UNIVERSITÄT ZU LÜBECK

Lifting in Gaussian Bayesian Networks

Mattis Hartwig

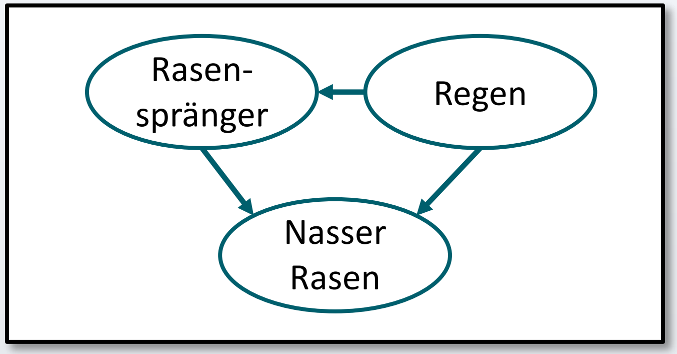
Institut für Informationssysteme, Universität zu Lübeck

Lübeck, 30. Januar 2023

Der Titel lässt sich in drei Bestandteile zerlegen

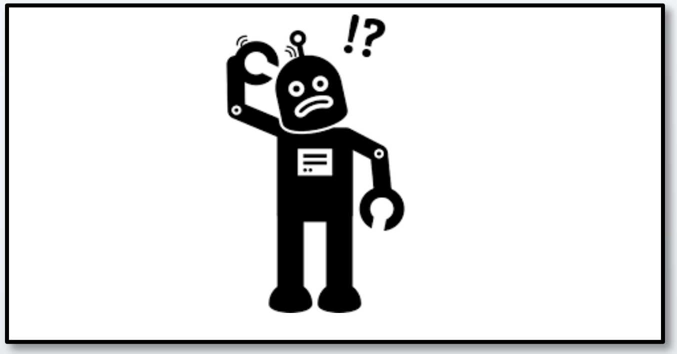
UNIVERSITÄT ZU LÜBECK

Probabilistic Graphical Models



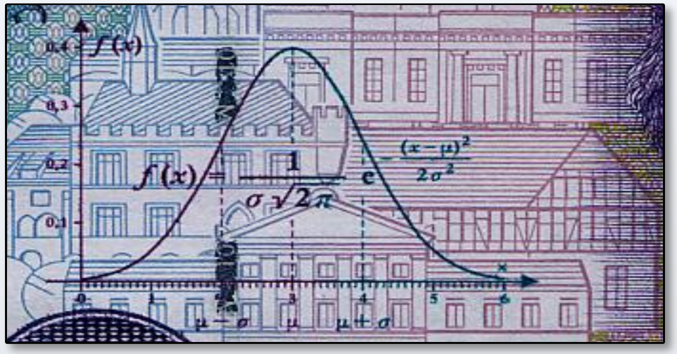
- Graphische Modelle sind eine **Kompakte Darstellung** einer Verbundwahrscheinlichkeit
- **Geeignet** um auch **kausale Beziehungen** zu modellieren

Efficient Query Answering



- Künstliche Agenten benötigen Antworten auf Modellanfragen
- **Schnellere Antwortzeit** erhöht **Reaktionsgeschwindigkeit** und damit Leistung des Agenten

Gaussian



- **Kontinuierliche Variablen** benötigen eine **Dichtefunktion**
- Normalverteilung hat gute Eigenschaften **in Anwendungen**

Spezifische Anwendungsszenarien bedürfen neuer Methoden.

