Probabilistic Robotics exam, spring 2012

May 30, 2012; 8.15–12.15.

The maximum number of points is 40. To pass the exam (with grade G), you need 20 points. To pass with distinction (with grade VG), you need 30 points. Martin will come by at around 10 o’clock to clarify any questions that need clarification. Allowed tools are pen, paper, and a calculator.

Good luck!

Question 1
4 points

Both SLAM with Rao-Blackwellised particle filters and graph-based SLAM have been covered in this course. For both approaches, give two examples of situations that would “break” the algorithm.

Solution

Examples for particle filters:

1. A state space with more than three or four degrees of freedom (for example, fully three-dimensional poses).

2. When a particle which at some point is assigned a low probability doesn’t get resampled in the next iteration, but would have turned out to represent a correct path and map at a later stage.

3. Updating the filter (resampling) when the robot is stationary and doesn’t get new motion updates. The filter would deteriorate to fill up with copies of one single particle.

Examples for pose-graph SLAM:

1. Front-end inserts an erroneous edge with a too high confidence. (For example, data association between two distinct locations that look very similar, or a scan-registration error.)

2. Failing to add a constraint between parts of the robot trajectory, leaving an unconnected graph.

3. A robot continuously running around a small loop, in which all features are observable from each pose. This would take away the sparsity of the graph, eventually making the back-end (the graph optimisation step) prohibitively slow.

Question 2
5 points

At the end of this exam is a multiple-choice question where each question has exactly two possible answers. Assume that a student knows the correct answer to a proportion $k$ of all the questions and makes a random guess for the remaining questions.

The teacher grading this exam observes that question two is correctly answered ($Z_2 = \text{correct}$) by this student. What was the probability that the student was guessing based on
this observation? Derive the formula for the conditional probability and calculate the actual percentage for \( k = 0.5 \).

**Solution**

Let the random variable \( X_2 \in \{ \text{guessing}, \neg \text{guessing} \} \) denote whether the student is guessing or not for question two.

Using the assumption we have for how many questions the student knows the correct answer to, the prior probability that the student is guessing is 

\[
p(X_2 = \text{guessing}) = 1 - k = 0.5.
\]

The conditional probability for providing a correct answer when guessing is

\[
p(Z_2 = \text{correct} \mid X_2 = \text{guessing}) = 0.5,
\]
given that each question has two possible answers. The probability of a correct answer when the student knows the answer is

\[
p(Z_2 = \text{correct} \mid X_2 = \neg \text{guessing}) = 1.
\]

Now, what is the posterior probability after an observation?

\[
p(X_2 = \text{guessing} \mid Z_2 = \text{correct}) = 
\]

\[
\frac{p(\text{correct} \mid \text{guessing})p(\text{guessing})}{p(\text{correct})} = 
\frac{p(\text{correct} \mid \text{guessing})p(\text{guessing})}{p(\text{correct} | \text{guessing})p(\text{guessing}) + p(\text{correct} \mid \neg \text{guessing})p(\neg \text{guessing})}
\]

This is the form of the conditional probability that was asked for in the question. The actual percentage for \( k = 0.5 \) is

\[
p(X_2 = \text{guessing} \mid Z_2 = \text{correct}) = \frac{0.5 \cdot 0.5}{0.5 \cdot 0.5 + 1 \cdot 0.5} = \frac{1}{3}.
\]

In other words, the estimated probability that this student was guessing was decreased from \( \frac{1}{2} \) to \( \frac{1}{3} \) when observing a correct answer.

