ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Dr. Francis Colas
Institute of Robotics and Intelligent Systems
Autonomous Systems Lab

ETH Zürich
CLA E 26
Tannenstraße 3
8092 Zürich Switzerland

fcolas@mavt.ethz.ch
www.asl.ethz.ch

Information Processing in Robotics
Solution Sheet 2
Topic: Graphical models and Hidden Markov Models

Exercise 1: Bayesian network and independence

(a) $P(A)P(B|A)P(C|A)P(D|B,C)P(E)P(F|D,E)$

(b)
- $B \perp\!\!\!\perp C \mid A$: true: 2 paths, both blocked;
- $B \perp\!\!\!\perp C \mid D$: false: neither paths are blocked;
- $B \perp\!\!\!\perp C \mid F$: false: neither paths are blocked;
- $B \perp\!\!\!\perp C \mid E$: false: path through A not blocked but path through D is;
- $A \perp\!\!\!\perp D \mid B$: false: path through C not blocked;
- $A \perp\!\!\!\perp D \mid B, C$: true: both paths blocked;
- $A \perp\!\!\!\perp E$: true: both paths blocked in F;
- $A \perp\!\!\!\perp E \mid D$: true: both paths blocked in F;
- $A \perp\!\!\!\perp E \mid F$: false: both paths unblocked by F;
- $A \perp\!\!\!\perp E \mid D, F$: false: both paths unblocked by F.

Exercise 2: Hidden Markov Models

We've seen several different uses of HMM in the class each with its own inference. Among those, we presented the forward algorithm that can be summarized by:

$$\begin{aligned} \alpha(S^t) &= P(O^0, \ldots, O^t, S^t), \\ \alpha(S^{t+1}) &= P(O^{t+1}|S^{t+1}) \sum_{S^t} P(S^{t+1}|S^t)\alpha(S^t). \end{aligned}$$

(a) $\alpha(S^{t+1}) = B(O^{t+1})' \circ A \times \alpha(S^t)$ where $B(O^{t+1})$ is the vector row in B corresponding to observation O^{t+1} and $\circ$ is the elemnetwise product of vectors.

(b) The forward algorithm can be restated:

$\alpha \leftarrow \pi$
for all $t \in [0; T]$ **do**
  $BO \leftarrow B[O^t]$
  $\alpha \leftarrow BO' \circ A \times \alpha$
end for

(c) See code.

Exercise 3: Wifi localization

(a) $\pi = (1/6, 1/6, 1/6, 1/6, 1/6m1/6)'$

(b) $A = \begin{pmatrix} .5 & 0 & .1 & 0 & 0 & 0 \\ 0 & .5 & .1 & 0 & 0 & 0 \\ .5 & .5 & .5 & .5 & .5 & .5 \\ 0 & 0 & .1 & .5 & 0 & 0 \\ 0 & 0 & .1 & 0 & .5 & 0 \\ 0 & 0 & .1 & 0 & 0 & .5 \end{pmatrix}$

(c) $B = \begin{pmatrix} .6 & 0 & .3 & .9 & .05 & 0 \\ .4 & .4 & .4 & .1 & .8 & .15 \\ 0 & .6 & .3 & 0 & .15 & .85 \end{pmatrix}$

(d) see code

(e) α values gets smaller and smaller as their sum is the probability of the sequence of observation. Fixed precision can be an issue as the numbers get smaller and smaller. Renormalizing the vector while keeping the norm can improve precision.

If we continue observing 2, the distribution will converge to a distribution expressing that E is the most probable before C. In that case, it behaves like a Markov Chain and converges to the fixed point of the update.