Im Folgenden schauen wir auf Gaußsche Bayessche Netze

Gaußsches Bayessches Netz (GBNs)

Beispiel

$$P(A) \sim \mathcal{N}(\mu_A, \sigma_A^2)$$
$$P(D) \sim \mathcal{N}(\mu_D, \sigma_D^2)$$
$$P(K|A, D) \sim \mathcal{N}(\mu_K + \beta_{A,K}(x_A - \mu_a) + \beta_{D,K}(x_D - \mu_D), \sigma_K^2)$$

Allgemeine Formel

$$P(X_i|\boldsymbol{Pa}(X_i)) \sim \mathcal{N}\left(\mu_i + \sum_{X_k \in \boldsymbol{Pa}(X_i)} \beta_{k,i}(x_k - \mu_k), \sigma_i^2\right)$$

Multivariate Verbundwahrscheinlichkeit (hier Normalverteilung)

$$\mathcal{N}(\mu, \Sigma)$$

Beispiel

$$\mu = \begin{pmatrix} \mu_A \\ \mu_D \\ \mu_K \end{pmatrix}$$

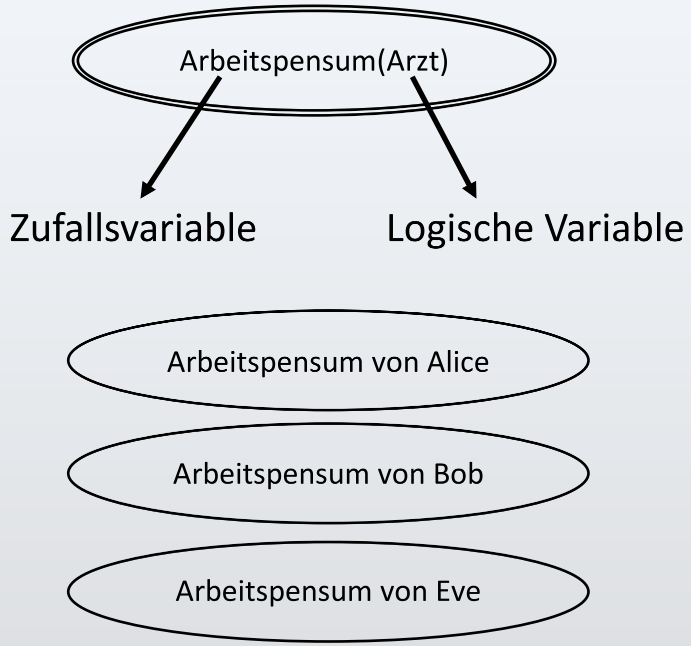
$$\Sigma = \begin{pmatrix} \sigma_A^2 & 0 & \beta_{A,K}\sigma_A^2 \\ 0 & \sigma_D^2 & \beta_{D,K}\sigma_D^2 \\ \beta_{A,K}\sigma_A^2 & \beta_{D,K}\sigma_D^2 & \beta_{D,K}^2\sigma_D^2 + \beta_{A,K}^2\sigma_A^2 + \sigma_K^2 \end{pmatrix}$$

