GWV - Grundlagen der Wissensverarbeitung

Blatt 8

Julian Tobergte, Melanie Budde,
Maximilian Bauregger, Mohammad Oslani

4. Dezember 2015

Exercise 8.2 Language Modelling

1.

Programm im Anhang.

2.

Nehmen wir z.B. den folgenden generierten Satz: „Linux gründende neue Herausforderung nicht mehr als Original-Sun-Speicherriegel. in diesen Ankündigungen gibt es auch mit der darauf, der.“ Grammatikalisch ist hier einiges im Argen. Es existieren sinnvolle Wortkombinationen, wie „nicht mehr als“ oder „in diesen Ankündigungen“. Eine Satzstruktur oder gar Sinn ist nicht vorhanden. Dies ergibt sich, da nur direkt benachbarte Worte miteinander verknüpft sind, Grammatik sich aber über den gesamten Satz erstreckt.

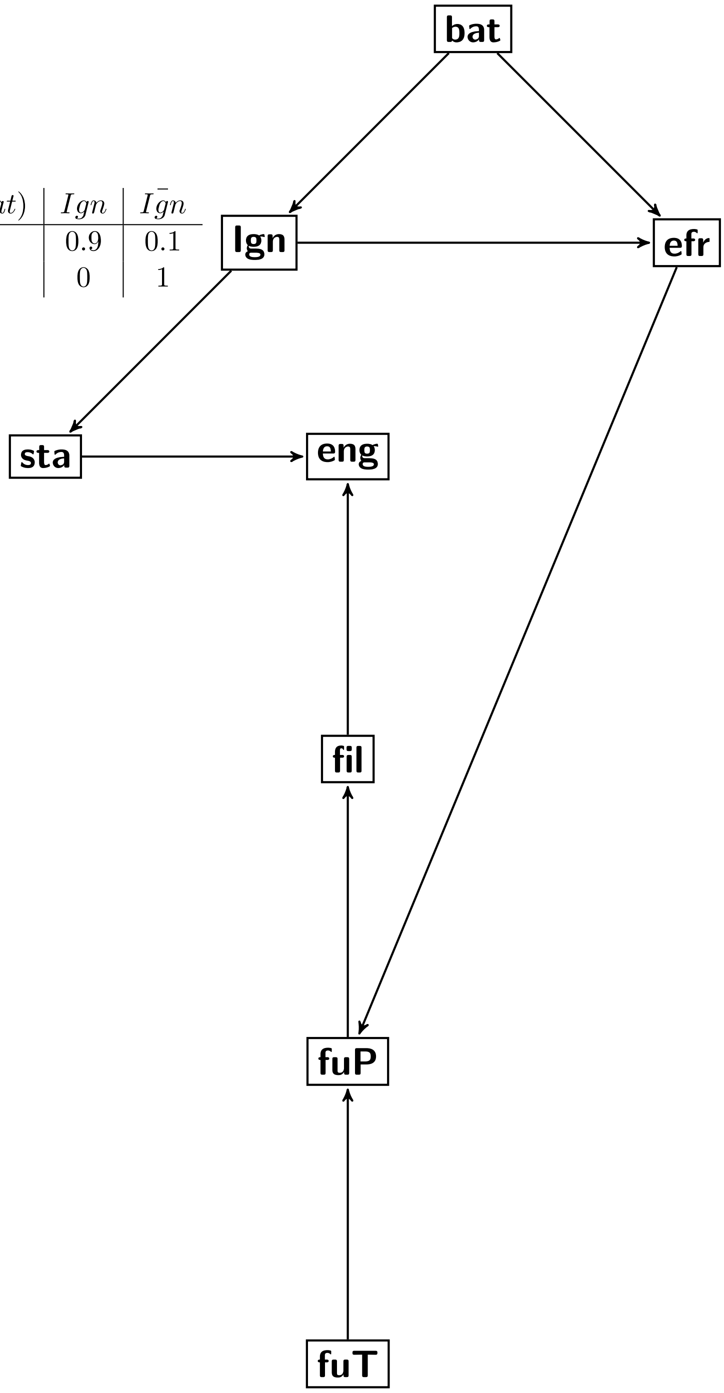
Exercise 8.3 Diagnosis (cont.)

