Motion Planning for the ATHLETE Rover with Reinforcement Learning

Patrick Mihelich

1. Introduction

Legged locomotion is attractive because it can enable a robot to traverse far more varied terrain than a wheeled rover is capable of. In the context of planetary exploration, this is especially attractive as the sites of greatest scientific interest tend to be characterized by difficult terrain. For example, it would be extremely difficult for a wheeled rover to make its way into a lunar crater in search of water.

Planning legged locomotion is, however, a more difficult problem the wheeled lomotion. Compared to wheeled robots, legged robots tend to have a much larger number of degrees of freedom. A planner for a legged robot therefore has to plan in a high-dimensional configuration space, placing considerable demands on the planner's efficiency.

The problem is further complicated when uneven terrain is considered. Classical motion planning techniques applied to legged locomotion generally assume a flat workspace with obstacles to be avoided. In the case of uneven terrain, there are no explicit obstacles, but a motion planner must take care to maintain the stability of the robot at all times while moving over sloped and uneven surfaces. For articulated legs, contact with the ground creates a closed-loop kinematic chain. Algorithms such as probabilistic roadmaps must be adapted to efficiently handle these closed-loop constraints, as a randomly sampled point in the configuration space has zero probability of exactly satisfying them.

Especially when the robot is intended for cooperative tasks with humans, an additional problem is planning trajectories which appear natural to a human observer. A planner for a humanoid bipedal robot, for example, might generate bizarre-looking arm motions which aid in balance, even on relatively flat terrain where such motions are unnecessary. A planner for legged locomotion should ideally encode constraints that encourage natural-looking motion.

The specific robot considered in this paper is the ATHLETE (All-Terrain Hex-Limbed Extra-Terrestrial Explorer) robot developed by the Jet Propulsion Laboratory (JPL). ATHLETE is intended to be a lunar rover, and is especially designed for movement over broken and uneven terrain. Its hexagonal frame is designed for carrying large payloads, or even a living capsule, so that it could be used for both transport and exploration by a lunar base. On flat terrain it can move quickly using wheeled locomotion. On uneven terrain it can fix its wheels and walk on its six articulated legs, and it is this mode of operation with which this paper is concerned. Each leg has six degrees of freedom. Adding six more DOF for the position and orientation of the chassis, ATHLETE has a total of 42 DOF.

Fig. 1. The ATHLETE rover

The goal of the present research is to develop a real-time on-line motion planning algorithm for ATHLETE that enables it to reach a goal location both quickly and safely. The approach used here is similar to that of Urmson et. al., who plan a global path using a traversability map and then select from a set of actions how to follow that path locally[3]. In this paper, the set of actions is defined to be a set of different gaits, and reinforcement learning is used to learn the action planner.

2. Approach

This paper proposes a three-part planning algorithm for legged locomotion over uneven terrain that uses a set of fixed gaits as a model for generating control actions. The use of fixed gaits drastically reduces the dimensionality of the configuration space and also results in natural-looking motion.

The full planning problem is decomposed into a global planner, an action planner appropriate for any state, and a set of gait planners:

(i) Field D* search is used to determine an approximate global trajectory for the robot's center of mass between the start and target locations, minimizing a measure of costs associated with the traversability of the terrain it crosses.
(ii) An action planner repeatedly selects a gait to use based on the traversability of the terrain currently occupied by the robot. Action selection is learned through reinforcement learning.
(iii) A gait planner determines the joint controls to step the legs according to the chosen gait in the direction of the next field D* waypoint.

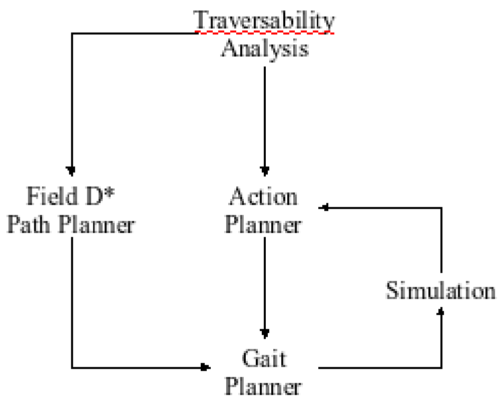


Fig. 2. Motion planning approach

For ATHLETE, we choose between two possible gaits, a "wave gait" and a "tripod gait." The wave gait moves each leg in turn keeping five legs on the ground at all times. The tripod gait moves three non-adjacent legs at a time, leaving only three legs on the ground at all times. Compared to the wave gait, the tripod gait trades stability for speed.