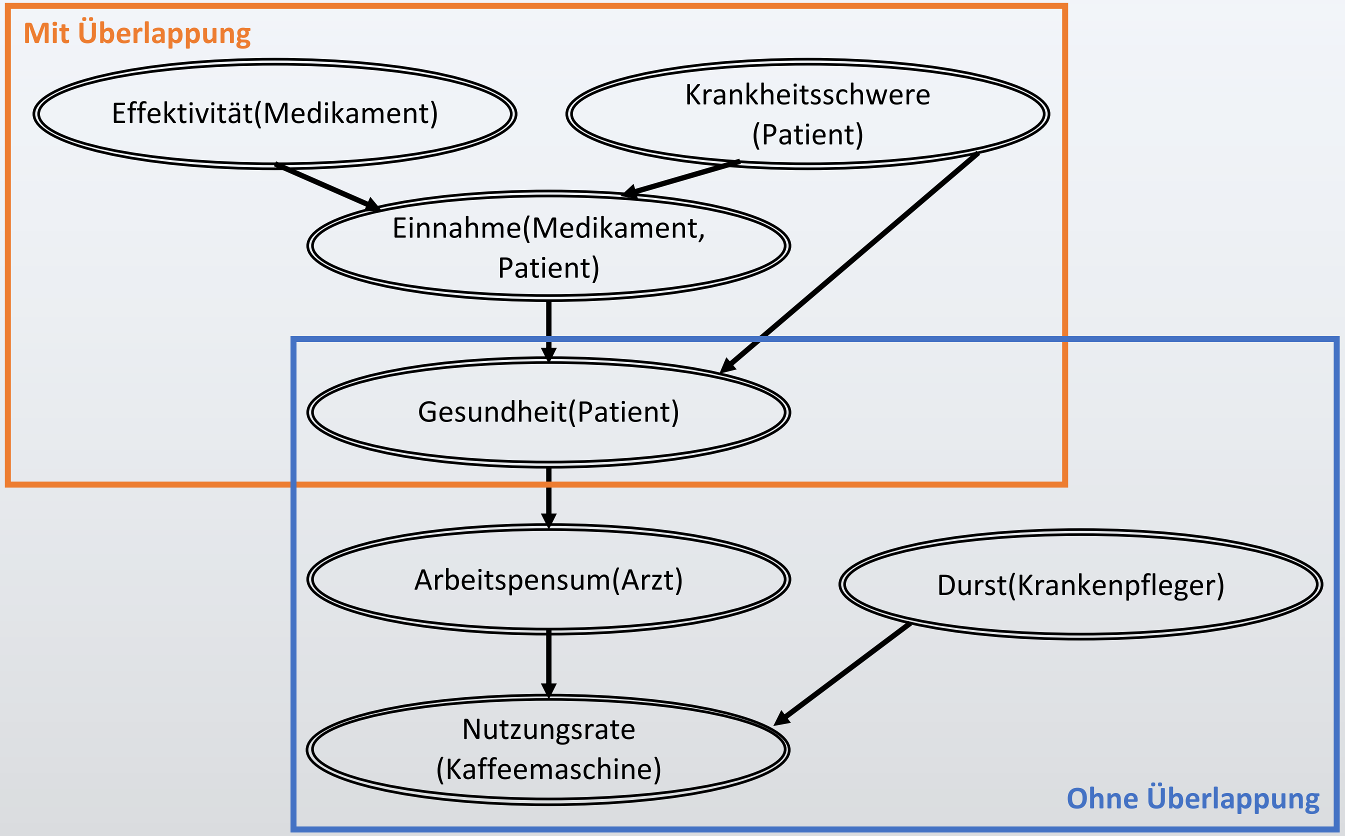
Ein Beispielmodell aus parametrisierten Zufallsvariablen

Anzahl — Parametrisierte Zufallsvariable

M

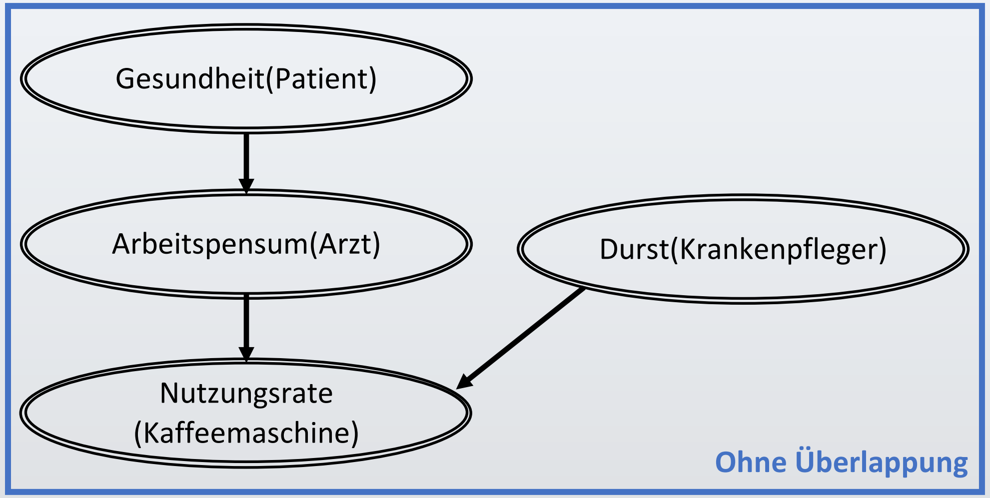
N





Die Überlappung bezieht sich auf die Sequenz von logischen Variablen.

