Informatics 2D. Tutorial 8
Probabilities and Bayesian Networks

Mihai Dobre, Alex Lascarides

Week 9

Notation

We follow the notation in Russell and Norvig 13.2. Random variables begin with an uppercase letter. Members of a random variable's domain (the possible values it can take on) start with a lowercase letter. For example, the random variable A can take on values $\{true,\ false\}$. By convention, the proposition of the form $A = true$ is abbreviated as a, while $A = false$ is abbreviated as $\neg a$.

Part 1: Basic Probability

In a model used in a diagnostic system of a car, there are three random variables; *Battery* that can be charged or dead, *Radio* that can work or not, and $Ignition$, which describes whether the car starts or not.

If the battery is dead, the car usually does not start. On the other hand, the radio may initially work if the battery is nearly dead but contains enough charge for the radio.

1. Given the probabilities in Figure 1 what is the probability of the following: (pay attention to the use of upper/lowercase)
   - $P(battery)$
   - $P(Radio)$
   - $P(radio \land \neg ignition)$
   - $P(\neg radio \lor \neg ignition)$
2. Derive Bayes rule from the Product Rule (see Figure 2)
3. Using the product rule in Figure 2 and the full joint distribution defined in Figure 1, compute:
$$P(Ignition \mid \neg radio)$$
4. Suppose that we know that *Battery* is the cause for $Ignition$ and also the cause of *Radio*, and that *Radio* is conditionally independent of $Ignition$ given *Battery*, what is the easiest way to write the full joint probability distribution:
$$P(Battery,\ Radio,\ Ignition)$$

	$battery$		$\neg battery$	
	$radio$	$\neg radio$	$radio$	$\neg radio$
$ignition$	0.7	0.085	0.002	0.003
$\neg ignition$	0.05	0.03	0.03	0.1

Figure 1: A full joint distribution for the *Battery*, *Radio*, and *Ignition* world.

Axioms

- $0 \leq P(a) \leq 1$
- $P(true) = 1$, $p(false) = 0$
- $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Properties derived from Axioms

- $\sum_{i=1}^{n} P(D = d_i) = 1$

Product Rule

- $P(a \wedge b) = P(a|b)\,P(b) = P(b|a)\,P(a)$

Bayes' Rule

- $P(a|b) = \dfrac{P(b|a)\,P(a)}{P(b)}$

Figure 2: Useful Probability Rules

Solutions

Note that the probabilities given in the the full joint distribution table sum up to 1, as required by the axioms of probability:

$$0.7 + 0.085 + 0.002 + 0.003 + 0.05 + 0.03 + 0.03 + 0.1 = 1$$

1. The probabilities are as follows:
   - $P(battery) = 0.7 + 0.085 + 0.05 + 0.03 = 0.865$
   - $P(Radio) = < P(Radio = true), \; P(Radio = false) >$ which are computed separately:

     $$P(Radio = true) = 0.7 + 0.05 + 0.002 + 0.03 = 0.782$$

     $$P(Radio = false) = 0.085 + 0.03 + 0.003 + 0.1 = 0.218$$

     Or use $P(Radio = false) = 1 - P(Radio = true)$.
   - $P(radio \wedge \neg ignition) = 0.05 + 0.03 = 0.08$
   - $P(\neg radio \vee \neg ignition) = 0.085+0.03+0.003+0.1+0.05+0.03 = 0.298$

2. We can directly derive Bayes rule from the product rule:

   $$P(a|b) \; P(b) = P(b|a)P(a) \;\; \Rightarrow P(a|b) = \frac{P(b|a)P(a)}{P(b)}$$

3. To compute $P(Ignition|\neg radio)$ we need to compute $P(ignition|\neg radio)$ and $P(\neg ignition|\neg radio)$ separately:

   $$P(ignition|\neg radio) = \frac{P(ignition \wedge \neg radio)}{P(\neg radio)}$$

   $$= \frac{0.085 + 0.003}{0.085 + 0.03 + 0.003 + 0.1} = 0.4037$$

   $$P(\neg ignition|\neg radio) = \frac{P(\neg ignition \wedge \neg radio)}{P(\neg radio)}$$

   $$= \frac{0.03 + 0.01}{0.085 + 0.03 + 0.003 + 0.1} = 0.5963$$

4. $P(Battery, Ignition, Radio) = P(Battery) \times P(Ignition|Battery) \times P(Radio|Battery)$

Part 2: Bayesian Networks

In a different model of the car, the alternator (A) can stop working due to an electric fault (E) or due to the breaking of the drive belt (D). The failure of the alternator causes complete discharge of the battery (B) that supplies current to the radio (R) and lights (L). The battery, the lights and the radio may also stop working for internal reasons.

