Machine Learning
Exercise 9

Marc Toussaint
Machine Learning & Robotics lab, U Stuttgart
Universitätsstraße 38, 70569 Stuttgart, Germany

July 10, 2014

1 Message passing

Consider three random variables A, B and C with joint distribution $P(A, B, C) = P(A)\ P(B|A)\ P(C|B)$. Let each RV be binary. We assume the CPTs are

$$P(A) = \begin{bmatrix} 1/3 & 2/3 \end{bmatrix}$$
$$P(B|A) = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$$
$$P(C|B) = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$$

a) Draw the Bayes Net for the joint $P(A, B, C)$. Draw also the factor graph that corresponds to $P(A, B, C) = f_1(A)\ f_2(A, B)\ f_3(B, C)$.

b) Compute (by hand on paper) all messages in this factor graph. These are (forward) $\mu_{1\to A}(A)$, $\mu_{2\to B}(B)$, $\mu_{3\to C}(C)$ and (backward) $\mu_{3\to B}(B)$, $\mu_{2\to A}(A)$.

c) Compute the posterior marginals (also called "beliefs") $P(A)$, $P(B)$, and $P(C)$ for each variable.

d) Assume we have an additional factor $f_4(A, C) = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$ that couples A and C. To what messages would loopy belief propagation converge to when we would iterate infinitely? Would it actually converge? Would it converge modulo a scaling of the messages? All these questions can be addressed by investigating the fixed point equations of loopy BP – what is the fixed point equation for, say, $\mu_{4\to A}$? (No numerical answers necessary for these questions.)

2 Sampling from a Gaussian

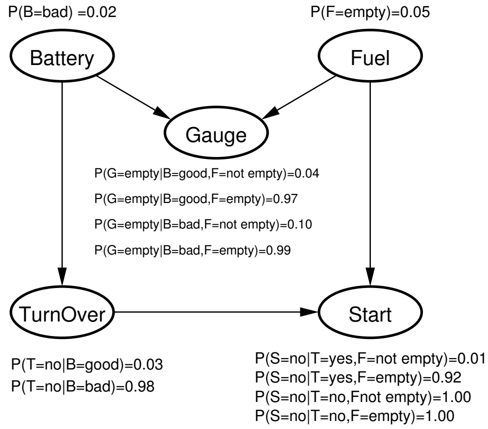
Consider the following simple model: There is a 1-dimensional Gaussian random variable x with $P(x) = \mathcal{N}(x|0, 1)$. There is a binary random variable Y with

$$P(Y\!=\!1\,|\,x) = \begin{cases} .9 & x > 0 \\ .1 & \text{otherwise} \end{cases} .$$

Use rejection sampling to compute a sample set representing the posterior $P(x\,|\,Y = 1)$. From this, compute an estimate of the posterior mean $\int_x x\ P(x\,|\,Y\!=\!1)$.

3 Sampling

Consider again the Bayesian network of binary random variables given below.



a) Condition on Start=no. Implement rejection sampling to collect a sample set $\mathcal{S} \sim P(B, F, G, T \mid S = \text{no})$ with K samples (e.g., $K = 1000$). Compute $P(F \mid S = \text{no})$.

b) Do the same using importance weighting instead of rejection sampling. Compare the results for varying K.