Multicriteria Evaluation of a Planetary Rover*

Gaurav S. Sukhatme[1] and George A. Bekey[1]

gaurav|bekey@robotics.usc.edu

[1]Department of Computer Science
Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, CA 90089-0781

Abstract

Evaluation of mobile robots is a multidimensional problem. There are many performance indices that can be used to compare and evaluate mobile robots. Examples are speed, stability etc. A natural framework for the evaluation of mobile robots is multicriteria evaluation - the science of evaluation functions whose output is nonscalar. In this paper we present a methodology for mission oriented mobile robot evaluation using a multicriteria framework. The methodology is illustrated by implementation on a four wheeled microrover class robot for mobility performance in a simulated Martian terrain. Our methodology makes use of nondimensional quantities which measure rover mobility with respect to its surroundings so that rovers of different sizes may be meaningfully compared. Ongoing and future work is discussed with regard to implementing the methodology on another robot built in our lab and predicting the values of the evaluation functions using validated simulations.

1 Introduction

There has been recent interest in mobile robot evaluation as evidenced by the work of Gat (Gat 1995), Sukhatme et al. (Sukhatme & Bekey 1995) and Matthies et al. (Matthies, Gat, Harrison, Wilcox, Volpe & Litwin 1995). Of particular interest is the issue of evaluating mobile robots for space applications such as planetary exploration. In this paper we argue that the proper framework for the evaluation of mobile robots involves the use of multiple, mission dependent criteria. A common thread of exasperation in the mobile robotics community is the paucity of well agreed upon metrics of performance that are considered standard. Progress in mobile robot evaluation will come to fruition when different metrics are proposed and debated and some standardization ensues. This paper is part of ongoing work to model the mobile robot evaluation problem in a multicriteria framework. We propose two evaluation metrics here which form the two criteria. We discuss how such 2-vectors may be compared and ranked using techniques from the multicriteria evaluation literature.

In December 1996 NASA will launch the pathfinder mission to Mars (Matthies et al. 1995). The mission will deploy an autonomous microrover on the surface of Mars. The robot's primary mission is to do spectroscopic analysis on Martian rocks of interest to scientists on Earth. Due to challenges such as round trip delay time the robot is equipped to perform a high degree of autonomous goal seeking and hazard avoidance behavior.

Missions such as the one above are strong motivation for reliable, repeatable and standard evaluation techniques. The evaluation techniques proposed in this paper are part of a project to develop mission oriented evaluation metrics suitable for robots built for extraterrestrial exploration. We use a simplified version of the Mars mission as a scenario in our experimental setup. During the course of our project we have built two robot rovers for exploration. The ultimate project goal is the development of performance metrics that can compare the mobility performance of different robots in a meaningful manner. In this paper we report on data taken from experiments with the wheeled robot called Marscar. Data from the legged

*This work is supported in part by Jet Propulsion Labs, California Institute of Technology under contract #959816.

robot MENOII and a comparison between the two robots is the subject of a future paper.

2 Previous Work

Mobile robotics lies at the intersection of many subjects. Of particular relevance to this work are previous proposed evaluation methods for mobile robots and the vast field of multicriteria evaluation. It is impossible to do full justice to either in a short paper of this nature. We however do include a short literature review of each.

Mobile Robot Evaluation

Static evaluation techniques are designed for measuring stability when the robot is stationary or when it is moving in a statically stable fashion. The primary method of choice is an energy based stability measure. Early work on robot stability was due to McGhee et al. (McGhee & Frank 1968) who defined the support polygon as the convex hull of the projections of all contacting points on a horizontal plane. A walker was considered stable if the projection of center of gravity on the horizontal plane lay within the support polygon.

In work by Nagy et al. (Nagy, Desa & Whittaker 1994) two modes of walker stability are characterized namely stance stability and walker stability. Both use the amount of energy needed to destabilize the walking robot as a measure of the stability of the robot. The energy stability margin was defined in (Messuri & Klein 1986) as the minimum work that must be done on a robot walker to tip it over an edge of a support boundary. In (Mahalingam & Whittaker 1989) the authors define a conservative support polygon with the motivation that the walking robot should retain its stability in the event of a single leg failure. In (Bares & Whittaker 1993) the authors discuss several evaluation criteria such as stride, stability etc. for comparing three configurations for the design of a walking robot.

