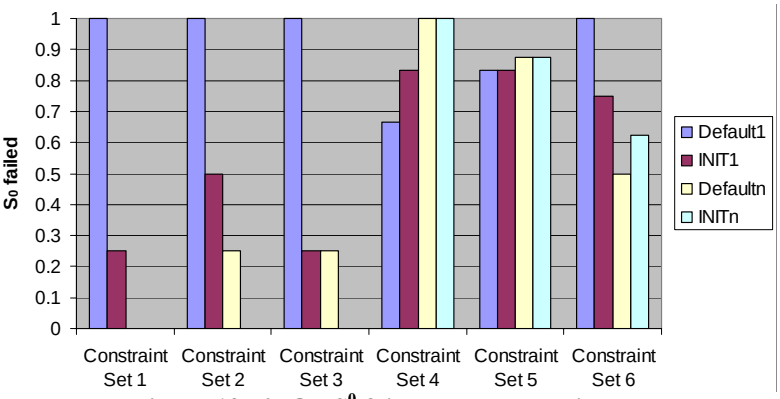

Figure 13: 6DOF $\mathbf{S}^0$ failures by constraint set

Figure 13 shows the percentage of $\mathbf{S}^0$ failures by constraint set. Here we see definite trends, with some constraint sets being apparently simple to entirely satisfy, while others had 100% $\mathbf{S}^0$ failure rates. Constraint Sets 2 and 3 had small numbers of noncompeting soft constraints and no hard constraints.

Constraint Set 2 had a soft numeric range limit on thrust; Constraint Set 3 was "a little quickly," which defuzzified into soft constraints on *max-speed* and *avg-speed*. With no other requirements, these constraints were solved much more successfully. $INIT^n$ solved them entirely for all obstacle cases; $Default^n$ had small errors with Obstacle Set 1 and Obstacle Set 4. Constraint Set 1 was very similar to Constraint Set 1 in the 2DOF case; we set a hard limit on *max-speed* and also the soft preference for "somewhat quickly." The hard limit was toward the low end of the fuzzy ranges that define "quickly," forcing the system to hit a small window of feature values that would satisfy both. While this constraint set gave the 2DOF case some problems, here it was entirely successfully solved in all cases that went to completion.

We also note that for these first three constraint sets, $INIT^1$ returns remarkably better initial solutions than $Default^1$. By happenstance, the default weights produce results that do not meet any of the $\mathbf{S^0}$, while $INIT^1$ achieves at least partial success in that. So we see here one practical application for *INIT*: in those cases where there are few or noncompeting constraints, it provides an excellent initial guess compared to using default weights. Constraint Sets 4 and 5 were clearly less successful. Constraint Set 4 included two hard upper limits on *max-acc* and *max-speed* and two soft range limits on *force* and *avg-speed*. It appears that when the suggested weights for the *force* and *avg-speed* ranges were combined, the force terms were much more sensitive to the change away from their own desired values. Further, the hard limit on *max-acc* was greatly exceeded in all initial cases. By the end of the iterations, the hard limits were all met, but *force* was failed in all cases: failed under the lower limit of the range. By requiring such a low *max-acc*, we were required to use less thrust than specified by the soft range. Similarly, all final *avg-speeds* failed low, as the trajectory had to go slowly enough to meet the upper limit on *max-speed* (a hard limit). Essentially, the stated soft constraints in Set 4 *had* to fail for the hard constraints to be met.

Constraint Set 5 added the soft fuzzy preference "moderately safely" to Constraint Set 4. This defuzzified into four more soft range constraints. We were practically guaranteed a certain failure rate, since the upper limit of "safely's" *avg-speed* constraint equaled the lower limit of the soft range constraint on *avg-speed* from Constraint Set 4. Of course, if we failed low on *avg-speed*, as we did for all of the Constraint Set 4 cases, we would be making the "safely" constraint, decreasing our overall failure rate. Soft constraints on *max-acc* and *max-speed* arising from "safely" were sometimes met when the hard constraints were met, again decreasing the failure rate. (And when they were failed, they failed low as in Constraint Set 4.) We saw very large failure margins which were greatly reduced by the end of the iterations.