$P(d) = 0.75$ $P(e) = 0.95$

d	e	$P(a)$
T	T	0.9
T	F	0.3
F	T	0.1
F	F	0.1

a	$P(b)$
T	0.8
F	0.2

b	$P(l)$
T	0.9
F	0.05

b	$P(r)$
T	0.8
F	0.1

Figure 3: Probabilities for the second car model (Part 2)

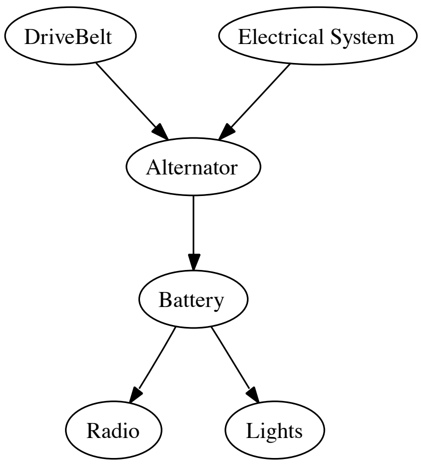

Figure 4: The Bayesian network

1. Draw the Bayesian network that represents the model of the car, showing the variables and the dependence/independence relationships between them.

2. Use the obtained network and the probabilities listed in Figure 3 to compute the probability of:

   - $P(d,\ e,\ a,\ b,\ \neg r,\ \neg l)$
   - $P(\neg d,\ e,\ \neg a,\ b,\ r,\ l)$

Solutions

1. The Bayesian network is shown in Figure 4

2. The joint distribution from the Bayesian network is:
   $P(E, A, D, B, R, L) = P(E) \times P(D) \times P(A|E, D) \times P(B|A) \times P(R|B) \times P(L|B)$

```
function ENUMERATION-ASK(X, e, bn) returns a distribution over X
  inputs: X, the query variable
          e, observed values for variables E
          bn, a Bayes net with variables {X} ∪ E ∪ Y   /* Y = hidden variables */

  Q(X) ← a distribution over X, initially empty
  for each value x_i of X do
      Q(x_i) ← ENUMERATE-ALL(bn.VARS, e_{x_i})
          where e_{x_i} is e extended with X = x_i
  return NORMALIZE(Q(X))
----------------------------------------------------------------------
function ENUMERATE-ALL(vars, e) returns a real number
  if EMPTY?(vars) then return 1.0
  Y ← FIRST(vars)
  if Y has value y in e
      then return P(y | parents(Y)) × ENUMERATE-ALL(REST(vars), e)
      else return Σ_y P(y | parents(Y)) × ENUMERATE-ALL(REST(vars), e_y)
          where e_y is e extended with Y = y
```

Figure 14.9 The enumeration algorithm for answering queries on Bayesian networks.

Figure 5: The enumeration algorithm

- $P(d,\ e,\ a,\ b,\ \neg r,\ \neg l) = 0.75 \times 0.95 \times 0.9 \times 0.8 \times (1-0.8) \times (1-0.9) = 0.01026$
- $P(\neg d,\ e,\ \neg a,\ b,\ r,\ l) = (1-0.75) \times 0.95 \times (1-0.1) \times 0.2 \times 0.8 \times 0.9 = 0.03078$

Part 3: Exact Inference in Bayesian Networks

To make a probability inference query means to compute the posterior probability distribution for a set of query variables given some observed event. X denotes the query variable, $\mathbf{E}$ denotes the set of evidence variables $E_1, ..., E_n$, $\mathbf{Y} = Y_1 \ldots Y_l$ denotes the non-evidence variables, also called hidden variables.

Conditional probability can be computed by summing terms from the full joint distribution:

$$P(X|\mathbf{e}) = \alpha P(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} P(X, \mathbf{e}, \mathbf{y})$$

where α is the normalisation factor.

Using the enumeration algorithm in Figure 5, Compute the conditional probability of:

$$P(D \mid A = true)$$

Solutions

$$P(D|a) \ = < P(D = true|a),\ P(D = false|a) >$$

We'll compute them separately:

- First, $P(D = true \mid a)$
$= P(d|a) = \alpha P(d, a) = \alpha \Sigma_E P(d, a, E)$
$= \alpha \Sigma_E P(d) P(a|d, E) P(E) = \alpha P(d) \Sigma_E P(a|d, E) P(E)$
$= \alpha P(d)[(P(a|d, e)P(e)) + (P(a|d, \neg e)P(\neg e))]$
$= 0.75[(0.9 \times 0.95) + (0.3 \times 0.05)] = \alpha 0.6525$

- Second, $P(D = false \mid a)$
$= P(\neg d|a) = \alpha P(\neg d, a) = \alpha \Sigma_E P(\neg d, a, E)$
$= \alpha \Sigma_E P(\neg d) P(a|\neg d, E) P(E) = \alpha P(\neg d) \Sigma_E P(a|\neg d, E) P(E)$
$= \alpha P(\neg d)[(P(a|\neg d, e)P(e)) + (P(a|\neg d, \neg e)P(\neg e))]$
$= \alpha 0.25[(0.1 \times 0.95) + (0.1 \times 0.05)] = \alpha 0.025$

where the normalisation factor α is computed from $1 = P(d|a) + P(\neg d|a) = \alpha 0.025 + \alpha 0.6525$, so:

- $P(d|a) = 0.6525 \times \frac{1}{0.6525+0.025} = 0.963$
- $P(\neg d|a) = 0.025 \times \frac{1}{0.6525+0.025} = 0.036$