This answer assumes a perfect “sensor model” — that is, that the teacher marks the answer as correct if and only if the answer actually is correct ;)

---

**Question 3**

Consider a world with only three possible robot locations: \( X = \{ x_1, x_2, x_3 \} \). Consider a Monte Carlo localisation algorithm which may use \( N \) samples among these locations. Initially, the samples are uniformly distributed over the three locations. (As usual, it is perfectly acceptable if there are less particles than locations.) Let \( Z \) be a Boolean sensor variable characterized by the following probabilities:

\[
\begin{align*}
p(z \mid x_1) &= 0.8 & p(\neg z \mid x_1) &= 0.2 \\
p(z \mid x_2) &= 0.4 & p(\neg z \mid x_2) &= 0.6 \\
p(z \mid x_3) &= 0.1 & p(\neg z \mid x_3) &= 0.9
\end{align*}
\]

In other words, we have a high probability of observing \( Z = z \) at location \( x_1 \), and a high probability of observing \( Z = \neg z \) at location \( x_3 \).

MCL uses these probabilities to generate particle weights, which are subsequently normalised and used in the resampling process. 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 three locations in X. Thus, the sampling process defines a probability distribution over X. With N = ∞ this distribution would be equal to true posterior.

2 points

a) Based on the prior uniform distribution, calculate the true posterior p(x_i | z) for each of the locations X = {x_1, x_2, x_3}.

3 points

b) Assume that you use only two particles: N = 2. There are 3^2 = 9 possible combinations possible for the initial particle set. The following table contains values which could be used to calculate the resampling probability for the new sample. Fill in the missing values, and compare the resulting probability distribution to the answer in question a). Are the distributions the same? In what way is the particle filter with two particles biased? Explain this difference.

| number | sample set | prob. of set | p(z | s) for each sample s | weights | prob. of resampling for each x_i |
|--------|------------|--------------|--------------------------|---------|---------------------------------|
| 1      | x_1, x_1  | 1/9          | 0.8, 0.8                 | 1/2, 1/2 | 1/9 0 0 |
| 2      | x_1, x_2  | 1/9          | ..., 0.4                 | 2/3, 1/3 | 2/27 1/27 0 |
| 3      | x_1, x_3  | 1/9          | ..., 0.1                 | 8/9, ... | ... 0 1/81 |
| 4      | x_2, x_1  | 1/9          | ..., ...                 | ... ... | ... 1/27 0 |
| 5      | x_2, x_2  | 1/9          | ..., ...                 | ... ... | ... ... ... |
| 6      | x_2, x_3  | 1/9          | ..., 4/5, 1/5            | ... ... | ... 1/45 |
| 7      | x_3, x_1  | 1/9          | 0.1, 0.8                 | 1/9, 8/9 | ... ... ... |
| 8      | x_3, x_2  | 1/9          | ..., ...                 | ... ... | ... ... ... |
| 9      | x_3, x_3  | 1/9          | ..., ...                 | ... ... | ... ... ... |
| Σ      |            |              |                          | 0.363+   | ... 1 |

Solution

a) This is another exercise in applying Bayes’ theorem.

\[
p(x_i | z) = \frac{p(z | x_i)p(x_i)}{p(z)} = \frac{p(z | x_i)p(x_i)}{\sum_{i=1}^{3} p(z | x_i)p(x_i)}
\]

\[
p(x_1 | z) = \frac{0.8 \cdot 1/3}{0.8 \cdot 1/3 + 0.4 \cdot 1/3 + 0.1 \cdot 1/3} = \frac{0.267}{0.433} = 0.616
\]

\[
p(x_2 | z) = \frac{0.4 \cdot 1/3}{0.8 \cdot 1/3 + 0.4 \cdot 1/3 + 0.1 \cdot 1/3} = \frac{0.133}{0.433} = 0.308
\]

\[
p(x_3 | z) = \frac{0.1 \cdot 1/3}{0.8 \cdot 1/3 + 0.4 \cdot 1/3 + 0.1 \cdot 1/3} = \frac{0.133}{0.433} = 0.308
\]

b) With a finite particle set, the filter is biased towards the prior distribution.