Constraint Set 6 was "very energy-saving," which decomposed into soft range constraints on *force* and *torque*. But since the only torque needed in the trajectories was that required to avoid obstacles, the trajectories all failed low; they could not use enough torque to satisfy the "low *torque*" constraint. "Low *force*" was more typically made, or failed high with very small margins (0.05, 0.02 N) for the completed cases ($Default_n$, $INIT_n$). Since the nonlinearity of the system is in its rotational dynamics, and since we wish to show that our architecture will work with nonlinear systems, we decided to rewrite the $\boldsymbol{x_f}$ to include explicit rotational changes rerun it over the set of obstacles under Constraint Set 6. The results are presented separately. For many of the runs, a solution was returned after between one and three iterations (Figure 12). For one or two iterations, the weights are converged upon with no overshoot. For those default weight cases that took three iterations, some converged monotonically to the correct weights, while others overshot on the second iteration and corrected with the third

C. Torque

In our original data set, only one of eight completed trajectories (4 $Default^n$, 4 $INIT^n$) met the "low *torque*" requirement. The rest were too low to be considered "low" by the standards of our fuzzy rule set. Since none of our goal states required a rotation, the only rotations required were those needed to avoid obstacles. These did not use sufficient torque to be considered low. So, to test the *torque WADJ* rules, we included a rotation change in each axis at the goal state and re-ran the tests. We also had some concerns about the possible interaction of *force* and *torque*. In *WADJ*, all features except *torque* are first checked for adjustment. Then the selected W_3/W_1 ratio is used together with the desired *torque* value to calculate a slope from Figure 9; that slope is then used to pick a W_2/W_1 ratio via the equation in Figure 10. The "low *force*" requirement was keeping W_3/W_1 small, and the heuristic is less well-conditioned for W_3/W_1 less than 1. Although meeting mixed constraints is an important goal, we also wanted to isolate the torque response to the *WADJ* process, since it is so different from the other *WADJ* heuristics. So we created additional test cases: Constraint Set 7, "low *torque*," and Constraint Set 8, "medium *torque*."

Figure 14 shows our overall failure rates for these three constraint sets (CS 7, 8, and the revised CS 6); each case had 16 runs (four soft constraints run over four obstacle sets). The *Default*^{1} and *INIT*^{1} cases are high again, not unexpectedly, and the failure rates for the completed runs are much lower. All seven failures at run completion were torque failing low; of those seven, four were from the "medium torque" Constraint Set 8. W_2/W_1 was continually adjusted down to discount it, to allow for greater torque in these cases, but what was required was that W_3/W_1 be increased. Since there was a tacit assumption that some other state feature would be relying on W_3/W_1 and that it may have been adjusted to affect that other feature, the torque *WADJ* never altered W_3/W_1, and W_2/W_1 could not be adjusted sufficiently before *time_limit* was reached. Our current *TPLAN* cannot handle direct maximization of trajectory qualities; it can only minimize. We have found that we can minimize features which are inversely related to our feature of interest for a maximizing effect; thus by penalizing time, we can usually force in increase in speed. Another *TPLAN* might allow for direct maximization of features.

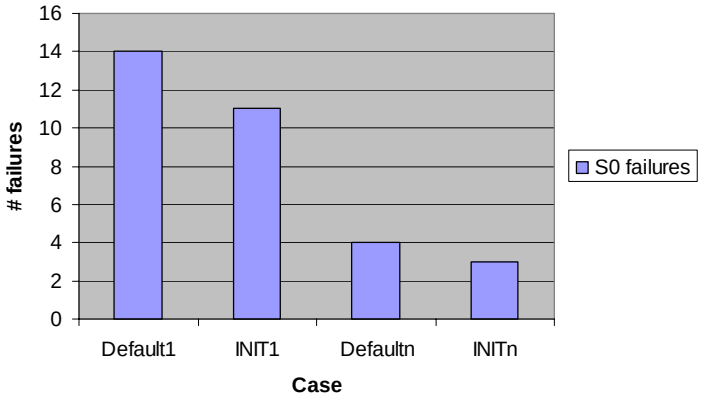

Figure 14: 6DOF S^0 failures for Constraint Sets 6, 7 and 8

Figure 15 shows the number of iterations required for these runs. All of those runs which took four or fewer iterations to return a solution returned a complete success. The utility of *INIT* is again shown in the large number of runs that returned successful trajectories after only one or two iterations; eight (2/3 of the total) of the trajectories created using *INIT* were solved in two or fewer iterations, while only 3 trajectories created using default weights met this standard.