Wilcox (Wilcox 1991) introduced a metric called the MCC (Mobility Characteristic Curve) to measure the ability of a robot to surmount obstacles. The obstacle was a cylinder of (theoretically) infinite length and diameter d which was buried to a depth $d/3$ in an inclined plane of slope s composed of loose sand. The MCC was defined as the plot with s on the horizontal axis and the diameter of the largest cylinder that the robot could surmount (in dimensionless units based upon its length) on the vertical axis. The proposed figure of merit was the area between the co–ordinate axes and the MCC. The two main achievements of this method were its independence of scale and easy reproducibility. Its chief drawback was that it used a simple obstacle geometry and did not evaluate the entire system in a mission oriented way.

Lietzau (Lietzau 1993) proposed a set of benchmarks to assess the performance of a Mars microrover. These benchmarks were divided into five categories namely, mobility, navigation and control, science, autonomy and environmental. A set of weights was assigned to these categories based upon their importance by the system designers and mission specialists. The weighted sum of the individual benchmarks was then proposed as a figure of merit. Lietzau's work is a thorough description of the individual subsystem tests that are a necessary part of evaluation but does not focus on the system level evaluation that we emphasize here. The main problem with doing isolated tests and then assigning a composite score using a weighted combination technique is that it presupposes the existence of a good model for the combination and involves some amount of arbitrary assignment of weights.

Gat (Gat 1995) proposed an evaluation technique that is both statistical and mission oriented using Rocky IV as a testbed on simulated Martian rock distributions (both physical and computer simulations). His technique measures the displacement of the rover as a function of time as it navigates autonomously from start to goal. The experiment is run several times over multiple rock distributions to infer a statistical distribution of the displacement with respect to elapsed time. In (Gat 1995) no scalarized metric or index of performance is proposed. The evaluation function is also not based on multiple cost functions with all their attendant complexities (such as dealing with vector outputs). (Gat 1995) is an excellent example of the usefulness of statistical evaluation techniques and the value of simple kinematic simulations to predict robot performance.

In (Matthies et al. 1995) the authors propose a variety of evaluation techniques to measure and improve the performance of a Mars microrover. The evaluation techniques proposed include the work of Gat (Gat 1995) and other metrics. The authors discuss detailed experimental studies of mobility performance based on goal designation, rover localization, hazard detection and path selection.

In (Sukhatme & Bekey 1995) the authors propose an evaluation technique which is statistical and uses multiple cost functions. In (Sukhatme & Bekey 1995) the evaluation technique was illustrated using simple kinematic simulations. It is the subject of the current paper to report the results of similar experiments on a real robot.

Multicriteria Evaluation

An intuitively appealing way to evaluate a complex system is to break it into parts and assign each part some "weight". The system is decomposed into parts so that each part has a simple structure and is easy to evaluate [1]. The composite system is then assigned a figure of merit using a weighted combination (possibly linear) of the parts. Such a decomposition is not possible for many complex systems. These systems may be viewed as an interaction of several subsystems which are *strongly coupled* to each other and are *not* independent.

The extension is therefore to model coupling between subsystems and continue using the above weighting technique. The result is a branch of systems theory called multicriteria evaluation. The domain of the cost function is now a set of vectors. Simply put, multicriteria evaluation is the branch of research that deals with non-scalar optimization i.e. the range of the evaluation function is a set of vectors.

One cannot therefore maximize the cost directly in a multicriteria problem (one ends up maximizing some scalarized function of the cost like its norm for instance) but one can still make some progress towards a solution. Research in multicriteria evaluation has focussed largely on generating the so called noninferior set which separates the feasible solutions from the infeasible ones and leaving the final choice of solution to a human designer/specialist. The notion of the noninferior set was introduced by Pareto (Pareto 1971) in the context of a multitude of objectives in economics. The non-inferior set is called the set of Pareto-optimal solutions.

Literature on methods to find the noninferior set is extensive; (Cohon & Marks 1974) is an example. All the basic techniques to find the non-inferior set rely on transforming the objective function into a scalar function (*scalarization*).

[1] In this intuitive argument assigning a figure of merit to each part is called "evaluating" it.

