Einführung in die
Künstliche Intelligenz
SS09 - Prof. Dr. J. Fürnkranz

TECHNISCHE UNIVERSITÄT DARMSTADT

Beispiellösung für das 5. Übungsblatt (30.06.2009)

Aufgabe 1 Wahrscheinlichkeitsrechnung

Aus der Aufgabenbeschreibung lässt sich folgendes entnehmen:

$$\begin{aligned}
P(N) &= 0.9\\
P(J \mid N) &= 0.4\\
P(S \mid N) &= 0.2\\
P(J,S \mid N) &= 0.1\\
P(J,S \mid \neg N) &= 0.8
\end{aligned}$$

In der Aufgabe war nun gefordert $P(N \mid J,S)$ zu berechnen. Dies lässt sich z.B. wie folgt berechnen:

$$\begin{aligned}
P(N \mid J,S) &= \frac{P(J,S \mid N)\,P(N)}{P(J,S)}\\
&= \frac{0.1 \cdot 0.9}{P(J,S)}\\
&= \frac{0.09}{P(J,S,N) + P(J,S,\neg N)}\\
&= \frac{0.09}{P(J,S \mid N)\,P(N) + P(J,S \mid \neg N)P(\neg N)}\\
&= \frac{0.09}{0.1 \cdot 0.9 + 0.8 \cdot 0.1}\\
&= \frac{0.09}{0.17} \approx 0.5294
\end{aligned}$$

Aufgabe 2 Bayes'sches Netz

a) Gesucht ist $P(W \mid G \wedge \neg L)$. Ein möglicher Lösungsweg, der oft das Ergebnis aus Aufgabe 4.3c verwendet, ist wie folgt:

$$\begin{aligned}
P(W \mid G,\neg L) &= \frac{P(G,W,\neg L)}{P(G,\neg L)}\\
&= \frac{P(G,W,\neg L,D) + P(G,W,\neg L,\neg D)}{P(G,W,\neg L) + P(G,\neg W,\neg L)}\\
&= \frac{P(G,W,\neg L,D) + P(G,W,\neg L,\neg D)}{(P(G,W,\neg L,D) + P(G,W,\neg L,\neg D)) + (P(G,\neg W,\neg L,D) + P(G,\neg W,\neg L,\neg D))}\\
&= \frac{P(G \mid W) \cdot P(W \mid \neg L,D) \cdot P(\neg L)P(D) + P(G \mid W) \cdot P(W \mid \neg L,\neg D) \cdot P(\neg L) \cdot P(\neg D)}{(P(G,W,\neg L,D) + P(G,W,\neg L,\neg D)) + (P(G,\neg W,\neg L,D) + P(G,\neg W,\neg L,\neg D))}\\
&= \frac{0.9 \cdot 0.5 \cdot 0.4 \cdot 0.6 + 0.9 \cdot 0.3 \cdot 0.4 \cdot 0.4}{(P(G,W,\neg L,D) + P(G,W,\neg L,\neg D)) + (P(G,\neg W,\neg L,D) + P(G,\neg W,\neg L,\neg D))}\\
&= \frac{0.1512}{0.1512 + (P(G,\neg W,\neg L,D) + P(G,\neg W,\neg L,\neg D))}\\
&= \frac{0.1512}{0.1512 + (P(G \mid \neg W) \cdot P(\neg W \mid \neg L,D) \cdot P(\neg L) \cdot P(D) + P(G \mid \neg W) \cdot P(\neg W \mid \neg L,\neg D) \cdot P(\neg L) \cdot P(\neg D))}\\
&= \frac{0.1512}{0.1976} \approx 0.7652
\end{aligned}$$

b)

L	W	D	G	$P(L,W,D,G)$
T	T	T	T	$0.2268 = P(G\,\vert\,W) \cdot P(W\,\vert\,D,L) \cdot P(D) \cdot P(L)$
T	T	T	F	$0.0252 = P(\neg G\,\vert\,W) \cdot P(W\,\vert\,D,L) \cdot P(D) \cdot P(L)$
T	T	F	T	$0.0864 = \ldots$
T	T	F	F	0.0096
T	F	T	T	0.0216
T	F	T	F	0.0864
T	F	F	T	0.0288
T	F	F	F	0.1152
F	T	T	T	0.108
F	T	T	F	0.012
F	T	F	T	0.0432
F	T	F	F	0.0048
F	F	T	T	0.024
F	F	T	F	0.096
F	F	F	T	0.0224
F	F	F	F	0.0896

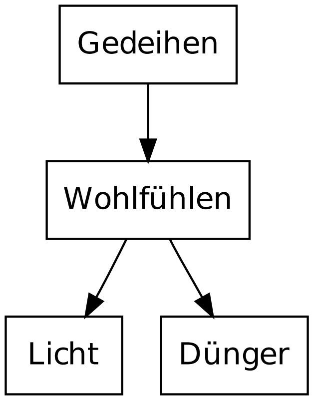
c) Gegebene Variablenreihenfolge : G, W, D, L.

1. Iteration: Variable G wird zum Netzwerk hinzugefügt. Da das Netzwerk bisher nur aus G besteht, findet keine Überprüfung nach möglichen Eltern statt.

