Lecturer: Prof. Kai Arras
Lab instructors: Till Hielscher, Dennis Rotondi, Fabio Scaparro
Submission instructions: see box
Submission and voting deadline in ILIAS: Nov. 08, 14:00

Socially Intelligent Robotics Lab
Institute for Artificial Intelligence
Faculty of Computer Science
University of Stuttgart

Submission Instructions: For the solution of theoretical tasks, use a header with your name and Matrikelnummer on each sheet and combine all files (pictures, scans, LaTeX-ed solutions) into a single `.pdf` document. For code solutions, if not stated otherwise, your code should execute correctly when called from a single `.m` or `.mlx` script (external functions are ok as long as they are called from the script). Each file you submit must include a header with your name and Matrikelnummer. Please add comments to make your code readable and to indicate to which task and subtask it refers to. For submission, all files should be included in a single `.zip` archive named as: `Ex02_YourLastname_Matrikelnummer.zip`. Remember to vote on the tasks that you solved and be ready to present them.

Exercise 02: Probability Refresher

Exercise 02.1: Conditional Probability

a) Suppose that we have three colored boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange and 0 limes, and box g contains 3 apples, 3 oranges and 4 limes. If a box is chosen at random with probabilities $p(r) = 0.2$, $p(b) = 0.2$, $p(g) = 0.6$, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box), then what is the probability of selecting an apple?

b) If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

Exercise 02.2: Bayes' Rule

a) Suppose we live in a world with only three possible robot positions $X = (x_1, x_2, x_3)$. Let Z be a Boolean sensor variable characterized by the following probabilities:

$$\begin{array}{ll} p(z|x_1) = 0.8 & p(\neg z|x_1) = 0.2 \\ p(z|x_2) = 0.4 & p(\neg z|x_2) = 0.6 \\ p(z|x_3) = 0.1 & p(\neg z|x_3) = 0.9 \end{array}$$

Suppose that the marginal distribution of the robot position is uniform, $p(x_i) = \frac{1}{3}$. Calculate the posterior $p(x_i|z)$ for each of the locations $X = (x_1, x_2, x_3)$.

Exercise 02.3: Expectation, Variance and Covariance

a) Suppose that the two variables x and z are statistically independent. Show that the mean and variance of their sum satisfies

$$\mathrm{E}[x+z] = \mathrm{E}[x] + \mathrm{E}[z] \tag{1}$$

$$\mathrm{Var}[x+z] = \mathrm{Var}[x] + \mathrm{Var}[z] \tag{2}$$

b) Show that if two variables x and y are independent, then their *covariance* is zero. The covariance is defined as

$$\mathrm{Cov}[x,y] = \mathrm{E}[(x-\mu_x)(y-\mu_y)] \tag{3}$$

c) John computes the *covariance matrix* of a random vector as follows. Bob tells him that he is wrong. Explain why.

$$C = \begin{bmatrix} 9 & 4 \\ 4 & 1 \end{bmatrix} \tag{4}$$

Exercise 02.4: Probability distributions and moments

This exercise requires you to import the provided Matlab file `distributions.mat` into your Matlab workspace. It will then contain a 10000×5 matrix called `data`.

a) Each of the five columns of *data* contains 10.000 values that were randomly sampled from unknown probability distributions. Without using Matlab's built-in functions, calculate the

- mean
- variance
- standard deviation
- 3rd central moment
- 4th central moment

for each column. Instead of using for-loops, use vectorized functions such as `sum` and `repmat`.

b) Now verify your results using the built-in functions `mean`, `var`, `std` and `moment`.

c) Plot each of the five distributions into the same figure using the `plot` command. Use $[-5, 20]$ as the `x`-axis limits and `'.-'` as the line style. You can use the `histc` command with e. g. 100 bins to compute a histogram of the relative frequencies of the values in each column of `data`. Normalize the frequencies such that for each distribution, they sum up to 1.0.

d) The points have been sampled by four different classes of distributions: Gaussian, Binomial, Poisson and Chi Squared. With the help of your previous knowledge or Google, determine for each of the five columns which was the generating distribution.

e) If you suspect any of the distributions to be Gaussian, what were the parameters originally used to build the distribution?

Exercise 02.5: Probabilistic Graphical Models

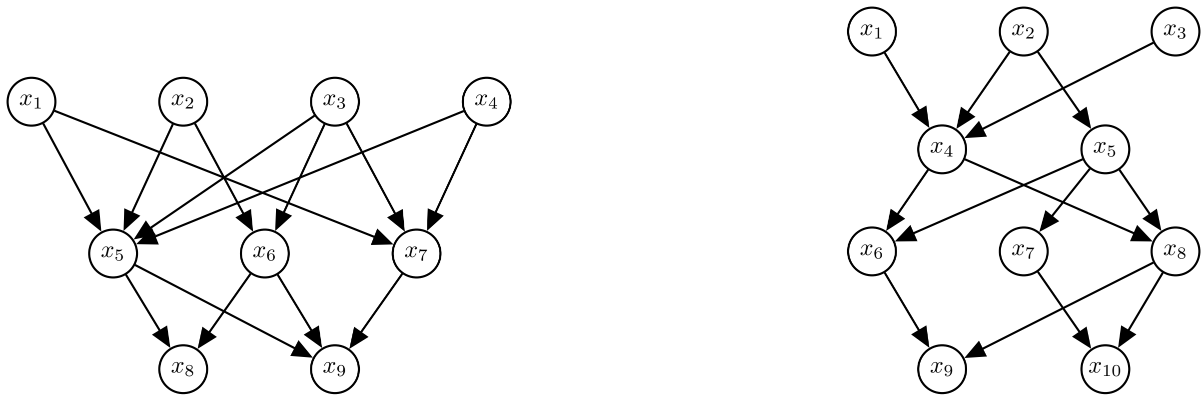

Figure 1: Probabilistic graphical models **a**(left) and **b**(right)

a) Consider a malfunctioning coffee-serving robot that you have to repair. From your experience you can state the following: a loose cable is a possible cause for loss of commands over the bus that controls the robot's arm and is also an explanation for high CPU load of the robot's built-in embedded computer (because, for example, a certain task throws an exception and restarts constantly). In turn, either of these could cause the arm to malfunction and spill coffee. An increased temperature of the robot's PC can also be explained by a high CPU load.
Represent these causal links in a probabilistic graphical model. Let a stand for LOOSE-CABLE, b for LOSS-OF-COMMANDS, c for HIGH-CPU-LOAD, d for SPILLED-COFFEE, and e for INCREASED-PC-TEMPERATURE.

b) From the given directed graphs **a** and **b** shown in Figure 1, find the corresponding distribution. Remember that:

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k) \tag{5}$$