3. Traversability analysis

Naturally the robot must be able to avoid obstacles, but in varied terrain it also essential that it prefer traversing flat, smooth regions to sloped, rough ones. We therefore compute a continuous measure of traversability for use by both the global planner, to maximize the traversability of the global trajectory, and by the action planner, to select a gait appropriate to the difficultly of the terrain. We construct a traversability map using a simplified form of the Morphin algorithm [2].

We discretize the terrain into a grid of cells of fixed size. Due to ATHLETE's size, we choose the cell size to be smaller than the area occupied by the robot to allow for enough granularity in the global path planned by Field D*.

For each cell, we sample points from the terrain and perform a least-squares plane-fitting to those points. The slope of the plane is used as a measure of the terrain slope, and the chi-squared error of the sample points from the fitted plane as a measure of the roughness of the terrain. Scaling these measures by *a priori* maximum allowable terrain slopes and roughnesses for the robot and multiplying them gives a measure of "goodness." Morphin multiplies the goodness by a measure of uncertainty to determine traversability; since we assume full knowledge of the terrain geometry, we simply equate traversability with goodness.

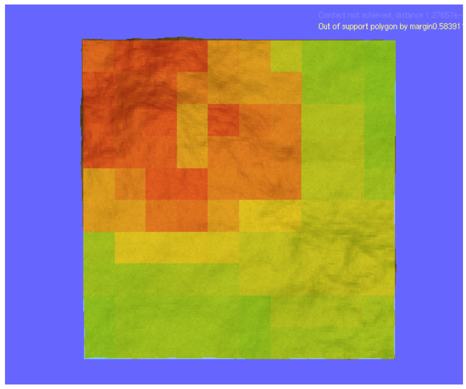


Fig. 3. Motion planning approach

4. Field D* Path Planning

For each terrain cell, we then calculate a weighted traversability average of surrounding cells (up to approximately the area of the robot) as a measure of the cost of moving the robot directly through that cell. Field D* search [1] is then used to plan a route through the terrain.

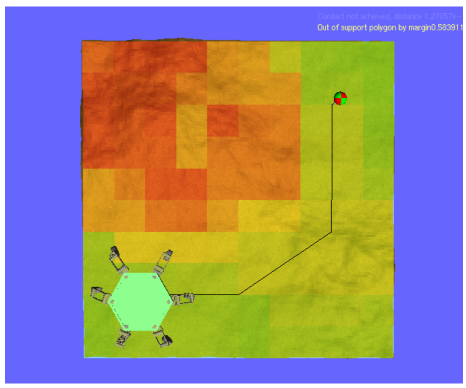


Fig. 4. Motion planning approach

5. Action Planner

Although we assume perfect knowledge of the environment and the robot state, the results of robot actions are nondeterministic. Due to deficiencies of our physics model, uncertainty in the controller, and unpredictable interactions with the environment, the state resulting from execution of a control action may differ from what was expected. This makes an open-loop planner unsuitable. Instead, we use reinforcement learning to develop a planner which maps states to control actions.

5.1. *Reinforcement learning*

The state of the robot includes its position and orientation, which gait the robot is currently employing, and the slopes and roughnesses of the terrain immediately surrounding the robot. For the purposes of action selection and learning, we reduce the state to a single feature, the minimum traversability of the terrain grid cells currently underlying the robot.

We take an action to be a single iteration of the chosen gait, in which all six legs take exactly one step. For example, a full iteration of the tripod gait involves two steps, each moving three legs in the same direction simultaneously. There is also a special action to select a different gait.

The reward function for an action incorporates the distance travelled, the stability of the robot during the action, and any "mishaps" that occurred. Stability can be measured by the area of the polygonal support region formed by the non-moving legs. Mishaps include exceeding the torque limits and self-collisions induced by over-extending one or more legs.

The planner is learned with value iteration. State transitions are learned by executing many actions in varied terrain in a physics simulation.

6. Gait Planner

The selected gait planner accepts a heading and generates the low-level control actions given to the motor controllers or physics simulation. The results of these control actions are used to learn the action planner.

The wave gait planner used moves each leg in turn, going counterclockwise around the frame of the rover. Each leg is moved in a parabolic arc in the desired direction until it comes in contact with the ground.