In einem Szenario, in dem sich die logischen Variablen nicht überlappen kommt, es zu einfachen Kovarianzmatrix-Strukturen



$$\underbrace{\Sigma = \left[\begin{array}{cc|ccc|cc|ccc}
1 & 0 & 0 & 0 & 0 & -2 & -2 & -12 & -12 & -12 \\
0 & 1 & 0 & 0 & 0 & -2 & -2 & -12 & -12 & -12 \\
\hline
0 & 0 & 2 & 0 & 0 & 0 & 0 & 4 & 4 & 4 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 4 & 4 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 4 & 4 & 4 \\
\hline
-2 & -2 & 0 & 0 & 0 & 11 & 8 & 57 & 57 & 57 \\
-2 & -2 & 0 & 0 & 0 & 8 & 11 & 57 & 57 & 57 \\
\hline
-12 & -12 & 4 & 4 & 4 & 57 & 57 & 370 & 366 & 366 \\
-12 & -12 & 4 & 4 & 4 & 57 & 57 & 366 & 370 & 366 \\
-12 & -12 & 4 & 4 & 4 & 57 & 57 & 366 & 366 & 370
\end{array}\right]}_{\text{Speicheraufwand NxN}}$$

$$\underbrace{\rho = \begin{bmatrix} 0 & 0 & -2 & -12 \\ 0 & 0 & 0 & 4 \\ -2 & 0 & 8 & 57 \\ -12 & 4 & 57 & 366 \end{bmatrix} \qquad \sigma = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}}_{\text{Speicheraufwand MxM}}$$

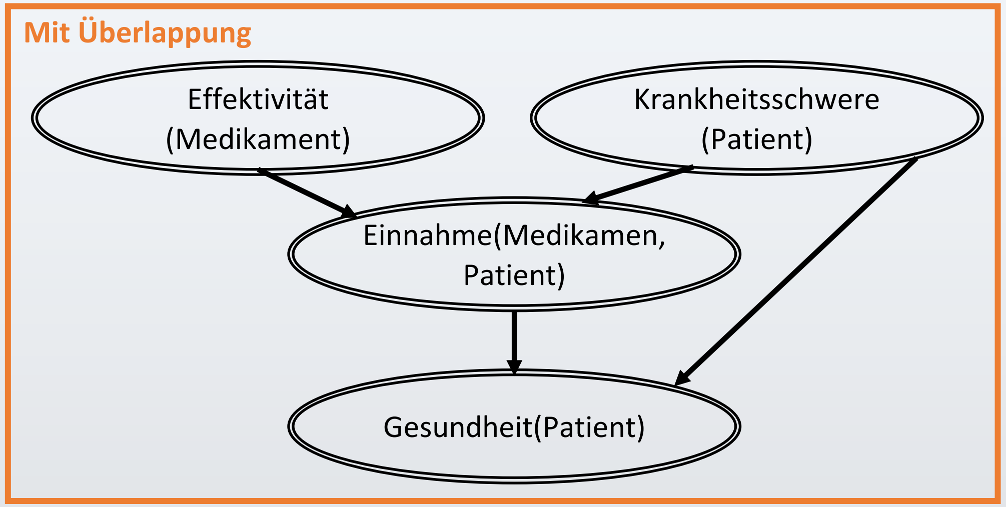
UNIVERSITÄT ZU LÜBECK

In einem Szenario, in dem sich die logischen Variablen nicht überlappen kommt, es zu einfachen Kovarianzmatrix-Strukturen

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -2 & -2 & -12 & -12 & -12 \\ 0 & 1 & 0 & 0 & 0 & -2 & -2 & -12 & -12 & -12 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 4 & 4 & 4 \\ -2 & -2 & 0 & 0 & 0 & 11 & 8 & 57 & 57 & 57 \\ -2 & -2 & 0 & 0 & 0 & 8 & 11 & 57 & 57 & 57 \\ -12 & -12 & 4 & 4 & 4 & 57 & 57 & 370 & 366 & 366 \\ -12 & -12 & 4 & 4 & 4 & 57 & 57 & 366 & 370 & 366 \\ -12 & -12 & 4 & 4 & 4 & 57 & 57 & 366 & 366 & 370 \end{bmatrix}$$

$$\rho = \begin{bmatrix} 0 & 0 & -2 & -12 \\ 0 & 0 & 0 & 4 \\ -2 & 0 & 8 & 57 \\ -12 & 4 & 57 & 366 \end{bmatrix}$$