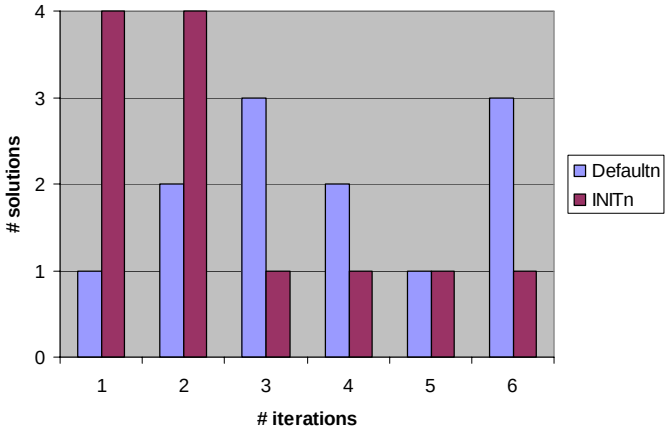

Figure 15: Iterations required for Constraint Sets 6, 7 and 8

We have demonstrated that *WADJ* heuristics can be developed for a deep space 6DOF domain with nonlinear dynamics. Our results were, if anything, better in the 6DOF domain than in our 2DOF domain, with smaller $\mathbf{S^0}$ failure margins and larger success margins for $\mathbf{H^0}$. The average number of iterations required to find a solution was commensurate with the 2DOF case, also arguing that the technique implemented in the architecture will scale well with the dynamic complexity of the domain. The surprising similarity of the 2DOF and 6DOF *WADJ* curves, even to the values of the coefficients, is noteworthy. That also argues for the potential for a general application to the optimization of dynamic systems.

The performance of our *TPLAN*, BVP4C2, was less robust and far more slow than we had hoped. Prior work on this algorithm required extensive hand-tuning of several sets of gains just to solve a single trajectory problem. We were running it, on average, 5.2 times per problem for 24 fairly different problems. So these difficulties are not entirely unexpected. In the future, however, a different *TPLAN* should be selected for work with nonlinear systems.

VI. Conclusions and Future Work

The 2DOF and 6DOF examples have shown that our architecture can optimize trajectories over hard constraints and natural language preferences. The *WADJ* heuristics consistently direct the weights toward values that meet hard and soft constraints and are robust to differences in initial weight sets. The fuzzy logic enables a more natural human interface, opening a route to easy tasking of autonomous agents by non-expert users (e.g., hospital staff commanding a robotic assistant, warfighters with a Future Combat System robot, the elderly using a companion robot). However, the ability to meet hard numeric constraints is not lost in adding the fuzziness. This allows the system to be used as an "automated graduate student," overseeing trajectory generation, rejecting those which do not

meet required hard constraints, and making intelligent adjustments to the weights to move the solution in the required direction. Substantial knowledge engineering and preprocessing was required to develop the fuzzy rules, the *WADJ* rules, and the *TPLAN* implementation. But once this offline process was completed, the system was applicable to a wide range of obstacle and constraint conditions with no further adjustments. This makes the architecture useful for robots operating long-term in a consistent environment, but not so useful for "one off" operations such as technology demonstrations.

In the future, a more sophisticated version of *INIT* could look at the constraint set as a whole and recognize potential conflicts. Currently, the system will execute many iterations to satisfy conflicting constraints. An early detection of this kind of possible conflict, or else a software monitor that detects a pattern of cycling back and forth would both be useful to have.

INIT should also be invariant to the order in which constraints are processed. In this implementation, the order in which the hard constraints are considered impacts the returned weight vector. After the soft constraints have been aggregated via a centroid computation, *INIT* cycles through the hard constraints and checks to see if the currently suggested weights are liable to meet them. If they are not, *INIT* adjusts the weights up or down as needed. If competing constraints are being considered, the last one addressed by *INIT* will be favored, rather than a median weight which might satisfy both.

