Quantitative Methoden Wissensbasierter Systeme

Probabilistische Netze und ihre Anwendungen

Robert Remus

Universität Leipzig
Fakultät für Mathematik und Informatik
Abteilung für Intelligente Systeme

23. Januar 2008

Übersicht

Übersicht

Bedingte Wahrscheinlichkeit

- Sei P eine **Wahrscheinlichkeitsfunktion**
- Seien A, B Formeln mit $P(A) > 0$
- Dann ist die **bedingte Wahrscheinlichkeit** von B gegeben A definiert als
$$P(B \mid A) = \frac{P(A \ \wedge \ B)}{P(A)}$$
- $P(B \mid A)$ gibt also die Wahrscheinlichkeit dafür an, dass wenn A **wahr** ist, auch B **wahr** ist

Der Satz von Bayes

- Sei P eine Wahrscheinlichkeitsfunktion und
- seien A, B Formeln mit $P(A) > 0, P(B) > 0$,
- dann gilt **der Satz von Bayes**

$$P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$$

Übersicht

1. **Bayes'sche Netze**
   - Wahrscheinlichkeitstheoretische Grundlagen
   - Graphentheoretische Grundlagen
   - Motivation Bayes'scher Netze
   - Beispiele Bayes'scher Netze
   - Definition eines Bayes'schen Netzes

2. **Inferenz in probabilistischen Netzen**
   - Bayes-Netze und ihre Potentialdarstellungen
   - Der permanente Cliquenbaum als Wissensbasis

Cliquen

- $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$ sei ein **ungerichteter Graph** mit
  - $\mathbf{V}$ Menge der Knoten
  - $\mathcal{E}$ Menge der Kanten
- Eine Menge $\mathbf{C} \subseteq \mathbf{V}$ heißt **Clique** von $\mathcal{G}$, wenn $\mathbf{C}$ eine maximal vollständige Menge ist, d.h.
  - wenn alle Knoten aus $\mathbf{C}$ paarweise durch eine Kante aus $\mathcal{E}$ miteinander verbunden sind und
  - wenn es keine vollständige Teilmenge von $\mathbf{V}$ gibt, die $\mathbf{C}$ echt enthält

Triangulierte Graphen

- $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$ sei ein **ungerichteter Graph**
- Eine **Sehne** eines Zyklus' $v_0, \ldots, v_n$ in $\mathcal{G}$ ist eine Kante zwischen zwei nicht aufeinanderfolgenden Knoten v_i, v_j, d.h. $|i - j| > 1$
- $\mathcal{G}$ heißt **trianguliert**, wenn jeder einfache Zyklus der Länge $l > 3$ eine Sehne besitzt

Fill-in-Graphen

- $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$ sei ein **ungerichteter Graph**
- α sei eine **lineare Ordnung** auf den Knoten von $\mathcal{G}$
- Der **Fill-in** von $\mathcal{G}$ bzgl. α ist die Kantenmenge $\mathcal{F}(\alpha)$, wobei
  - $(v, w) \in \mathcal{F}(\alpha)$ gdw. $(v, w) \notin \mathcal{E}$ und es einen Weg zwischen v, w gibt, der außer v, w nur Knoten enthält, die bzgl. der Ordnung α v, w nachgeordnet sind
- Der **Fill-in-Graph** von $\mathcal{G}$ bzgl. α ist der Graph

$$\mathcal{G}(\alpha) = \langle \mathbf{V}, \mathcal{E} \cup \mathcal{F}(\alpha) \rangle$$

- $\mathcal{G}(\alpha)$ ist **trianguliert**

Morale Graphen

- $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$ sei ein **DAG**[1]
- Der **morale Graph** $\mathcal{G}_m$ zu $\mathcal{G}$ wird folgendermaßen konstruiert:
  - Sind zwei Knoten u, v Elternknoten eines gemeinsames Kindknotens w, d.h. $u, v \in pa(w)$ und sind u, v noch durch keine Kante verbunden, füge die Kante (u, v) oder die Kante (v, u) zu $\mathcal{E}$ hinzu
  - Dadurch entsteht ein gerichteter Graph $\mathcal{G}_m^d$
  - $\mathcal{G}_m$ ist der zu $\mathcal{G}_m^d$ gehörige ungerichtete Graph

