Universität des Saarlandes
Fachrichtung 6.2 – Informatik
J. Siekmann, S. Autexier, K. Fischer, E. Melis, C. Zinn
M. Schiller, E. Schulz, R.Vollmann

12th Theoretical Assignment in Artificial Intelligence (WS 2006/2007)

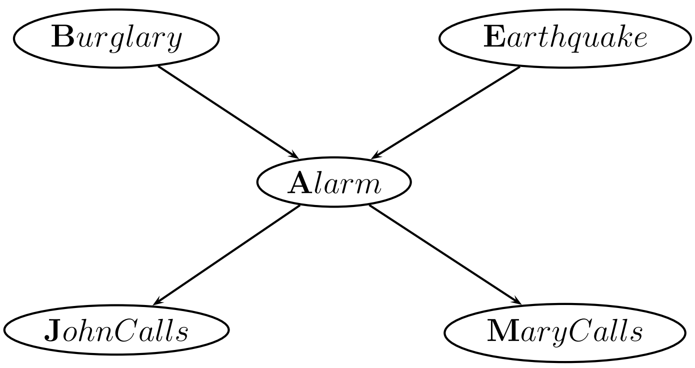
Issued: February 1st, 2007

Exercise 12.1 Consider the following set of training examples:

Instance	Classification	a_1	a_2
1	+	T	T
2	+	T	T
3	−	T	F
4	+	F	F
5	−	F	T
6	−	F	T

1. What is the entropy of this collection of training examples with respect to the target function classification?
2. What is the information gain of a_2 relative to these training examples?
3. What is the information gain of a_1 relative to these training examples?
4. Construct a decision tree with the decision-tree-learning algorithm from the lecture for the given set of training examples! Use information gain as the criterion for choosing the "best" attributes for splitting the tree!
5. Assume that *Instance 6* is an error with your measurement, and should in fact be classified as a + and not as a −. Does this change your decision tree? How?

Exercise 12.2 Consider the example from exercise 11.4 on the last exercise sheet.



The *a priory* probabilities and conditional probabilities are:

P(B)
.001

P(E)
.002

B	**E**	**P(A)**
T	*T*	.95
T	*F*	.94
F	*T*	.29
F	*F*	.001

A	**P(J)**
T	.90
F	.05

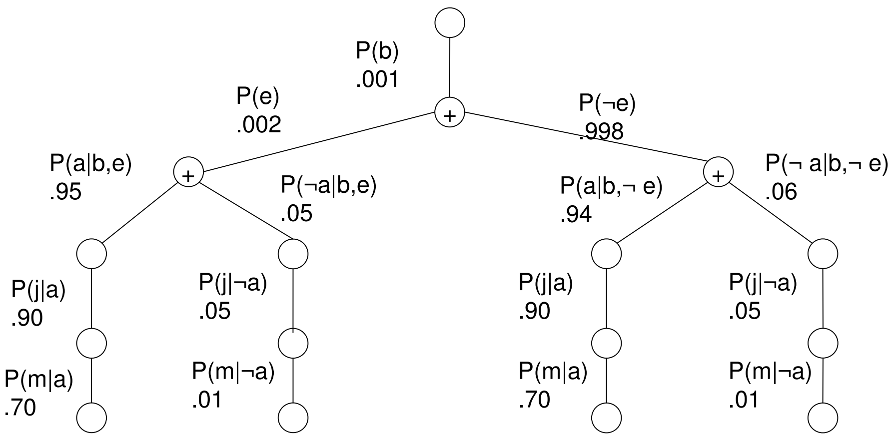
A	**P(M)**
T	.70
F	.01

In the lecture, we have seen the ENUMERATION-ASK algorithm (reproduced from Russell/Norvig: AIMA, 2003):

```
function ENUMERATION-ASK(X,e,bn) returns a distribution over X
    inputs:  X, the query variable
             e, observed values for variables E
             bn, a Bayesian network with variables {X}∪E∪Y

    Q(X) ← a distribution over X, initially empty
    for each value x_i of X do
        extend e with value x_i for X
        Q(x_i) ← ENUMERATE-ALL(VARS[bn],e)
    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
```

1. Consider the query $P(B|j,m)$. When calling ENUMERATION-ASK (where $X := B$, $e := \{j,m\}$, bn is the given network), ENUMERATE-ALL is called to compute $Q(b)$. This evaluation will produce a tree, which is reproduced below.

   Use the ENUMERATE-ALL algorithm step by step to evaluate ENUMERATE-ALL({B,E,A,J,M},{j,m,b}) – i.e. follow the steps of the algorithm that produce the given tree!



2. Use ENUMERATION-ASK to answer the query $P(B|j,m)$! You can use the solution from part 1., and then do the missing parts.

   One note about normalization: Remember that when you calculate the distribution of some variable, lets say D, where D can take the values d and $\neg d$, for example, then the probabilities of the possible values of D need to add up to 1. Furthermore, $P(d|e_1,...,e_n) + P(\neg d|e_1,...,e_n) = 1$. If the computation involves an unknown normalization constant α, such that $P(d|e_1,...,e_n) = \frac{x}{\alpha}$, and $P(\neg d|e_1,...,e_n) = \frac{y}{\alpha}$, α can be computed easily by knowing that $P(d|e_1,...,e_n) = P(\neg d|e_1,...,e_n) = 1$. This value α is what NORMALIZE is supposed to determine. The result of normalization is the computed probability distribution multiplied by α (look also at p. 476 and p. 505 in Russell/Norvig).