2. Iteration: In der zweiten Iteration wird W hinzugefügt und überprüft, ob eine Kante von G nach W gelegt werden kann, also ob eine direkter Einfluß von G auf W vorliegt. Dazu wird überprüft, ob W unabhängig von G ist, d.h. ob $P(W \mid G) = P(W)$ gilt. Aus der Tabelle bestimmen wir mittels *Marginalization*
$P(W) = 0.2268 + 0.0252 + 0.0864 + 0.0096 + 0.108 + 0.012 + 0.0432 + 0.0048 = 0.516$ und
$P(W \mid G) = \frac{P(W,G)}{P(G)} = \frac{0.2268+0.0864+0.108+0.0432}{0.2268+0.0864+0.0216+0.0288+0.108+0.0432+0.024+0.0224} \approx 0.9092.$
Da W und G somit nicht unabhängig sind, kann G als Elternteil von W angesehen werden und es wird eine Kante von G nach W gelegt.

3. Iteration: Variable D wird zum Netzwerk hinzugefügt und es wird erneut ermittelt, welche Variablen als Elternteile in Frage kommen: $P(D \mid G, W) = P(D)$?
$P(D \mid G, W) = \frac{0.2268+0.108}{0.2268+0.0864+0.108+0.0432} \approx 0.7209$
$P(D) = 0.6$ (aus der Aufgabenstellung)
Da D nicht unabhängig von G, W ist, kann mindestens eins der beiden Variablen als Elternknoten benutzt werden.
$P(D \mid G, W) = P(D \mid G)$?
$P(D \mid G, W) \approx 0.7209$
$P(D \mid G) = \frac{0.2268+0.0216+0.108+0.024}{0.2268+0.0864+0.0216+0.0288+0.108+0.0432+0.024+0.0224} \approx 0.6778$
Die obige Überprüfung besagt, dass D von W abhängig ist. Es wird eine Kante von W nach D eingefügt. Es bleibt noch zu prüfen, ob auch D von G abhängig ist.
$P(D \mid G, W) = P(D \mid W)$?
$P(D \mid G, W) \approx 0.7209$
$P(D \mid W) = \frac{0.2268+0.0252+0.108+0.012}{0.2268+0.0252+0.0864+0.0096+0.108+0.012+0.0432+0.0048} \approx 0.7209$
Da D nicht von G abhängig ist, wird keine neue Kante erzeugt.

4. Iteration: Als letzte Variable wird L zum Netzwerk hinzugefügt. Wir überprüfen erneut, von welchen Variablen L abhängig ist.
$P(L \mid G, W, D) = P(L)$?
$P(L \mid G, W, D) = \frac{0.2268}{0.2268+0.108} \approx 0.6774$
$P(L) = 0.6$
$P(L \mid G, W, D) = P(L|D)$?
$P(L \mid G, W, D) \approx 0.6774$
$P(L \mid D) = 0.6$ (*Beachten Sie, dass D unabhängig von L ist*)
$P(L \mid G, W, D) = P(L \mid W)$?
$P(L \mid G, W, D) \approx 0.6774$
$P(L \mid W) \approx 0.6774$
$\rightarrow$ G und D sind unabhängig für L und das Wissen über W ist aus Alice's Sicht ausreichend um schließen zu können, ob eine Planze genug Licht bekommen hat oder nicht. Als letzte Kante wird somit eine Kante von W nach L ins Netzwerk eingefügt.



Beachten Sie, dass die Konstruktionsvorschrift auf den Folien keine konkrete Angaben macht, in welcher Reihenfolge die möglichen Variablen für PARENTS(X_i) *durchgeführt werden. Darüberhinaus ist nicht angegeben, was bei mehrdeutigen $Parents$ (eine offensichtliches Kriterium ist bzgl. der Kompaktheit, die Anzahl der Variablen in* PARENTS(X_i)*) zu tun ist. Die Folie hat diesbezüglich eher einen deklarativen Charakter, d.h. es beschreibt was erreicht werden soll und nicht wie es erreicht wird.*

Aufgabe 3 Approximative Inferenz in Bayes'sche Netzen

a) Die approximierte gemeinsame Wahrscheinlichkeitstabelle ergibt sich durch einfaches *Abzählen* der Stichproben.

C	S	R	W	$P(C,S,R,W)$
T	T	T	T	$\frac{2}{40}$
T	T	T	F	0
T	T	F	T	0
T	T	F	F	0
T	F	T	T	$\frac{16}{40}$
T	F	T	F	$\frac{3}{40}$
T	F	F	T	0
T	F	F	F	$\frac{2}{40}$
F	T	T	T	$\frac{3}{40}$
F	T	T	F	0
F	T	F	T	$\frac{5}{40}$
F	T	F	F	0
F	F	T	T	$\frac{1}{40}$
F	F	T	F	0
F	F	F	T	0
F	F	F	F	$\frac{8}{40}$

b) Bei REJECTION-SAMPLING werden nur Stichproben betrachtet, die konsistent mit der gegebenen Evidenz sind. Das heisst, alle Stichproben, die mit der Evidenz (hier $Sprinkler = true$) übereinstimmen, werden zur Abschätzung der Abfrage herangezogen. Aus der Aufgabenstellung sind zehn Stichproben relevant, von denen für fünf Stichproben $Rain = true$ gilt. Somit wird die Wahrscheinlichkeit $P(Rain|Sprinkler = true)$ als 50 % approximiert, bzw. $\mathbf{P}(Rain|Sprinkler = True) = (0.5, 0.5)$.