The *WADJ* process could also be rendered more sophisticated. After the *INIT* cycle, the "adverbial modifiers" like "very" or "somewhat" are lost in the weight adjustment process. The endpoints of the fuzzy regions for the soft constraints are fixed, without regard to the strength of the user's preference. Nor are they currently considered when deciding which of several competing soft constraints must fail. The assumption has been that the *INIT* process would put the solution in approximately the correct region in weight-space, and further iterations would reflect that. That assumption does not necessarily hold, as the *WADJ* rules can cause oscillations of initially very large, then decreasing, magnitude in weight space. Something that preserves the knowledge of soft preference strength past the *INIT* phase would help this adhere more closely to true user preference, and perhaps reduce total iterations needed.

A more sophisticated notion of error margins in *FEXT* might also be of use here. A *WADJ* algorithm that seeks to minimize the entire vector of errors, rather than each error individually, would be computationally more expensive (an optimization within an optimization) but could yield superior results with fewer iterations.

Finally, other forms of *WADJ* specific to other cost functionals could be explored. A cost functional based on a linear quadratic regulator (LQR), in which components of the state vector like the velocities are directly penalized, could replace the time component of the cost functionals used here. Of course, these new terms would still have weighting terms and the relationships between them would have to be investigated, following the procedures outlined here.

We would like to augment *EVAL* with an understanding of the adverbial modifiers, as mentioned above, so that preferences the user described as weaker would be violated in favor of meeting more strongly-held preferences. Additionally, some mechanism whereby the original set of limits $\mathbf{L}^0$ can be revisited and perhaps altered by the architecture is an avenue of further research. There could be cases where the slight easement of a limit could lead to an overall acceptable solution; we would like to be able to identify these cases and flag them for the user. In this vein, the addition of "firm" versus "hard" or "soft" constraints might be considered: those constraints which the user very greatly prefers to be met, but which do not indicate total failure if failed.

Some optimization routines use negative weights in the cost functional, to allow certain terms (e.g., a quality measure) to be maximized. Users must be very careful when doing this, because it becomes possible for the term to grow without bound as time goes to infinity. The cost goes to negative infinity, dominated by this term times its negative constant. If the user has determined that, due to the properties of his particular problem, this will not happen, then such a term can be used. This work does not investigate the possibility of adding such terms, and we could look to that in the future as well.

Acknowledgements

This work was performed in part at the Naval Research Laboratory under funding from the Office of Naval Research under work order N0001404WX30001.

References

[1] Johan Andersson. "A survey of multiobjective optimization in engineering design." Technical report LiTH-IKP-R-1097, Department of Mechanical Engineering, Linköping University, Linköping, Sweden, 2000.