3 The Evaluation Methodology

Our evaluation methodology uses animal ecology as an inspiration. As an example Skellam (Skellam 1951) studied the spread of muskrats over Central Europe. He was able to show that the dispersal of this species over time is a diffusive process with no drift. In other words over small periods of time muskrats effectively execute a random walk. The study of such diffusive processes is however of a statistical nature - one is not able to predict where a given animal will be at any point in its life, rather one is able to predict what fraction of the muskrat population is within a particular radius of some (known) start location.

It has been pointed out that modeling robot behavior is difficult to a large extent because of the uncertainty associated with slippage and sensor data. In (Gat 1995) it is pointed out that one may however develop a simple kinematic model of a robot and run it several times over simulated terrain to measure its mobility in a statistical sense. The result may then be compared against several runs of the real robot. The increased number of runs serves to reduce the demands on the fidelity of the simulated model.

In our experiments we start the robot at a particular location and command it to go to some goal location. The robot autonomously navigates to the goal using a combination of reactive obstacle avoidance and dead reckoning. In the process an overhead vision system is used to measure the position of the robot at regular time intervals. Many such trials are done with different rock placements of the same density (i.e. the number of rocks of a particular diameter is held constant). Using this data we are able to plot the average displacement from the start location as a function of time. To go back to the analogy with the muskrat study, each trial simulates a new member of a large population. The second group of trials we have done measures the energy consumed by the robot and measures the position of the robot at regular energy consumption intervals. These isoenergy measurements are also repeated over several trials and rock placements. As a consequence we are able to plot the average displacement of the robot from its start location as a function of its energy consumption.

The nature of our results is the same as the muskrat study - we cannot say where the robot will be for a particular value of time elapsed (or energy consumed), but we can calculate the likelihood of the robot being at a particular distance from its start location for some given value of time elapsed (or energy consumed). It is these probabilities that are used to define two metrics of performance.

Using the trials described above, two graphs are plotted. The first is the probability of reaching a particular distance r_0 from the start location as a function of elapsed time t. The second is the same probability plotted as a function of the energy e consumed by the robot. The quantity r_0 is written as kl where the constant k is mission dependent (Longer missions imply a larger k value for evaluation purposes). We denote the two probability functions as $P_{kl}(t)$ and $P_{kl}(e)$ respectively. The quantity l is a nondimensional length. It is defined as the ratio of the length of the robot to the diameter of the smallest rock that the robot considers an obstacle. In the trials reported here $l = 8$ i.e. the robot is 8 times as long as the smallest rock that it considers an obstacle. We consider relatively short missions here and use $k = 0.25$. The time t is also nondimensional - we measure time in terms of the length of the robot l and its average speed v. Nondimensional time is defined as the product of actual time elapsed t' and the average speed of the robot v divided by the length of the robot L. In other words $t = t'v/L$. In the same spirit the nondimensional energy e is defined as the actual energy consumed e' divided by the product of the robot mass m, the acceleration due to gravity g and the real length of the robot L. So $e = e'/mgL$. For the trials reported here the relevant quantities have the following values.

Quantity	*Symbol*	*Value*
mass	m	5 kg
length	L	0.4 m
width	W	0.3 m
average speed	v	0.22 cm/s

We define the **time** performance metric τ to be the area under the first curve from $t = 0$ to $t = t_0$ where t_0 is chosen using a mission oriented criterion (how long a mission is being planned). The **energy** performance metric η is defined as the area under the second curve from $e = 0$ to $e = e_0$ where e_0 is also chosen using a mission oriented criterion. The two metrics are thus defined as

$$\tau = \int_0^{t_0} P_{kl}(t)\ dt \tag{1}$$

Figure 1: Marscar in a simulated martian environment

$$\eta = \int_0^{e_0} P_{kl}(e)\ de \tag{2}$$

Note that both metrics are bounded above by t_0 and e_0 respectively. An 'ideal' robot would achieve these values; practical designs have lower scores. The reason behind defining the two metrics using nondimensional quantities is that it is desirable to use metrics which measure rover performance with respect to its surroundings. If this is done a meaningful comparison can be made between a variety of rover designs. We plan to use the same metrics to measure performance of another rover built in our lab which is legged instead of wheeled.