[1] Directed Acyclic Graph

Übersicht

Motivation Bayes'scher Netze - Beispiel

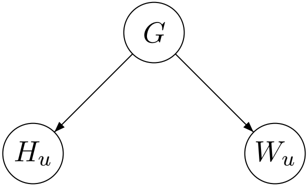
- die Zufallsvariablen A, B bezeichnen die Ergebnisse der Würfe zweier gleicher, fairer **Münzen**
- die Zufallsvariable C bezeichnet das Klingeln einer **Glocke** gdw. die Münzen nach einem Wurf das gleiche Bild zeigen
- A, B sind **unabhängig**
- A, B sind **bedingt abhängig** unter Kenntnis von C
- ein Graph G müsste nun die Unabhängigkeit von A, B und die Abhängigkeit von A, B, C, d.h. **einseitige Abhängigkeiten**, zum Ausdruck bringen
- dies ist jedoch mit einem **ungerichteten** Graphen G' nicht möglich

Übersicht

Beispiele Bayes'scher Netze I

„Holmes & Watson in London"

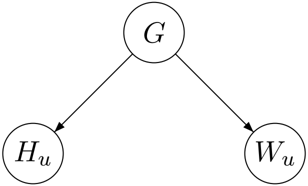
- G, H_u, W_u seien **zweiwertige Zufallsvariablen**
  - G : glatte Straßen
  - H_u: Holmes hat einen Unfall
  - W_u: Watson hat einen Unfall



Beispiele Bayes'scher Netze I

„Holmes & Watson in London“

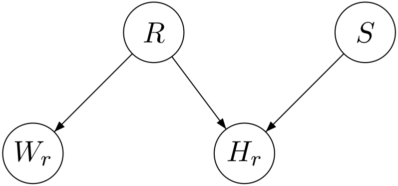
- G, H_u, W_u seien **zweiwertige Zufallsvariablen**
  - G : glatte Straßen
  - H_u: Holmes hat einen Unfall
  - W_u: Watson hat einen Unfall



Beispiele Bayes'scher Netze II

„Holmes & Watson in Los Angeles"

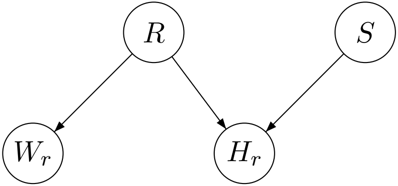
- R, S, H_r, W_r seien **zweiwertige Zufallsvariablen**
  - R: es hat geregnet
  - S: der Rasensprenger hat sich eingeschaltet
  - H_r: Holmes' Rasen ist nass
  - W_r: Watsons Rasen ist nass



Beispiele Bayes'scher Netze II

„Holmes & Watson in Los Angeles“

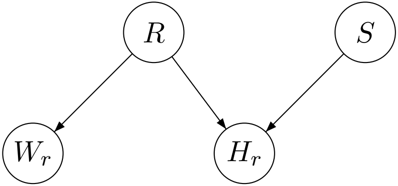
- R, S, H_r, W_r seien **zweiwertige Zufallsvariablen**
  - R: es hat geregnet
  - S: der Rasensprenger hat sich eingeschaltet
  - H_r: Holmes' Rasen ist nass
  - W_r: Watsons Rasen ist nass



Beispiele Bayes'scher Netze II

„Holmes & Watson in Los Angeles“

- R, S, H_r, W_r seien **zweiwertige Zufallsvariablen**
  - R: es hat geregnet
  - S: der Rasensprenger hat sich eingeschaltet
  - H_r: Holmes' Rasen ist nass
  - W_r: Watsons Rasen ist nass



Übersicht

1. **Bayes'sche Netze**
   - Wahrscheinlichkeitstheoretische Grundlagen
   - Graphentheoretische Grundlagen
   - Motivation Bayes'scher Netze
   - Beispiele Bayes'scher Netze
   - Definition eines Bayes'schen Netzes
2. **Inferenz in probabilistischen Netzen**
   - Bayes-Netze und ihre Potentialdarstellungen
   - Der permanente Cliquenbaum als Wissensbasis

Definition eines Bayes'schen Netzes

- $\mathbf{V}$ sei eine Menge von Aussagenvariablen
- P sei eine gemeinsame Verteilung über $\mathbf{V}$
- $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$ sei ein **DAG**
- für jedes $A_i \in \mathbf{V}$ bezeichne
  - $pa(A_i) \subseteq \mathbf{V}$ die Menge aller **Elternknoten**,
  - $de(A_i) \subseteq \mathbf{V}$ die Menge aller **Nachkommen** und
  - $nd(A_i) \subseteq \mathbf{V}$ die Menge aller **Nicht-Nachkommen**

  von A_i
- $\mathcal{B} = \langle \mathbf{V}, \mathcal{E}, P \rangle$ heißt dann **Bayes'sches Netz**, wenn für jedes A_i gilt:
  - $A_i \perp\!\!\!\perp_P nd(A_i) \mid pa(A_i)$[2]