[2] Bart Kosko and Satoru Isaka. "Fuzzy logic." *Scientific American* 269, pages 76-81, July 1993.
[3] Vilém Novák. "Fuzzy logic: applications to natural language." In Stuart C. Shapiro (editor), *Encyclopedia of Artificial Intelligence, Second Edition,*, pages 515-521, John Wiley & Sons, New York, 1992.
[4] Dennis Perzanowski, Alan C. Schultz and William Adams. "Integrating natural language and gestures in a robotics domain." In *Proceedings of the IEEE International Symposium on Intelligent Control*, pages 247-252. National Institute of Standards and Technology, Gaithersburg, MD, 1998.
[5] Barbara Tversky and Paul U. Lee. "How space structures language." In C. Freksa, C. Habel, and K. F. Wender (editors). Spatial cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge. pages 157-175, Springer-Verlag, Berlin, 1998.
[6] Alicia Abella and John R. Kender. "Qualitatively describing objects using spatial prepositions." In Proceedings of IEEE Workshop on Qualitative Vision, pages 33-38, New York, 1993.
[7] Patrick Oliver, Toshiyuki Maeda and Jun-ichi Tsujii. "Automatic depiction of spatial descriptions." *Spatial Reasoning* 2, pages 1405-1410, 1994.
[8] Amitabha Mukerjee. "Neat vs. scruffy: a survey of computational models for spatial expressions." In Patrick Olivier and Klaus-Peter Gapp (editors), *Computational Representation and Processing of Spatial Expressions*, Kluwer Academic Press, Boston, MA, 1998.
[9] Oussama Khatib. "Real-time obstacle avoidance for manipulators and mobile robots." *The International Journal of Robotics Research* 5(1), pages 90-98, 1986.
[10] Jean-Claude Latombe. *Robot Motion Planning*. Kluwer Academic Publishers, Boston, MA, 1991.
[11] Ronald C. Arkin. *Behavior-Based Robotics.* The MIT Press, Cambridge, MA, 1998.
[12] Juan C. Velásquez. "An emotion-based approach to robotics." In *Proceedings of the 1999 IEEE/RSJ International Conference on Intelligent Robotics and Systems*, pages 235 – 240, Kyongju, Korea, 1999.
[13] Maxim Likhachev, Michael Kaess, and Ronald C. Arkin. "Learning behavioral parameterization using spatio-temporal case-based reasoning." in *Proceedings of the 2002 IEEE International Conference on Robotics and Automation (ICRA 2002),* pages 1282 – 1289, Washington, DC, 2002.
[14] Juan C. Santamaría and Ashwin Ram. "Learning of parameter-adaptive reactive controllers for robotic navigation." In Proceedings of the World Multiconference on Systems, Cybernetics, and Informatics (CSI '97), Caracas, Venezuela, 1997.
[15] Eric Aaron, Harold Sun, Franjo Ivančić, and Dimitris Metaxas. "A hybrid dynamical systems approach to intelligent low-level navigation." In *Proceedings of Computer Animation 2002*, 154-163, San Antonio, TX, 2002.
[16] Siome Goldstein, Menelaos Karavelas, Dimitris Metaxas, Leonidas Guibas, Eric Aaron, and Ambarish Goswami. "Scalable nonlinear dynamical systems for agent steering and crowd simulation." *Computers and Graphics* 25(6), 983-998, 2001.
[17] Daniel Shapiro and Pat Langley. "Separating skills from preference: using learning to program by reward." In *Proceedings of the Nineteenth International Conference on Machine Learning*, pages 570-577, Sydney, Australia, 2002.
[18] Paolo Fiorini and Zvi Shiller. "Motion planning in dynamic environments using velocity obstacles." *International Journal of Robotics Research*, 17(7), pages 760-772, 1998.
[19] Zvi Shiller, F. Large and S. Sekhavat. "Motion planning in dynamic environments: obstacles moving along arbitrary trajectories." In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2001)*, Seoul, Korea, 2001.
[20] Jianwei Zang, Jörg Raczkowsky and Andreas Herp. "Emulation of spline curves and its applications in robot motion control." In *Proceedings of the IEEE Conference on Fuzzy Systems*, pages 831-836, 1994.
[21] Jung-Hoon Hwang, Ronald C. Arkin and Dong-Soo Kwon. "Mobile robots at your fingertips: Bezier curve on-line trajectory generation for supervisory control." In *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003),* pages 1444-1449, Las Vegas, Nevada, 2003.
[22] Steven M. LaValle and James J. Kuffner, Jr. "Randomized kinodynamic planning." International Journal of Robotics Research, 20(5), pages 378-400, 2001.
[23] Paolo Fiorini and Zvi Shiller. "Time optimal trajectory planning in dynamic environments." *Journal of Applied Mathematics and Computer Science*, 7(2), pages 101-126, 1997.
[24] Zvi Shiller and Yu-Rwei Gwo. "Dynamic motion planning of autonomous vehicles." IEEE Transactions on Robotics and Automation, 7(2), pages 241-249, 1991.
[25] Ian Garcia and Jonathan P. How. "Trajectory optimization for satellite reconfiguration maneuvers with position and attitude constraints."
[26] John H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, 1975.