4 The Experimental Setup

The robot on which the testing was done is called Marscar (See Figure 1). It is a 4 wheeled microrover with Ackerman steering. The front wheels are actuated together with a single motor. The same holds for the rear wheels. The robot is $0.4m$ in length and $0.35m$ in width. It stands $0.24m$ high. It has a ground clearance of $0.04m$. Its total mass (with all sensors mounted) is $5kg$. All sensing and navigation decisions are made onboard. The power source (a high amphour $12V$ battery) is the only offboard component. The following are the sensors onboard the robot.

- A digital compass which is interfaced to the microcontroller on board. The compass has a resolution of approximately 0.1^o.

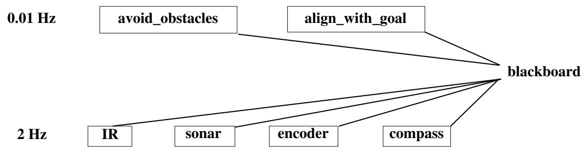

Figure 2: The software architecture

- A forward looking sonar for detecting obstacles. The sonar is used to gauge distance to a frontal obstacle. Empirical tests show that it is accurate to within 5 cm for echos from rocks without any special postprocessing of the sonar echo.
- Two forward looking infrared emitter/detectors pairs which have a maximum range of 10 cm. These are binary sensors and are use to detect whether there is an obstacle directly in front of either front wheel.
- Two shaft encoders on the front and rear motor shafts which detect shaft rotation which is then converted to linear distance traveled by the wheels.
- A current sensor that measures the amount of current drawn by the robot. This sensor is accurate to 0.5% over the range 0-70 A.

The robot uses dead reckoning interspersed with reactive obstacle avoidance to navigate to the goal. The navigation algorithm is given below:

```
Loop until at goal:
    If obstacle in front
        Detour
    Else
        Head towards goal
```

The robot starts its traverse in the simulated outdoor terrain (a sandbox of size $3.5m \times 3.5m$) at a location known to it in a frame of reference rigidly attached to the sandbox. The goal position is also specified to the robot in the same reference frame before the start of a traverse. The robot essentially runs two processes - the first is called `avoid_obstacles()` and causes the robot to move forward in the absence of obstacles. If the robot detects an obstacle in front of it it will back up and turn and try to detour around the obstacle. The second process is a goal seeking process - it interrupts periodically to align the robot's length axis with the line joining the robot center and

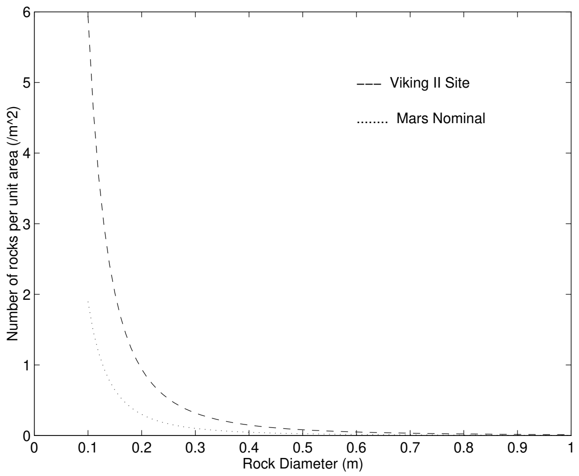

Figure 3: The exponential Moore distribution

the goal. This is called the `align_with_goal()` process. The robot keeps track of its current position as a result of each movement using its encoder - which it converts to linear position change of the body in the appropriate direction using its compass. In addition there are a set of low level sensor processes that run more frequently and update a set of global variables that store the robot position, direction etc. These variables are readable and writable by the two higher level processes. A schematic of the software architecture is shown in Figure 2.

There is currently no correction for wheel slip in the system. Indeed wheel slip is not detected at this stage. Such a navigation scheme suffers from the classical problem of drift in dead-reckoning.

The different rock placements used for the trials reported here were all drawn from a standard model of expected rock distribution on Mars (Moore & Jakosky 1989). The model is an exponential decay in the diameter of the rocks given in the following equation. N is the cumulative number of rocks per unit area and D is rock diameter.

$$N = aD^{-b} \tag{3}$$

From imagery taken at the landing site of an earlier Mars mission (Viking II) the value of b has been found to be 2.66. Indirect (thermal) measurements are used to obtain modal values for a for the entire planet. This is the so-called Mars nominal value $a = 0.00415$ used in the trials reported here. For the Viking II lander site $a = 0.013$. Figure 3 shows a plot of N vs. D for the Viking II lander site and the Mars nominal values.

