Learning to Plan

Ian Baldwin
Department of Engineering
Oxford
ib@robots.ox.ac.uk

Paul Newman
Department of Engineering
Oxford
pnewman@robots.ox.ac.uk

1. Introduction

•Many techniques exist for solving motion-planning problems: **Decomposition-based** methods, **potential-field methods**, and **sampling-based approaches**

•We presents a novel way to bias the sampling domain of sampling-based planners by learning from example plans

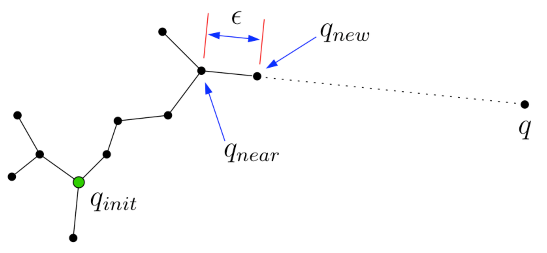
•We introduce the concept of a **Semantic Field**

•We show how the field can be trained using expert data, pruned according to information content and inserted into a regular RRT to produce efficient plans

2. Sampling-based Planning

Rapidly-Exploring Random Trees [RRT]

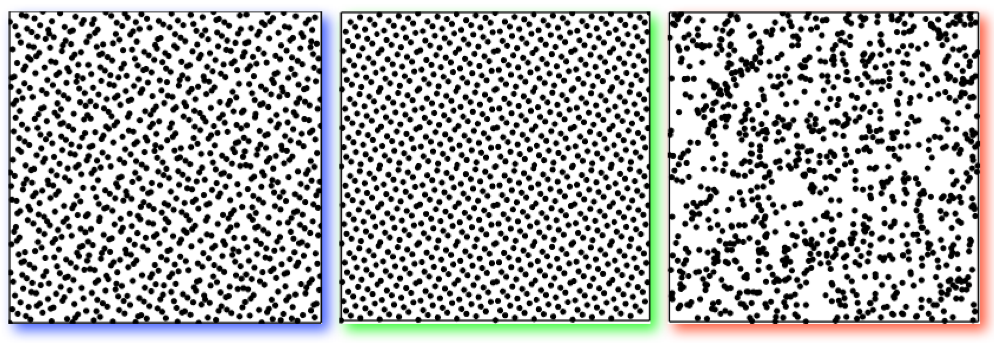
•RRTs[1] are a form of sampling-based planner, used for solving complex path-planning queries[2][3][4]



•The planner generates an approximation of the free-space (C-free) of the environment through **sampling**

•The performance of the planner is highly dependent on the sampling strategy used[5][6]

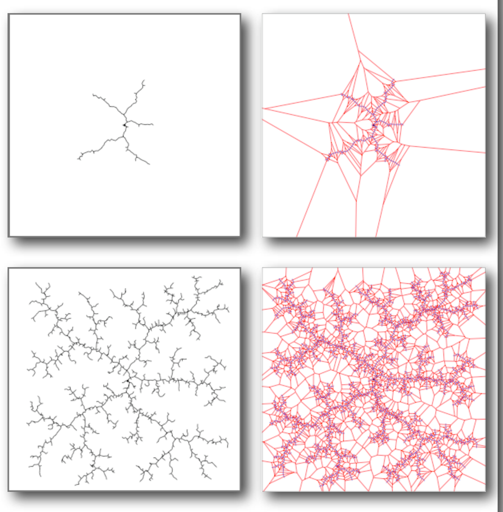
Sampling Strategies



•Sampling strategies can be **deterministic** (Halton or Hammersley points sets), or random

•Deterministic sampling methods generate low-**discrepancy** points[7]

•However, in some cased the **discrepancy** is a poor measure of point equi-distribution, and therefore random sampling strategies (which generalize better to higher dimensions) are used