The full table for N = 2 looks like this:
<table>
<thead>
<tr>
<th>number</th>
<th>sample set</th>
<th>prob. of set</th>
<th>( p(z \mid s) ) for each sample ( s )</th>
<th>weights</th>
<th>prob. of resampling for each ( x_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( x_1, x_1 )</td>
<td>1/9</td>
<td>0.8, 0.8</td>
<td>1/2, 1/2</td>
<td>1/9, 0, 0</td>
</tr>
<tr>
<td>2</td>
<td>( x_1, x_2 )</td>
<td>1/9</td>
<td>0.8, 0.4</td>
<td>2/3, 1/3</td>
<td>2/27, 1/27, 0</td>
</tr>
<tr>
<td>3</td>
<td>( x_1, x_3 )</td>
<td>1/9</td>
<td>0.8, 0.1</td>
<td>8/9, 1/9</td>
<td>8/81, 0, 1/81</td>
</tr>
<tr>
<td>4</td>
<td>( x_2, x_1 )</td>
<td>1/9</td>
<td>0.4, 0.8</td>
<td>1/3, 2/3</td>
<td>2/27, 1/27, 0</td>
</tr>
<tr>
<td>5</td>
<td>( x_2, x_2 )</td>
<td>1/9</td>
<td>0.4, 0.4</td>
<td>1/2, 1/2</td>
<td>0, 1/9, 0</td>
</tr>
<tr>
<td>6</td>
<td>( x_2, x_3 )</td>
<td>1/9</td>
<td>0.4, 0.1</td>
<td>4/5, 1/5</td>
<td>0, 4/45, 1/45</td>
</tr>
<tr>
<td>7</td>
<td>( x_3, x_1 )</td>
<td>1/9</td>
<td>0.1, 0.8</td>
<td>1/9, 8/9</td>
<td>8/81, 0, 1/81</td>
</tr>
<tr>
<td>8</td>
<td>( x_3, x_2 )</td>
<td>1/9</td>
<td>0.1, 0.4</td>
<td>1/5, 4/5</td>
<td>0, 4/45, 1/45</td>
</tr>
<tr>
<td>9</td>
<td>( x_3, x_3 )</td>
<td>1/9</td>
<td>0.1, 0.1</td>
<td>1/2, 1/2</td>
<td>0, 0, 1/9</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.457, 0.363, 0.180</td>
</tr>
</tbody>
</table>

(For an even more limited particle set, \( N = 1 \), the table would look as follows.)

<table>
<thead>
<tr>
<th>number</th>
<th>sample set</th>
<th>prob. of set</th>
<th>( p(z \mid s) ) for each sample ( s )</th>
<th>weights</th>
<th>prob. of resampling for each ( x_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( x_1 )</td>
<td>1/3</td>
<td>0.8</td>
<td>1</td>
<td>1/3, 0, 0</td>
</tr>
<tr>
<td>2</td>
<td>( x_2 )</td>
<td>1/3</td>
<td>0.4</td>
<td>1</td>
<td>0, 1/3, 0</td>
</tr>
<tr>
<td>3</td>
<td>( x_3 )</td>
<td>1/3</td>
<td>0.1</td>
<td>1</td>
<td>0, 0, 1/3</td>
</tr>
<tr>
<td>( \Sigma )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.333, 0.333, 0.333</td>
</tr>
</tbody>
</table>

**Question 4**

5 points

The following figure shows the set of possible positions \((x, y)\) at time \((t + 1)\) of a mobile robot using an odometry-based motion model; i.e., the action \( u \) is given by the odometry information \((d_{rot1}, d_{rot2}, d_{trans})\). The errors in \( d_{rot1} \) and \( d_{trans} \) are assumed to follow zero-mean uniform density functions \( p_{rot1}(a) \), \( p_{trans}(d) \) respectively. Give the analytical expressions of the two density functions!

![Possible resulting positions \((x, y)\) at time \((t + 1)\) and pose \((x, y, \Theta)\) at time \(t\)](image)

**Solution**

The probability density function for the rotation should be

\[
p_{rot1}(a) = \frac{1}{\alpha} \quad \text{for} \quad -\frac{\alpha}{2} \leq a \leq \frac{\alpha}{2},
\]

\[
p_{rot1}(a) = 0 \quad \text{otherwise}.
\]
The function for the translation should be

\[ p_{\text{trans}}(d) = \frac{1}{r_1 - r_2} \]

for \[-r_1 - r_2 \leq d \leq \frac{r_1 - r_2}{2} ,\]

\[ p_{\text{trans}}(d) = 0 \]

otherwise.