The data collection proceeds using an overhead camera connected to a frame grabber attached to a PC. The program that runs on the PC operates in two modes. The first (or *time*) mode is when the program takes a picture of the robot at a fixed time interval (30 sec). These pictures are then postprocessed after the experiment to extract the position of the robot at fixed (30 sec) time intervals. The second (or *energy*) mode is when the program takes a picture when the robot tells it to. The robot has onboard a current meter that measures the current draw of the robot. Using this the robot keeps a running total of its energy consumption and triggers a image capture at fixed (2.5 kJ) energy consumptions increments. These pictures are also postprocessed after the experiment to extract the position of the robot at fixed (2.5 kJ) energy-consumption intervals. The experimental setup enables the measurement of the robot displacement (from the start location) as a function of **a.** time elapsed and **b.** energy consumed by the robot.

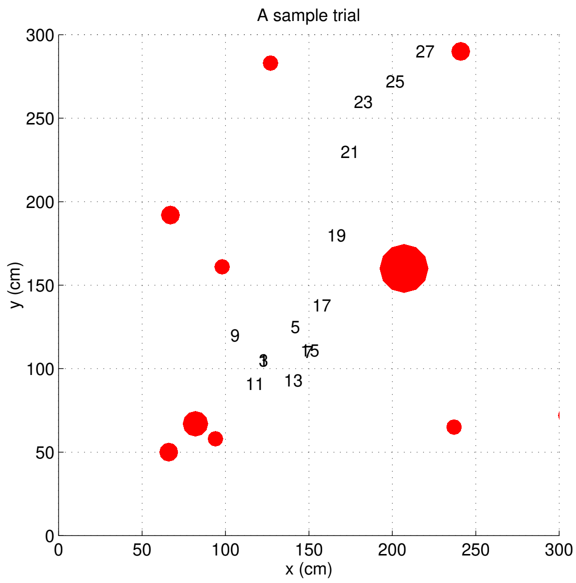

Figure 4: A typical trajectory

5 Experimental Results

Figure 4 below shows a sample trajectory of the robot as a function of time. Increasing numbers show successive positions of the robot with increasing elapsed time. The start position of this particular run was $(125, 100)$ and the desired target was $(250, 250)$. Due to error in dead reckoning there is a discrepancy between the actual final position of the robot and the desired final position. No attempt has been made at this stage to improve the dead reckoning performance. In this trial the robot first heads towards the goal and confronts the obstacle at $(200, 150)$. It backs up and tries to detour a couple of times. The `align_with_goal()` process causes a reorientation of the major axis of the robot at step 13 and another reorientation occurs at step 23. The trial ends at step 27 when the robot believes it has reached its commanded goal.

A total of 6 such trials was done for the isochronal data and a total of 6 trials was done for the isoenergy data. Two different rock placements (both Mars nominal) were used with 3 trials in each placement for both data sets.

The average displacement of the robot as a function of time is shown in Figure 5. The dots represent a second order best fit polynomial to the average plot. The average displacement of the robot as a function of the energy consumed is shown in Figure 6. The dots are, once again, a best fit polynomial. The figures below show the probability of reaching a distance kl vs. nondimensional time and nondimensional energy.

The time figure of merit τ is the area under the curve in Figure 7 and the energy figure of merit η is the area under the curve in Figure 8. From the data collected the values are $\tau = 1.98$ and $\eta = 420$.

6 Conclusions

In this paper an experimental methodology to evaluate the performance of mobile robots has been presented. The methodology basically measures the the mobility of the robot as a function of two resources, namely elapsed time and energy consumed. The methodology has several desirable features namely

- It is mission oriented
- It uses normalization for robot dimensions with respect to surroundings