Bayessches Netz
Es sind nur 2 Tabellen eingezeichnet, da sich der Rest genau gleich verhält.

$P(bat) = 0.9$

$P(Ign \mid bat)$	Ign	$\overline{Ign}$
bat	0.9	0.1
$\overline{bat}$	0	1

$P(Efr \mid bat, Ign)$	Efr	$\overline{Efr}$
bat, Ign	0.9	0.1
$bat, \overline{Ign}$	0	1
$\overline{bat}, Ign$	0	1
$\overline{bat}, \overline{Ign}$	0	1



Zur Berechnung benötigt man den Satz der totalen Wahrscheinlichkeit.

Hierbei sind viele Summanden 0, daher ausgelassen.

$$\begin{aligned}
P(bat) &= 0.9 \\
P(sta) &= 0.9 \cdot P(Ign) = 0.9 \cdot 0.9 \cdot 0.9 \\
&= 0.729 \\
P(eng) &= 0.9 \cdot P(sta) \cdot P(fil) = 0.9 \cdot 0.729 \cdot 0.9 \cdot 0.9 \cdot 0.9 \cdot 0.9 \cdot 0.9 \\
&\approx 0.387 \\
P(eng|fuP) &= 0.9 \cdot P(sta) \cdot 0.9 \cdot 1 \cdot 0.9 \cdot 0.9 \cdot 0.9 \\
&\approx 0.430
\end{aligned}$$

Exercise 8.4 Bayesian Probabilities

Variablen

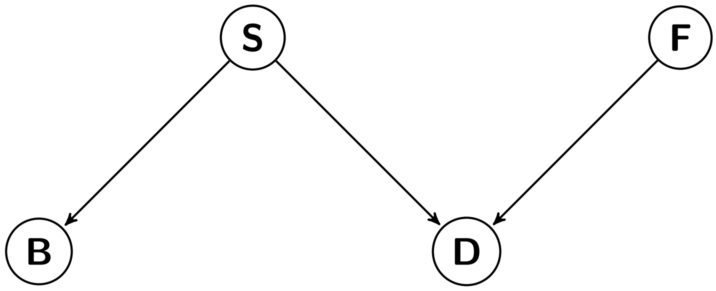
$$\begin{aligned}
&S := \text{Schmuggler} && (s = \text{ist Schmuggler}, \bar{s} = \text{ist kein Schmuggler}) \\
&B := \text{Bellen des Hundes} && (b = \text{Hund bellt}, \bar{b} = \text{Hund bellt nicht}) \\
&F := \text{Fieber} && (f = \text{hat Fieber}, \bar{f} = \text{hat kein Fieber}) \\
&D := \text{Schwitzen} && (d = \text{Person schwitzt}, \bar{d} = \text{Person schwitzt nicht})
\end{aligned}$$

Gegebene Wahrscheinlichkeiten

$$\begin{aligned}
&P(s) = 0.01 && \Rightarrow P(\bar{s}) = 0.99 \\
&P(b|s) = 0.8 && \Rightarrow P(\bar{b}|s) = 0.2 \\
&P(b|\bar{s}) = 0.05 && \Rightarrow P(\bar{b}|\bar{s}) = 0.95 \\
&P(d|\bar{s}, \bar{f}) = 0 && \Rightarrow P(\bar{d}|\bar{s}, \bar{f}) = 1 \\
&P(d|s, \bar{f}) = 0.4 && \Rightarrow P(\bar{d}|s, \bar{f}) = 0.6 \\
&P(d|\bar{s}, f) = 0.6 && \Rightarrow P(\bar{d}|\bar{s}, f) = 0.4 \\
&P(d|s, f) = 0.8 && \Rightarrow P(\bar{d}|s, f) = 0.2 \\
&P(f) = 0.013 && \Rightarrow P(\bar{f}) = 0.987
\end{aligned}$$