---

**Question 5**

4 points

Grid maps, as their name suggests, are grids where the value of each cell represent the probability of that cell being occupied. The probability of each cell is updated using a binary Bayes filter. What are the two main simplifying assumptions used in such a filter to keep the problem tractable (i.e., less computationally complex than when using no assumptions)?

**Solution**

The two main simplifying assumptions are:

- The occupancy of each cell is independent of the other cells given measurement data
- The pose of the robot is fully known

An extra assumption is that

- the map is static (does not change over time)

Any combination of two of the three assumptions above has been counted as a correct answer to the question.

---

**Question 6**

5 points

Explain the different steps of Kalman filters. Divide them into prediction and correction and give an explanation on how the state and uncertainty evolves.

**Solution**

The idea was to check the overall concept of a Kalman Filter and rather not only to give a set of equations (possible only memorized). However, since this was not explicitly said in the question, answers containing only equations are considered OK.

Your answer should include written text (or equations) describing the KF.

---

**Question 7**

5 points

Extend the previous question with respect to EKF localization. What sensor data is typically used and where? What models of uncertainty needs to be provided and where are they used?
Solution

The following topics should be included in your answer to get the full five points:

• Based on equation or in text that linearization is used in the prediction and correction step (2 points)

• Discussion on sensor data — range / bearing / etc. — or that a measurement equation was provided (1 point)

• Models of uncertainty — odometry (control input) covariance matrix in the prediction step (1 point)
  – measurements covariance matrix in the correction step (1 point)

Question 8

True or false? Correct answer is +1 point per question. A false answer results in −1 point, but you cannot get negative total points for this question.

1. If two variables A and B are independent, then they remain independent given knowledge of any other variable C.

2. Bayes filters assume conditional independence of sensor measurements taken at different points in time given the current and all past states.

3. A likelihood-field range-sensor model is more accurate than a beam model.

4. Occupancy grid maps in log-odds form are numerically more stable than in probability form.

5. In certain degenerate cases a particle filter could still work even with a single particle.

6. Given the accumulated 3D rotation error between two nodes a and b in a pose graph, encoded by a matrix $R(x, y, z)$, distributing the rotation error over the n nodes between a and b can accurately be done by applying $R_{i}(\frac{x}{n}, \frac{y}{n}, \frac{z}{n})$ to each node i between a and b.

7. When using the normal-distributions transform (NDT) to represent laser scan data, an advantage of using a mixture of a normal and a uniform distribution (as opposed to just a normal distribution) is that spurious scan points don’t “inflate” the normal distribution unnecessarily.

Solution

1. false

   The fact that $p(A | B) = p(A)$ does not necessarily mean that $p(A | B, C) = p(A)$. For example, if $A$ denotes the outcome of a roll of one die, $B$ denotes the outcome of another die, and $C$ denotes whether the sum of the two dice is odd or even, knowing the outcome of both $B$ and $C$ does change the probability distribution for $A$.

2. true

   This is the Markov assumption.

3. false

   Using likelihood fields is an approach that can be used for speeding up evaluation of the sensor model in a grid map, and make the model smoother, but it is not accurate.
4. true
   Log odds go from minus to plus infinity, probabilities go from 0 to 1 (with 1 being the unstable value).

5. true
   The degenerate case would be the one of fully deterministic motion, and with a known start pose.

6. false
   This is discussed in the paper by Grisetti et al. Three-dimensional distributions are noncommutative, so the interpolation has to be done with something like a slerp instead.

7. true
   Trying to fit a single normal distribution to points in a plane, plus a couple of extraneous points (for example, from a person passing by) would result in a distribution with large covariance, which would not in a meaningful way describe the underlying surface.