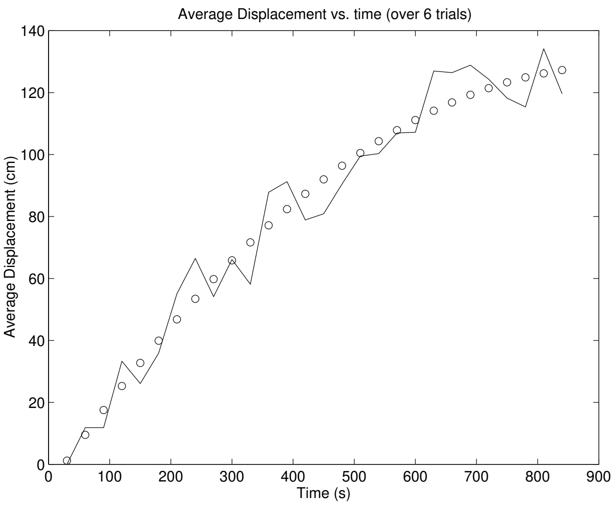

Figure 5: Average displacement vs. time

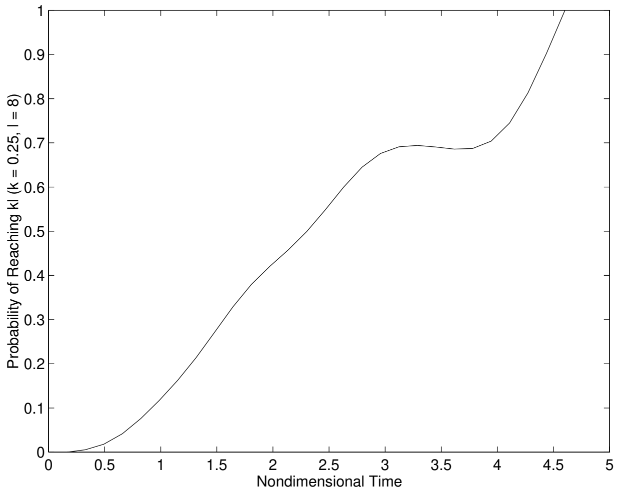

Figure 7: Probability of reaching kl vs. time

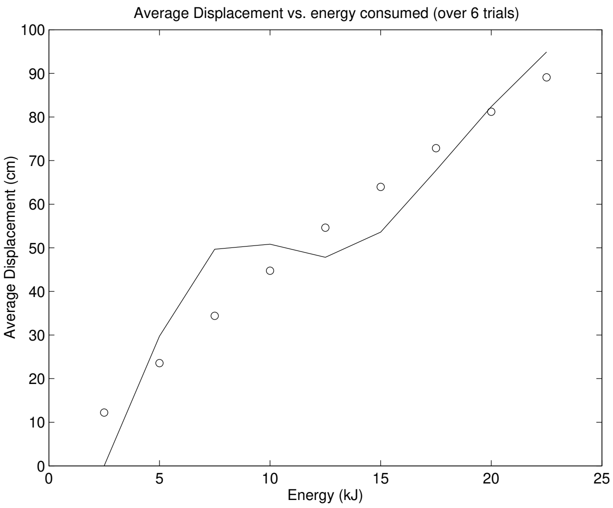

Figure 6: Average displacement vs. energy

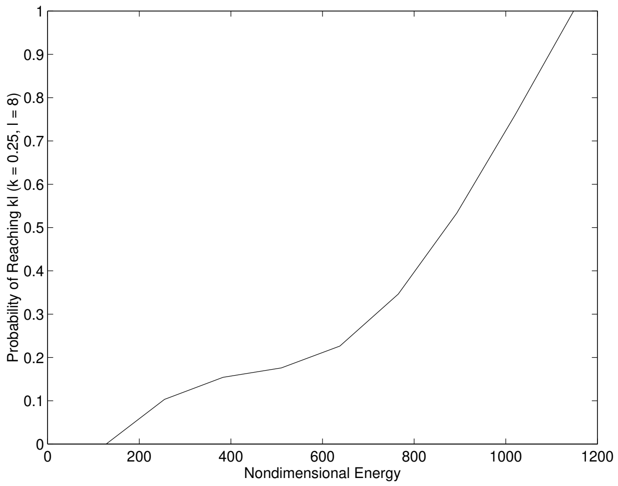

Figure 8: Probability of reaching kl vs. energy

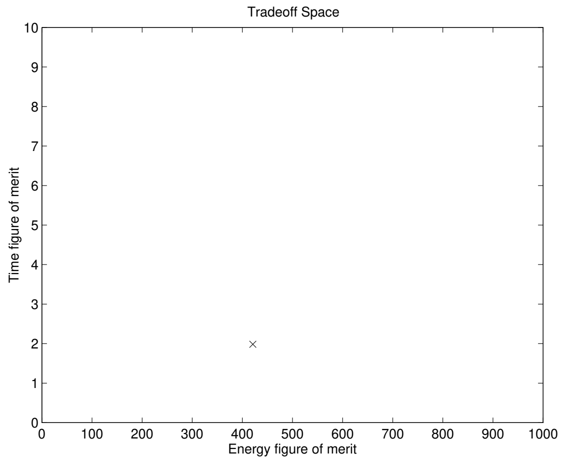

Figure 9: Time and energy figures of merit

- Multiple dimensions of evaluation are used

The results shown here do not yet address the issue of how to scalarize the two criteria that make up the evaluation function. One suggestion is to use a conversion factor between energy and time (based on current technology) and use that to convert one figure of merit into the units of the other. This scalarized quantity is amenable to ordering and one can compare robots using it. It is our plan to populate the so-called tradeoff space with other designs so that a comparison may be made between them. Work in the future includes more trials with the existing robot and similar trials with the legged robot built in our lab. Scalarizing the results of the evaluation is also an important component of future work. One of the main difficulties with doing the kinds of trials described here is that they are time consuming. Work is underway to develop simulations of the robots and to perform similar trials (in much larger numbers) in simulation. We are also interested in studies of sensitivity and scale using these metrics. Sensitivity issues include the effects of small parameter changes (examples are mass, ground clearance etc.) on the values of the metrics. Scaling issues include the effects of large changes in parameters (such as length, width etc.) on the values of the metrics proposed in this paper.

Acknowledgments

The authors would like to thank S. Hayati, G. Rodriguez, C. Weisbin and B. Wilcox for insightful discussions over the past few months. The authors also thank S. Cozy and S. Goldberg for experimental support and M. McHenry for his comments.

References