Bayessches Netz

$P(s) = 0.01$ $P(f) = 0.013$



$P(B\|S)$	b	$\bar{b}$
s	$0,8$	0.2
$\bar{s}$	0.05	0.95

$P(D\|S,F)$	d	$\bar{d}$
s, f	$0,8$	0.2
$s, \bar{f}$	0.4	0.6
$\bar{s}, f$	0.6	0.4
$\bar{s}, \bar{f}$	0	1

Formeln

Gesetz der totalen Wahrscheinlichkeiten:

$$P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$$

Verallgemeinert:

$$\sum_{j=1}^{\infty} P(A|B_j) \cdot P(B_j) = P(A)$$

Satz von Bayes:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

Aufgaben

- Explaining Away

  In dem Falle, dass eine Person ein Schmuggler ist, bellt der Hund sehr wahrscheinlich. Die Erklärung dafür, dass der Hund bellt, ist also wahrscheinlich, dass die Person ein Schmuggler ist.

- The probability that a person is a smuggler given the observation that the drug dog is barking:

Gesucht wird der Wert für $P(s|b)$.

$$
\begin{aligned}
P(b|s) &= 0.8 \\
P(s) &= 0.01 \\
P(b) &= P(b|s) \cdot P(s) + P(b|\bar{s}) \cdot P(\bar{s}) \\
&= 0.8 \cdot 0.01 + 0.05 \cdot 0.99 \\
&= 0.0575
\end{aligned}
$$

$$
\begin{aligned}
P(s|b) &= \frac{P(b|s) \cdot P(s)}{P(b)} \\
&= \frac{0.8 \cdot 0.01}{0.0575} \\
&\approx 0.13913
\end{aligned}
$$

- The probability that a suspect is sweating (without any prior observation):

  Gesucht wird der Wert für $P(d)$.
  Mit Hilfe des Satzes der Totalen Wahrscheinlichkeit ist dieser Wert wie folgt zu berechnen.

$$
\begin{aligned}
P(d) = &P(d|s,f) \cdot P(s) \cdot P(f) + P(d|s,\bar{f}) \cdot P(s) \cdot P(\bar{f}) \\
&+ P(d|\bar{s},\bar{f}) \cdot P(\bar{s}) \cdot P(\bar{f}) + P(d|\bar{s},f) \cdot P(\bar{s}) \cdot P(f) \\
P(d) = &0.8 \cdot 0.01 \cdot 0.013 + 0.4 \cdot 0.01 \cdot 0.987 + 0 + 0.6 \cdot 0.99 \cdot 0.013 \\
\approx &0.011774
\end{aligned}
$$

- The probability that a person is a smuggler given both the observations that that person is sweating and that the drug dog barked at him or her:

  Gesucht wird der Wert für $P(s|b,d)$.
  Dafür müssen mehrere Zwischenergebnisse benutzt werden.

$$
\begin{aligned}
P(s|b,d) &= \frac{P(b,d|s) \cdot P(s)}{P(b,d)} \\
P(b,d|s) &= P(b|s) \cdot P(d|s) \\
P(d|s) &= P(d|s,\bar{f}) \cdot P(d|s,f) \\
&= 0.4 \cdot 0.8 = 0.32 \\
P(b,d|s) &= 0.8 \cdot 0.32 \\
P(b) &= P(b|s) \cdot P(s) + P(b|\bar{s}) \cdot P(\bar{s}) \\
&= 0.0575 \\
P(s|b,d) &= \frac{0.8 \cdot 0.32 \cdot 0.01}{0.011774 \cdot 0.0575} \\
&= 378
\end{aligned}
$$