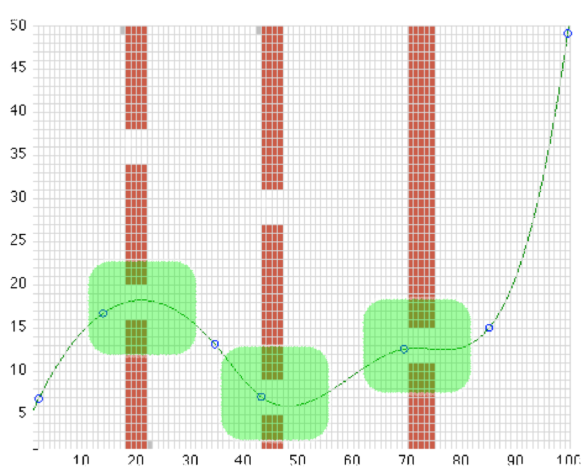
•The sampling strategy employed directly impacts the performance

•We seek to learn **better** sampling distributions based on the trajectories we have observed from vehicles under expert control

3. Semantic Fields

•We assume that a human operator uses semantic knowledge to influence their planning strategies near identifiable objects

•A simple example grid-world is shown in the adjacent figure, with a corresponding "expert path"



• The goal is to maneuver a holonomic point robot from the start position (lower left) to the goal position at the (top right)

•Situated at the centre of each opening in the environment is a **semantic field** (green), with an influence over a neighboring group of cells, and the expert is modeled as a stochastic planner who chooses samples within this region

•The semantic field can be considered to be a random variable with **k** states, where **k** is the number of cells in the discrete field

Constructing the fields

•We observe a set of **N** exemplar paths through the field

•Each path is discretized and each element is considered to be an independent observation, and therefore a state of the random variable

•The distribution that describes the counts of these observations is the Multinomial:

$$\begin{aligned} Mult(s_1, s_2 \ldots \mid \mu, N) &= \frac{N!}{s_1! \cdots s_k!}\mu_1^{s_1}\cdots\mu_k^{s_k} \\ &= \frac{N}{s_1 \ldots s_k}\prod_{k=1}^{k}\mu_k^{s_k} \end{aligned}$$

•This describes the counts over the semantic field cells $S=\{s_1,..,s_k\}$ for N discrete observations with cell probabilities $\mu=\{\mu_1,..,\mu_k\}$

•The goal of the planner during the learning phase is to approximate the multinomial distribution that best describes the collection of exemplar paths shown during the training phase (denoted as $Z=\{z_1,...,z_k\}$

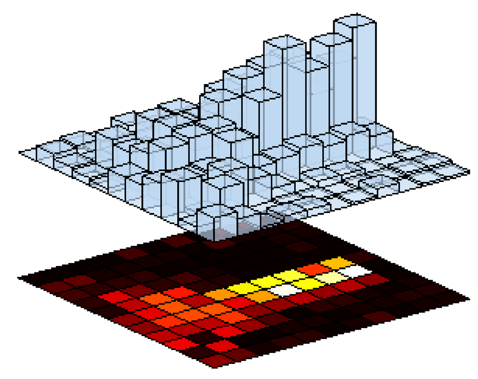
•This can be considered to be a likelihood term, and utilizing the conjugate prior of the Multinomial distribution, the Dirichlet:

$$Dir(\mu_1,\ldots,\mu_K;\alpha_1,\ldots,\alpha_K) = \frac{1}{B(\alpha)}\prod_{k=1}^{K}\mu_k^{\alpha_k - 1}$$

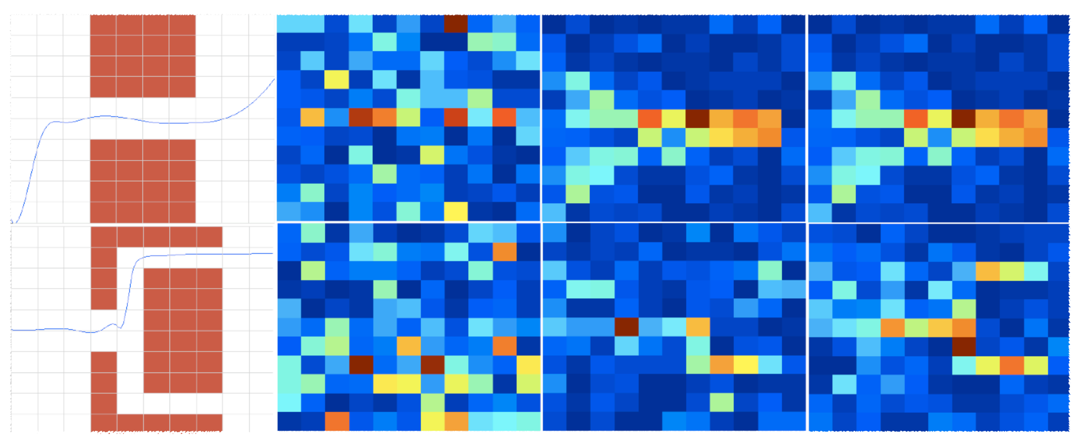
(where $\alpha=\{\alpha_1,...,\alpha_K\}$ are the "shape" parameters, which can be interpreted as "frequency counts" of observed variables), we can form the posterior distribution:

$$p(\mu \mid \mathcal{Z}, \alpha) = Dir(\mu \mid \alpha + \mathcal{Z})$$

•The figure below shows a draw from the posterior distribution, with **1000** samples drawn from a Multinomial distribution parameterized by the posterior:



•The figure below shows two exemplar paths through different "doorway" configurations (also shown is the evolution of samples from the posterior distribution during the learning process):



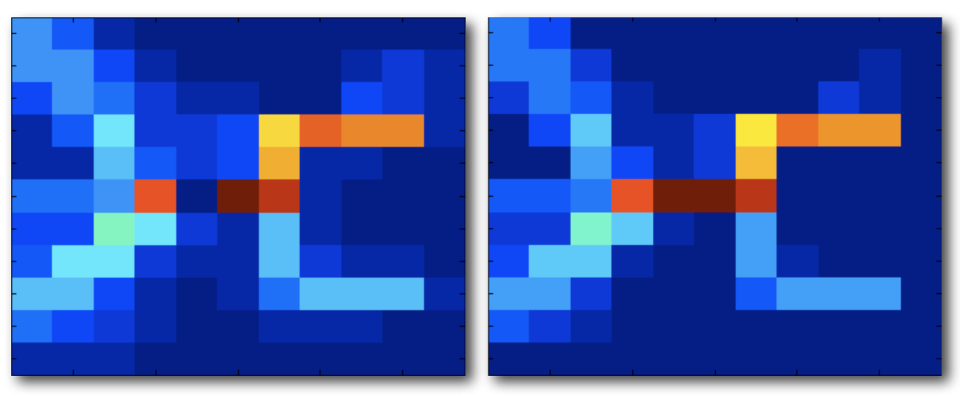
Utilizing Mutual Information

•During the learning phase, the semantic fields were bounded in a predetermined way (for the example grid-world, this was an **11x11** grid), which may not be optimal during planning

•Therefore there is a need to prune the semantic field, in order to definitively specify the region in which there is expected to be a measurable gain from utilizing the learned bias

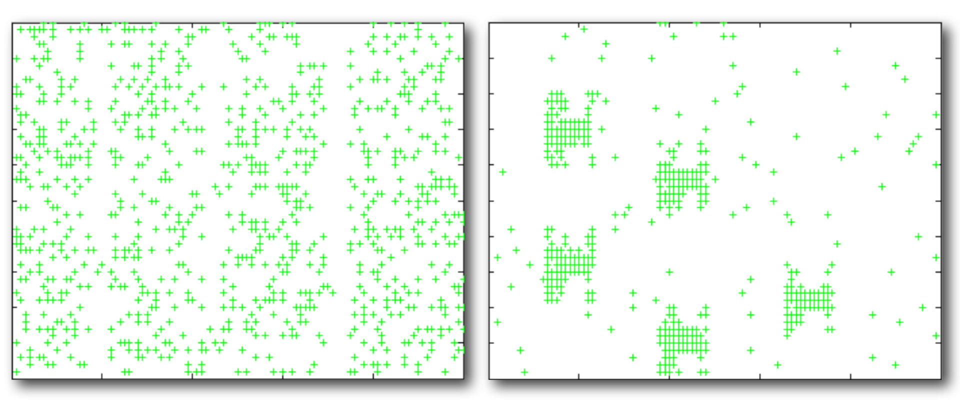
•One way of doing this is to analyze the information content of the field, specifically by evaluating the Mutual Information (**MI**) between adjacent states in the grid:

$$MI(X;Y) = \sum_{y \in Y}\sum_{x \in X} p(x,y) log \frac{p(x,y)}{p(x)p(y)}$$



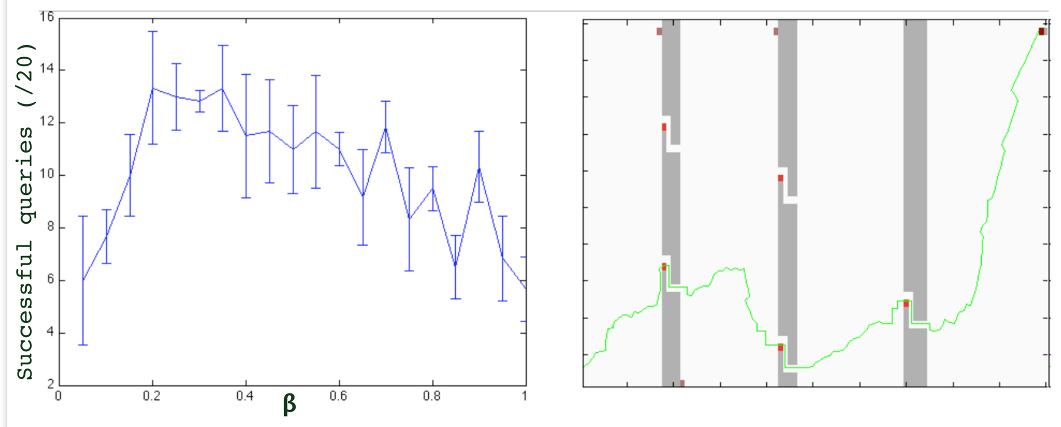
•The trimmed field (right) shows that states with a low MI content relative to their neighbors have been removed

•To ensure that the planner is **probabilistically complete**, a free weighting parameter is introduced between the distributions generated by the semantic fields, and a **uniform distribution**



•The above image shows **1000** samples generated uniformly (left), and with the semantic-field distributions (right)

Planner Performance



•To evaluate the performance of the algorithm, it was necessary to evaluate the weighting parameter (termed β)

•The above figure(left) shows the results of varying β and running multiple queries (for the adjacent grid world)

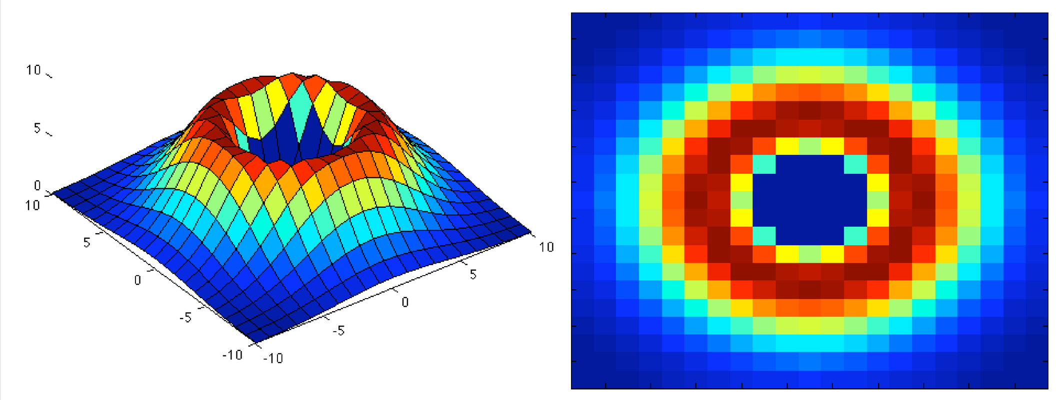
•Extreme weighting values led to poor planning results