The tripod gait planner used alternates between movements of three non-adjacent legs. During each movement, it attempts to shift the frame of the body a fixed distance along the plane of the body (to accommodate sloping terrain). Inverse kinematics are used to determine where to place the moving legs so that they end in contact with the ground.

7. Results

Fig. 5 shows the results of the planning framework in simulation on a plot of uneven terrain. Field D* plans a global path skirting the hill to the northwest. Starting on relatively flat terrain, the action planner initially selects the tripod gait. It has learned to prefer this gait on relatively easy terrain, where additional stability is less essential, for its speed advantage over the wave gait. On reaching the bumpier area in the southeast, the action planner switches to the slower but surer wave gait. Upon reaching flatter ground close to the goal, it reverts back to the tripod gait.

8. Summary

This paper presents a framework for motion planning of legged robots over uneven terrain. Reinforcement learning is used to learn a planner which selects from a set of predefined gaits according to the difficulty of the terrain.

In future work, more choice in actions could be allowed. In particular, the action planner could be allowed to exercise some control over the heading of the robot. More state features would also be useful, to allow the action planner greater knowledge of the terrain and perhaps take into account the current pitch and roll of the robot.

The use of predefined gaits necessarily limits the difficulty of the terrain that can be safely traversed using this planning framework. It could be useful to incorporate a full footfall planner as an additional action to be chosen on particularly difficult terrain.

The framework could also be extended to uncertain environments without considerable difficulty. The full Morphin algorithm supports updating traversability measures when new data is acquired. Instead of using arbitrarily sampled points from a known terrain mesh, points derived from sensor data could be used. Field D* likewise supports efficient replanning when costs of grid cells are updated. It has been shown to be at least two orders of magnitude faster than repeated application of A* in this respect.

algorithm (described in "Motion planning for a six-legged lunar robot", Hauser, Bretl, Latombe and Wilcox).

9. Acknowledgments

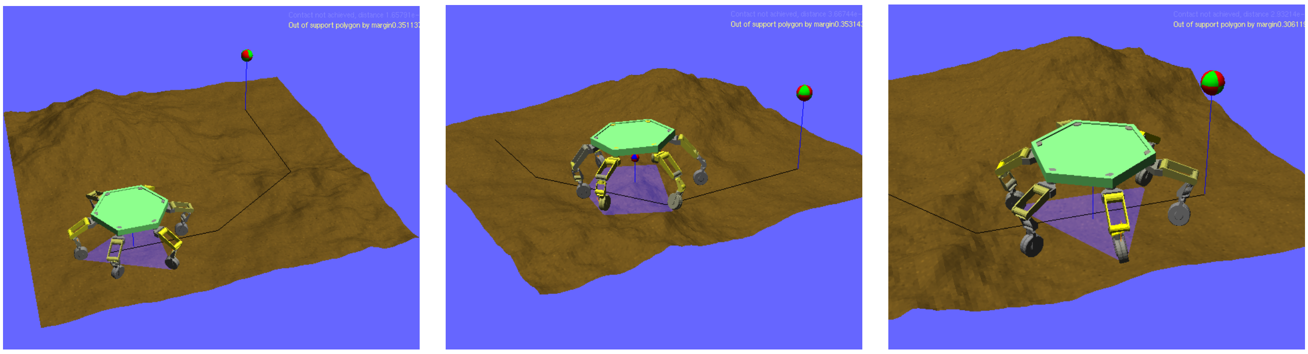
I would like to thank Kris Hauser for his robotics simulation software, which was invaluable to this research. I would also like to thank NASA Ames Research Center and JPL for the opportunity to work with ATHLETE, and Jean-Claude Latombe for his comments and feedback.

References

[1] D. Ferguson and A. Stentz, "Field D*: An Interpolation-based Path Planner and Replanner," Proc. International Symposium on Robotics Research, October 2005.

[2] S. Singh et al. "Recent Progress in Local and Global Traversability for Planetary Rovers," Proc. IEEE International Conference on Robotics and Automation, San Francisco, USA, April 2000.

[3] C. Urmson, R. Simmons, I. Nesnas, "A Generic Framework for Robotic Navigation," Proc. IEEE Aerospace Conference, Big Sky Montana, March 2003.



(a) Tripod gait (b) Wave gait (c) Tripod gait

Fig. 5. ATHLETE selects gait based on difficulty of terrain