[2] d.h. A_i ist probabilistisch unabhängig von der Ausprägung seiner Nicht-Nachkommen unter Kenntnis der Ausprägung seiner Elternknoten

Berechnung der gemeinamen Verteilung P

Wendet man die **Kettenregel** an und lässt in sie die **Unabhängigkeitsannahmen** $A_i \perp\!\!\!\perp_P nd(A_i) \mid pa(A_i)$ einfließen, zerfällt die Verteilung P in ein handliches Produkt:

- $\mathcal{B} = \langle \mathbf{V}, \mathcal{E}, P \rangle$ sei Bayes'sches Netz
- die gemeinsame Verteilung P lässt sich dann durch

$$P(\mathbf{V}) = \prod_{V \in \mathbf{V}} P(V \mid pa(V))$$

mit $P(V \mid \emptyset) = P(V)$ darstellen

Übersicht

Bayes-Netze und ihre Potentialdarstellungen

- $\mathcal{B} = \langle \mathbf{V}, \mathcal{E}, P \rangle$ sei **Bayes'sches Netz** mit DAG $\mathcal{G} = \langle \mathbf{V}, \mathcal{E} \rangle$
- $\mathcal{G}_u$ sei eine **Triangulierung des moralen Graphen** $\mathcal{G}_m$ von $\mathcal{G}$
- $\{C_i \mid 1 \leq i \leq p\}$ seien die **Cliquen** von $\mathcal{G}_u$
- Wähle für jedes $V \in \mathbf{V}$ eine Clique $clq(V) \in \{C_i \mid 1 \leq i \leq p\}$, so dass $V \cup pa(V) \subseteq clq(V)$ gilt
- Definiere für $1 \leq i \leq p$

$$\psi(\mathbf{C}_i) = \prod_{V : clq(V) = C_i} P(V \mid pa(V))$$

- $\{C_1, \ldots, C_p; \psi\}$ ist dann eine **Potentialdarstellung** von P

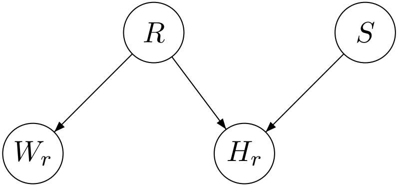
Übersicht

Der permanente Cliquenbaum als Wissensbasis

Cliquenbaum und **Potentialdarstellung** bilden zusammen die Wissensbasis.

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung I

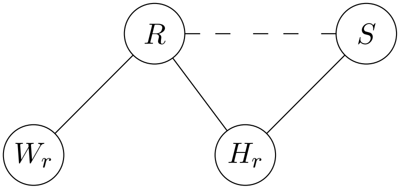
Gegeben sei das Bayes'sche Netzwerk $\mathcal{B} = \langle \mathbf{V}, \mathcal{E}, P \rangle$ mit DAG $(\mathbf{V}, \mathcal{E})$ aus dem „Holmes & Watson in Los Angeles Beispiel":



1. Bilde den **moralen Graphen** $\mathcal{G}_m$ von $(\mathbf{V}, \mathcal{E})$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung II

- *Lösung:* Der morale Graphe $\mathcal{G}_m$ von $(\mathbf{V}, \mathcal{E})$:



2. Triangulierung von $\mathcal{G}_m$: Bestimme mittels **Maximum Cardinality Search** eine lineare Ordnung α auf den Knoten $V \in \mathbf{V}$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung III

- *Lösung:* Eine lineare Ordnung α auf den Knoten in $\mathbf{V}$:

$$\begin{aligned} \alpha(W_r) &= 1 \\ \alpha(R) &= 2 \\ \alpha(S) &= 3 \\ \alpha(H_r) &= 4 \end{aligned}$$

2. Triangulierung von $\mathcal{G}_m$: Berechne den **Fill-in-Graphen** $\mathcal{G}(\alpha)$ von $\mathcal{G}_m$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung IV

- *Lösung:* Der Fill-in-Graph $\mathcal{G}(\alpha)$ von $\mathcal{G}_m$ enthält keine neuen Kanten, es gilt also $\mathcal{G}(\alpha) = \mathcal{G}_m = \mathcal{G}'$