In Szenarien, in denen sich die logischen Variablen überlappen, ist die Struktur etwas komplizierter



3-dimensionaler Tensor

$\underbrace{\hspace{6em}}_{\text{Speicheraufwand NxN}}$

$\underbrace{\hspace{6em}}_{\text{Speicheraufwand MxMx}2^S}$

Formel für die Struktur von jedem Block

$$\Sigma_{gr(Y_s,Y_t)} = \sum_{\boldsymbol{q}_{s,t} \in \boldsymbol{\Phi}_{s,t}} \rho_{s,t}^{\boldsymbol{q}_{s,t}} \bigotimes_{L \in \boldsymbol{L}_s \cup \boldsymbol{L}_t} \mathbf{J}^{qexp(\boldsymbol{q}_{s,t},L)}_{dim(L,\boldsymbol{L}_s)\times dim(L,\boldsymbol{L}_t)}$$

Ziel ist es, Anfragen zu beantworten und Berechnungen weitestgehend mit der angehobenen Repräsentation durchzuführen

Anfrage (Query)

$$P(Q|E = e)$$

Anfrageterme

Evidenz

Antwort

$$P(Q|E = e) \sim \mathcal{N}(\mu^*, \Sigma^*)$$

$$\mu^* = \mu_Q + \Sigma_{QE}\Sigma_{EE}^{-1}(e - \mu_E),$$
$$\Sigma^* = \Sigma_{QQ} - \Sigma_{QE}\Sigma_{EE}^{-1}\Sigma_{EQ}.$$

Lösungsweg

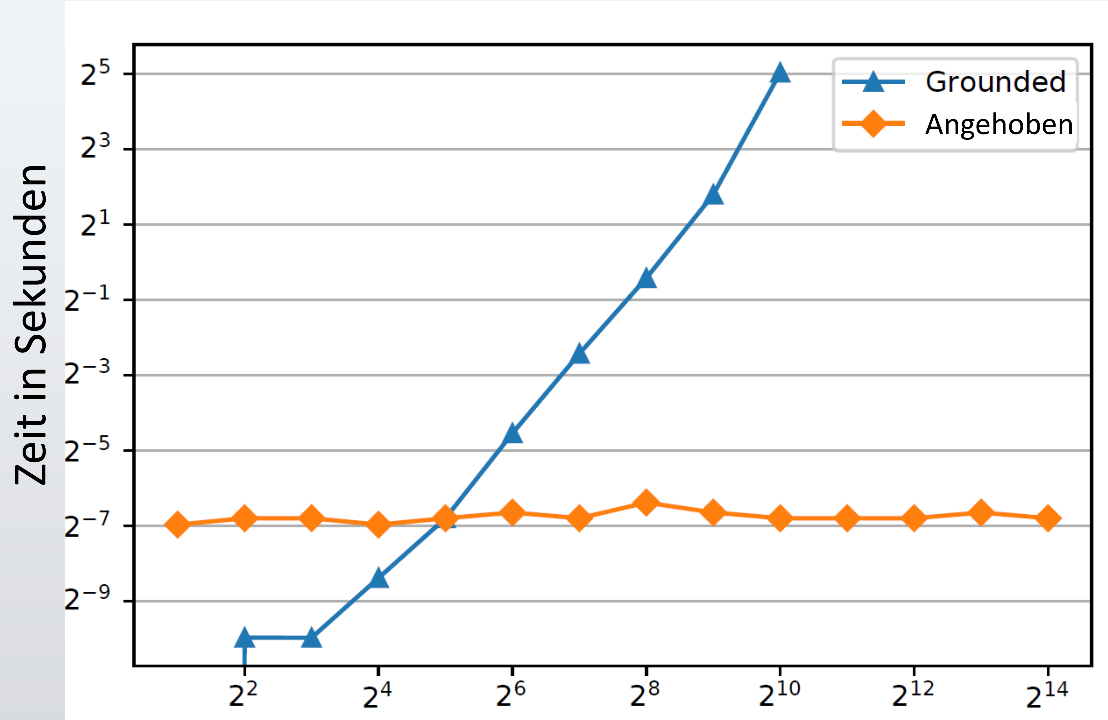
Nötige Operationen mit der angehobenen Repräsentation