•Optimal values lie in the **[0.2;0.4]** range

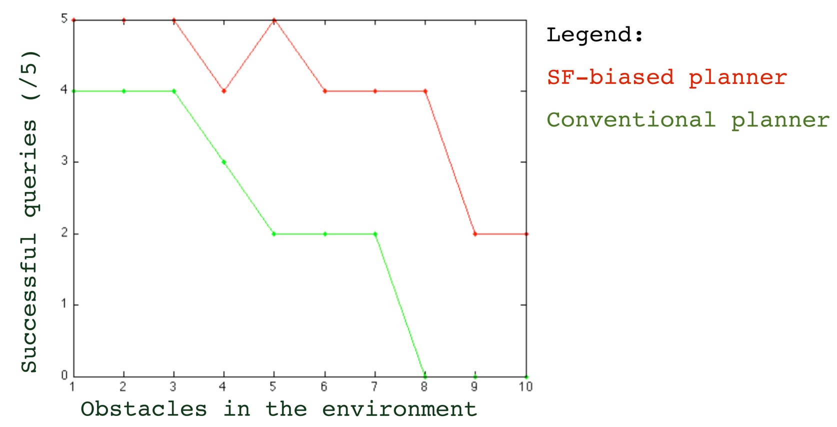
4. Incorporating transients

•To be effective in a realistic environment, transient obstacles must be incorporated

•The figure below shows a prior sampling distribution placed over constant-velocity, constant-heading randomly placed obstacles in the environment (red corresponds to higher probability):



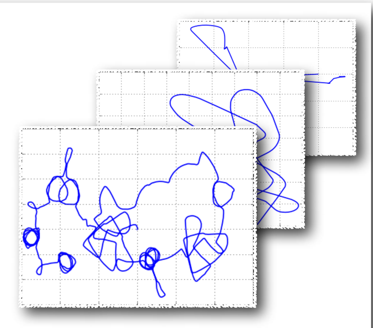
•The planner therefore tends to select samples around the **periphery** of an obstacle



• The above figure shows the performance of the planner for a sequence of queries with increasingly more dynamic objects

• The system is capable of solving planner queries even with a relatively large number of transients

5. Future Work



•Analysis and verification on existing datasets, collected over several years

•This technique can be generalized to more complicated environments, or incorporated into other planner types (for example **Probabilistic Road Maps**[8])

•Further work includes incorporating non-parametric methods into the sampler (for example, Gaussian Processes) to better approximate the distribution over continuous workspaces

•As the SF's represent expert domain knowledge, the concept could be incorporated in graph-based planning algorithms (e.g D*[9])

6. References

[1] Steven M. Lavalle. "Rapidly-exploring random trees: A new tool for path planning". Technical report, 1998.

[2] Marcelo Kallmann, Amaury Aubel, Tolga Abaci, and Daniel Thalmann. "Planning collision-free reaching motions for interactive object manipulation and grasping". Eurographics, 22:313–322, 2003.

[3] Juan Cortes and Thierry Simeon. "Sampling-based motion planning under kinematic loop-closure constraints." In In 6th International Workshop on Algorithmic Foundations of Robotics, pages 59–74. Springer-Verlag, 2004.

[4] Jongwoo Kim and J.P. Ostrowski. "Motion planning a aerial robot using rapidly-exploring random trees with dynamic constraints." In Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on, volume 2, pages 2200–2205 vol.2, Sept. 2003.

[5] Zheng Sun, Ieee Member, David Hsu, Tingting Jiang, Hanna Kurniawati, John H. Reif, and Ieee Fellow. "Narrow passage sampling for probabilistic roadmap planning", 2005.

[6] Shawna L. Thomas, Marco Morales, Xinyu Tang, and Nancy M. Amato. "Biasing samplers to improve motion planning performance." In ICRA, pages 1625–1630, 2007.

[7] Steven M LaValle. "Planning Algorithms." 2004.

[8] Lydia E. Kavraki, Petr Svestka, Jean-Claude Latombe, and Mark H. Overmars. "Probabilistic roadmaps for path planning in high-dimensional configuration spaces." In IEEE International Conference on Methods and Models in Automation and Robotics, 2005, pages 566–580. MorganKaufmann, 1997

[9] Anthony Stentz and Is Carnegle Mellon. "Optimal and efficient path planning for unknown and dynamic environments." International Journal of Robotics and Automation, 10:89–100, 1993.

Sequence showing the effect of increasing the weighting parameter, β