Bares, J. & Whittaker, W. (1993), 'Configuration of autonomous walkers for extreme terrain', *International Journal of Robotics Research* **12**(6), 535–559.

Cohon, J. L. & Marks, D. H. (1974), Multiobjective analysis in water resource planning, *in* R. de Neufville & D. H. Marks, eds, 'Systems Planning and Design', Civil Engineering and Engineering Mechanics, Prentice-Hall, pp. 304–321.

Gat, E. (1995), 'Towards principled experimental study of autonomous mobile robots', *Autonomous Robots.*

Lietzau, K. R. (1993), Mars micro rover performance measurement and testing, Master's thesis, Massachusetts Institute of Technology, Department of Aeronautics and Astronautics.

Mahalingam, S. & Whittaker, W. L. (1989), Terrain adaptive gaits for walkers with completely overlapping leg workspaces, *in* 'Proc. Robots 13'.

Matthies, L., Gat, E., Harrison, R., Wilcox, B., Volpe, R. & Litwin, T. (1995), 'Mars microrover navigation: Performance evaluation and enhancement', *Autonomous Robots.*

McGhee, R. B. & Frank, A. A. (1968), 'On the stability properties of quadruped creeping gaits', *Mathematical Biosciences* **3**(3/4), 331–351.

Messuri, D. A. & Klein, C. A. (1986), 'Automatic body regulation for maintaining stability of a legged vehicle during rough terrain locomotion', *IEEE Journal of Robotics and Automation* **1**(3), 132–141.

Moore, H. J. & Jakosky, B. M. (1989), 'Viking landing sites, remote-sensing observations and physical properties of martian surface materials', *Icarus* **81**, 164–184.

Nagy, P., Desa, S. & Whittaker, W. L. (1994), 'Energy-based stability measures for reliable locomotion of statically stable walkers: Theory and application', *The International Journal of Robotics Research* **13**(3), 272–287.

Pareto, V. (1971), *Manual of Political Economy: Translation of the French Edition (1927)*, The McMillan Press Ltd.

Skellam, J. G. (1951), 'Random dispersal in theoretical populations', *Biometrika* **38**, 196–218.

Sukhatme, G. S. & Bekey, G. A. (1995), Mission reachability for extraterrestrial rovers, *in* 'Proc. 1995 IEEE Int. Conf. on Robotics and Automation'.

Wilcox, B. (1991), 'Mobility characteristic curve', Jet Propulsion Labs, IOM 3472-91-019.