- Addition
- Multiplikation von Teilmatrizen
- Invertierung von Teilmatrizen

Notiz: Es gibt gewisse Symmetrie-Bedingungen in den Queries, damit diese mit der angehobenen Repräsentation gerechnet werden können.

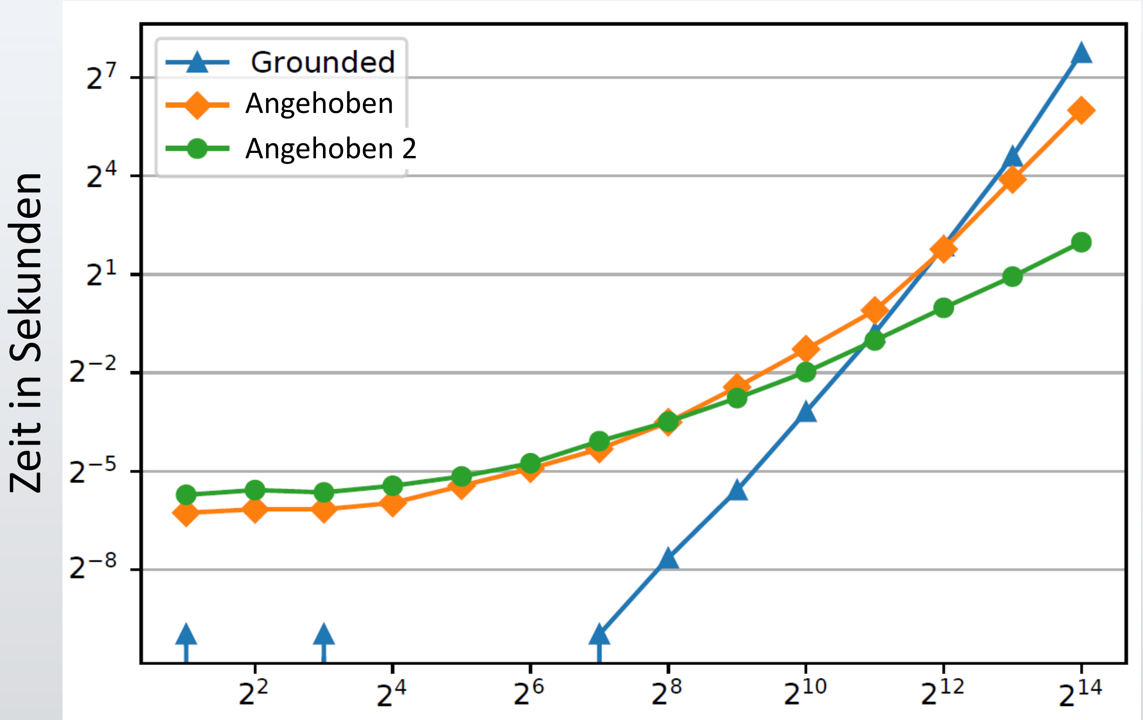
Die theoretischen Analysen zu den implementierten Algorithmen wurden ebenfalls ausführlich evaluiert

Konstruktion der Verbundwahrscheinlichkeit



Kardinalität der logischen Variablen

Beantwortung von Fragen



Kardinalität der logischen Variablen

Ausblick

- Angehobene Repräsentationen für hybride Modelle
- Behandlung von unsymmetrischen Queries
- Lifted Message Passing
- Causality
- Approximate indistiguishability

Danke! Jetzt ist Zeit für Fragen.