[27] Francis Ysidro Edgewood. *Mathematical Physics*. P. Keagan, London, 1881.
[28] Vilfredo Pareto. *Cours D'Economie Politique*, volume I and II. F. Rouge, Lausanne, 1896.
[29] M. Farina and P. Amato. "On the optimal solution definition for many-criteria optimization problems." In *Proceedings of the NAFIPS-FLINT International Conference 2002*, pages 233--238, New Orleans, LA, 2002.
[30] Arturo H. Aguirre, Salvador B. Rionda, Carlos A. Coello Coello, Giovanni L. Lizárraga, and Efrén M. Montes. "Handling constraints using multiobjective optimization concepts." *International Journal for Numerical Methods in Engineering*, 59(15), pages 1989-2017, 2004.
[31] Oliver de Weck and Marshall B. Jones. "Isoperformance: Analysis and Design of Complex Systems with Known or Desired Outcomes." In Proceedings of the 14th Annual International Council on Systems Engineering, Toulouse, France, June 2004.
[32] Oliver de Weck, David W. Miller and Gary E. Moiser. "Multivariable Isoperformance Methodology for Precision Opto-Mechanical Systems." 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Denver, CO, April 2002.
[33] Il Yong Kim and Oliver de Weck. "Adaptive Weighted-Sum Method for Bi-Objective Optimization: Pareto Front Generation." Structural and Multidisciplinary Optimization, 29 (2), pages 149-158, February 2005.
[34] Tom Schouwennaars, Bart De Moor, Eric Feron and Jonathan How. "Mixed integer programming for multi-vehicle path planning." In *Proceedings of the European Control Conference*, pages 2603-2608, 2001.
[35] Arthur Richards and Jonathan How. "Performance evaluation of rendezvous using model predictive control." In *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Austin, Texas, 2003.
[36] Ian Garcia and Jonathan P. How. "Trajectory optimization for satellite reconfiguration maneuvers with position and attitude constraints." In Proceedings of the IEEE American Control Conference, pages 889-895, 2005.
[37] A. E. Bryson Jr. and Y. C. Ho. Applied *Optimal Control*. Blaisdell, Waltham, MA, 1969.
[38] Donald E. Kirk. *Optimal Control Theory: An Introduction*. Dover Publications, Inc., Mineola, NewYork, 2004 (reprint of 1970 edition by Prentice-Hall, Inc., Englewood Cliffs, NJ).
[39] S. Roberts and J. Shipman. *Two-Point Boundary Value Problems: Shooting Methods*. Elsevier, New York, 1972.
[40] Oscar von Stryk. "Numerical solution of optimal control problems by direct collocation." In R. Bulirsch, A. Miele, J. Stoer, and K.-H. Well (editors), *Optimal Control – Calculus of Variations, Optimal Control Theory and Numerical Methods, number 111 in International Series of Numerical Mathematics*, Birkhauser, Basel, 1993.
[41] Carl Glen Henshaw. "A unification of artificial potential function guidance and optimal trajectory planning." In *Proceedings of the 28th AAS Annual Rocky Mountain Guidance and Control Conference*, pages 219-234, Breckenridge, Colorado, 2005.
[42] Jacek Kierzenka and Lawrence Shampine. "A BVP solver based on residual control and the Matlab PSE." *ACM Transactions on Mathematical Software* 27(3), pages 299-316, 2001.
[43] J. Lennon. An Architecture for the Autonomous Generation of Preference-Optimized Trajectories. PhD dissertation, University of Maryland Department of Aerospace Engineering, 2006.
[44] J. Lennon and E. Atkins. "Intelligent Weight Selection for Trajectory Optimization." In *Proceedings of InfoTech@Aerospace*, Arlington, VA, September 26-28, 2005.
[45] J. Lennon and E. Atkins. "Multi-Objective Spacecraft Trajectory Optimization with Synthetic Agent Oversight." *Journal of Aerospace Computing, Information and Communication 1(1)*: 1-20, 2004.
[46] Carl Glen Henshaw. *A Variational Technique for Spacecraft Trajectory Planning*. PhD dissertation, University of Maryland Department of Aerospace Engineering, 2003.
[47] Panagiotis Tsiotras. "Stabilization and optimality results for the attitude control problem." *AIAA Journal of Guidance, Control, and Dynamics*, 19(4): pages 772-779, July-August 1996.