Assignment # 1 — Due: Mon, Nov. 5, 2001

Stanford CS 329 Probabilistic Robotics, Fall 2001
Please submit your answers per Email to thrun@cs.cmu.edu.

Written Assignment #1

1 Question 1: Biasdness of Particle Filters

In class, we discussed in length the fact that Monte Carlo Localization (and particle filters) are biased for finite sample sets, as a result of the way particles are resampled. In this question, you are asked to quantify this bias.
To simplify things, consider a world with 4 possible robot locations: $S = \{s_1, s_2, s_3, s_4\}$:

s_1	s_2
s_3	s_4

Initially, we draw $N \geq 1$ samples uniformly from among those locations—-as usual, it is perfectly acceptable if more than one sample is generated for any of the locations S. Let z now be our first actual sensor measurement. Suppose that z is characterized by the following conditional probabilities:

$$
\begin{aligned}
p(z|s_1) &= 0.8 & \qquad p(\neg z|s_1) &= 0.2 \\
p(z|s_2) &= 0.4 & \qquad p(\neg z|s_2) &= 0.6 \\
p(z|s_3) &= 0.1 & \qquad p(\neg z|s_3) &= 0.9 \\
p(z|s_4) &= 0.1 & \qquad p(\neg z|s_4) &= 0.9
\end{aligned}
$$

As explained in class, these probabilities are used to generate importance factors, which are subsequently normalized and used for resampling. For simplicity, let us assume we only generate one new sample in the resampling process, regardless of N. This sample might correspond to any of the four locations in S. Thus, the sampling process defines a probability distribution over S.
Questions:

1.1 What is the resulting probability distribution over S for this new sample? Answer this question separately for $N = 1, \ldots, 10$, and for $N = \infty$. Your answers have to be exact (truncation errors are acceptable).

1.2 What is the KL divergence between the those probability distributions and the "true" posterior, derived from Bayes filters? Again, answer this question separately for $N = 1, \ldots, 10$, and for $N = \infty$. The KL divergence between a distribution $\hat{p}$ and a "true" distribution p is given by

$$KL(\hat{p}, p) = \sum_i \hat{p}_i \log \frac{\hat{p}_i}{p_i}$$

Again, your answers have to be exact (up to truncation errors).

1.3 Prove the correctness of your answers for $N = 1$, $N = 2$, and $N = \infty$.

1.4 What modification of the problem formulation would guarantee that the specific estimator above is unbiased even for finite values of N? Provide at least two such modifications (each of which should be sufficient).

Hint: I wrote a deterministic *program to calculate some of these results.*

2 Question 2: One-Dimensional Kalman Filters

In class, we defined Bayes filter via the following recursive update equation:

$$p(x_t|u_{1\ldots t}, z_{1\ldots t}) \quad = \quad p(z_t|x_t) \int p(x_t|u_t, x_{t-1})\ p(x_{t-1}|u_{1\ldots t-1}, z_{1\ldots t-1})\ dx_{t-1}$$

Suppose all variables x_t, z_t, and u_t (for all t) are one-dimensional continuous variables in $\Re$. Furthermore, assume that both the motion model $p(x_t|u_t, x_{t-1})$ and the perceptual model $p(z_t|x_t)$ are linear with iid (identically independently distributed) Gaussian noise:

$$\begin{aligned} x_t &= Au_t + Bx_{t-1} + \mu_t \\ z_t &= Cx_t + \nu_t \end{aligned}$$

where μ_t and ν_t are iid Gaussian random variables with zero mean and variance σ, that is, all μ_t and ν_t are distributed according to $N(0; \sigma)$ for some fixed σ.
Furthermore, let us assume that the posterior $p(x_{t-1}|u_{1\ldots t-1}, z_{1\ldots t-1})$ at time $t-1$ is also normal distributed, with mean m and variance ρ.
Questions:

1. Prove that the posterior $p(x_t|u_{1\ldots t}, z_{1\ldots t})$ is also a Gaussian.

2. Derive the exact equations for calculating this Gaussian.

Please do not use matrix notation in any of these tasks—this is not necessary. *Hint 1: It might make sense to do the second question first. Hint 2: You might consider consulting the literature on Fourier transforms, which provides a simple answer for convolving two Normal distributions.*

3 Question 3: Beyond Probabilities?

The key idea of probabilistic robotics is to maintain probability distributions over unknown quantities such as robot poses and maps. Can you imagine situations where a probability distribution might be insufficient to accurately characterize the state of knowledge? If yes, describe one. If not, argue why no such situation might exist.

4 Question 4: EM for Mapping Forests

In class, we talked about how to use the expectation maximization (EM) algorithm for generating 3D maps from range measurements taken at known poses. In this question, you are asked to derive a similar algorithm, but with two differences:

1. All sensor measurements z_i are in a single plane, i.e., we are back to a two-dimensional mapping problem. Since the robot poses are assumed to be known, it may be convenient to think of z_i as a location in x-y space.
2. All objects in the world are trees with known radius r. Since this is a two-dimensional problem, each tree will show up as a circle of radius r in the final map.
3. For simplicity, let us assume that the number of trees J is known a priori, and that the measurement noise is Gaussian. In particular, there is no need to consider other sources of noise or objects other than trees.

Your questions:

1. Provide (and derive) the generative model.
2. What is the expected log likelihood that is being maximized in EM? Please provide a derivation.
3. Derive all necessary equations for the E-step.
4. Lay out your solution for the M-step. If you can't find a closed form solution, you might provide an algorithm for improving the map (this is known as Generalized EM).

Suggestion: You might want to use or Tom Mitchell's book *Machine Learning* (McGraw Hill 1997) or the paper at

http://www.cs.cmu.edu/∼thrun/papers/thrun.3D-EM.html

as a starting point. Simply follow the outline of the math provided there and modify it to accommodate circular objects with radius r.

Good luck!