3. Ordnung der Cliquen: Bestimme die **Cliquen** $\mathbf{C}_i$ von $\mathcal{G}'$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung V

- *Lösung:* Die Cliquen $\mathbf{C}_i$ von $\mathcal{G}'$:

$$\begin{aligned}\mathbf{C}_1 &= \{R, W_r\}\\ \mathbf{C}_2 &= \{R, S, H_r\}\end{aligned}$$

3. Ordnung der Cliquen: **Ordne** die Cliquen $\mathbf{C}_i$ nach dem gemäß α jeweils größten in ihnen vorkommenden Knoten

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung VI

- *Lösung:* Ordnung der Cliquen $\mathbf{C}_i$ ist gegeben durch $(\mathbf{C}_1, \mathbf{C}_2)$

4. Bestimme für $1 \leq i \leq p$ die Mengen $\mathbf{R}_i$ und $\mathbf{S}_i$:

$$\mathbf{S}_i = \mathbf{C}_i \cap (\mathbf{C}_1 \cup \ldots \cup \mathbf{C}_{i-1})$$
$$\mathbf{R}_i = \mathbf{C}_i \setminus \mathbf{S}_i$$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung VII

- *Lösung:* die Mengen $\mathbf{R}_i$ und $\mathbf{S}_i$ für $1 \leq i \leq p = 2$:

$$\begin{aligned}
\mathbf{S}_1 &= \mathbf{C}_1 \cap \mathbf{C}_0 = \emptyset \\
\mathbf{S}_2 &= \mathbf{C}_2 \cap \mathbf{C}_1 = \{R\} \\
\mathbf{R}_1 &= \{R, W_r\} \\
\mathbf{R}_2 &= \{S, H_r\}
\end{aligned}$$

5. Bestimme für jedes $i > 1$ ein $j < i$, so dass $\mathbf{S}_i \subseteq \mathbf{C}_j$

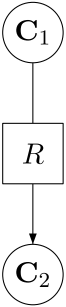
Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung VIII

- *Lösung:* Ein $j < i$ für $i > 1$, so dass $\mathbf{S}_i \subseteq \mathbf{C}_j$: $j = 1$, da $\mathbf{S}_2 = \{R\} \subseteq \{R, W_r\} = \mathbf{C}_1$. $\mathbf{C}_1$ heißt jetzt *Elternclique* von $\mathbf{C}_2$

6. Bilde anhand der im vorherigen Punkt festgelegten Eltern-Kind-Beziehungen einen **Cliquenbaum**

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung IX

- *Lösung:* Der Cliquenbaum:



7. Bestimme zu jedem $V \in \mathbf{V}$ eine Clique $clq(V)$ mit $\{V\} \cup pa(V) \subseteq clq(V)$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung X

- *Lösung:* Cliquen $clq(V)$ mit $\{V\} \cup pa(V) \subseteq clq(V)$ für jedes $V \in \mathbf{V}$:

$$\begin{aligned} clq(W_r) &= \mathbf{C}_1 \\ clq(R) &= \mathbf{C}_2 \\ clq(S) &= \mathbf{C}_2 \\ clq(H_r) &= \mathbf{C}_2 \end{aligned}$$

8. Definiere für $1 \leq i \leq p$

$$\psi(\mathbf{C}_i) = \prod_{V:clq(V)=\mathbf{C}_i} P(V \mid pa(V))$$

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung XI

- *Lösung:* Definition von $\psi(\mathbf{C}_i)$ für $i = 1$

$$\psi(\mathbf{C}_1) = \psi(\{R, W_r\}) = P(W_r \mid R)$$

also

R	W_r	$\psi(\{R, W_r\})$
0	0	0.8
0	1	0.2
1	0	0
1	1	1

Erzeugung des permanenten Cliquenbaums mit Potentialdarstellung XII

- *Lösung:* Definition von $\psi(\mathbf{C}_i)$ für $i = 2$:

$$\psi(\mathbf{C}_2) = \psi(\{R, S, H_r\}) = P(R)P(S)P(H_r \mid R, S)$$

also

R	S	W_r	$\psi(\{R, S, H_r\})$
0	0	0	0.72
0	0	1	0
0	1	0	0.008
0	1	1	0.072
1	0	0	0
1	0	1	0.18
1	1	0	0
1	1	1	0.02