

Diplomvortrag

Anpassung von Data-Warehouse-Techniken für den Einsatz unsicherer Verkehrsdaten

entstanden im Rahmen des OVID-Projektes

Daniel Härder

Institut für Programmstrukturen und Datenorganisation

Verantwortlicher Betreuer: Prof. Dr.-Ing. Klemens Böhm

Betreuender Mitarbeiter: Dipl.-Inform. Heiko Schepperle

Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Imperfekte Daten

- Ungenaue Daten:
 - Werte können nur durch Intervall oder ähnlich grobkörnige Einheit angegeben werden
- Unscharfe Daten:
 - Bei Kategorien, die nicht scharf voneinander abgegrenzt werden können (linguistische Variablen)
- Unsichere Daten:
 - Nur Angabe eines Wahrscheinlichkeitswertes möglich, der aussagt, wie sicher die Daten sind

Daten im Verkehrsumfeld

- Verkehrsdaten: → **nur geringfügig mit Unsicherheiten behaftet**
 - Erhebung durch Sensoren, Daten über aktuelle Verkehrslage
- Daten über die Verkehrsinfrastruktur: → **sehr sichere Daten**
 - Verkehrsnetze, Parkräume, ÖPNV, usw.
- Daten über Ereignisse und Störungen: → **teilweise unsicher**
 - Einflussnahme auf Verkehrsfluss / Verkehrsinfrastruktur
- Verkehrsprognosedaten: → **am stärksten mit Unsicherheit behaftet**
 - Prognosen für zukünftige Verkehrsdaten auf Basis der vorhandenen Daten

Langfristige Verkehrsprognose

Eine langfristige Verkehrsprognose kann in textueller Form beispielsweise folgendermaßen aussehen:

Mit einer Wahrscheinlichkeit von ca. 70 % kommt es auf dem Autobahnabschnitt X am Freitag, dem 22.04.2005, zwischen 15:00 und 16:00 Uhr zu einem erhöhten Verkehrsaufkommen ohne Stauungen!

→ Welches Verkehrsaufkommen ist in den restlichen 30% der Fälle?

Langfristige Verkehrsprognose

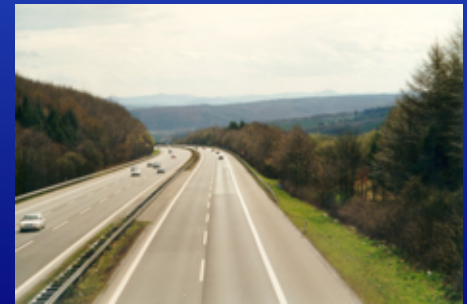
Eine langfristige Verkehrsprognose kann in textueller Form beispielsweise folgendermaßen aussehen:

• Vollständige Prognose:

→ Mit einer Wahrscheinlichkeit von ca. $\left. \begin{array}{c} 20\% \end{array} \right\}$ kommt es auf dem

Autobahnabschnitt X am Freitag, dem 22.04.2005, zwischen 15:00 und

16:00 Uhr zu $\left. \begin{array}{c} \text{freiem Verkehrsfluss} \end{array} \right\} !$



Langfristige Verkehrsprognose

Eine langfristige Verkehrsprognose kann in textueller Form beispielsweise folgendermaßen aussehen:

• Vollständige Prognose:

→ Mit einer Wahrscheinlichkeit von ca. $\left. \begin{matrix} 20\% \\ 70\% \end{matrix} \right\}$ kommt es auf dem

Autobahnabschnitt X am Freitag, dem 22.04.2005, zwischen 15:00 und

16:00 Uhr zu $\left. \begin{matrix} \text{freiem Verkehrsfluss} \\ \text{erhöhtem Verkehrsaufkommen} \end{matrix} \right\} !$



Langfristige Verkehrsprognose

Eine langfristige Verkehrsprognose kann in textueller Form beispielsweise folgendermaßen aussehen:

• Vollständige Prognose:

→ Mit einer Wahrscheinlichkeit von ca. $\left\{ \begin{array}{l} 20\% \\ 70\% \\ 10\% \end{array} \right\}$ kommt es auf dem

Autobahnabschnitt X am Freitag, dem 22.04.2005, zwischen 15:00 und

16:00 Uhr zu $\left\{ \begin{array}{l} \text{freiem Verkehrsfluss} \\ \text{erhöhtem Verkehrsaufkommen} \\ \text{Stauungen} \end{array} \right\} !$



Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Verkehrsprognose

Relation im probabilistischen relationalen Modell mit fehlenden Wahrscheinlichkeiten von Barbará et al. :

Autobahnabschnitt	Datum	Zeitfenster	Verkehrsaufkommen
X	22.04.2005	14	0,25 [frei] 0,65 [lebhaft] 0,1 [gestaut]
X	22.04.2005	15	0,2 [frei] 0,7 [lebhaft] 0,1 [gestaut]
X	22.04.2005	16	0,05 [frei] 0,6 [lebhaft] 0,25 [gestaut] 0,1 [*]

Modell von Barbará et al.

Eigenschaften des probabilistischen relationalen Modells von Barbará et al.:

- Relationen nicht in erster Normalform
- Angabe eines festen Wahrscheinlichkeitswertes
- Nur deterministische Schlüssel erlaubt
- Summe der Wahrscheinlichkeiten muss 1,0 ergeben
→ wird durch das Konzept der fehlenden Wahrscheinlichkeiten erreicht

Verkehrsprognose

Relation im probabilistischen relationalen Modell mit Wahrscheinlichkeitsintervallen von Lakshmanan et al. :

Autobahnabschnitt	Datum	Zeitfenster	Verkehrsaufkommen
$[X]$ $h_1(X) = [1, 1]$	$[22.04.05]$ $h_2(22.04.05) = [1, 1]$	$[14]$ $h_3(14) = [1, 1]$	$[frei, lebhaft, gestaut]$ $h_4(frei) = [0.2, 0.3]$ $h_5(lebhaft) = [0.6, 0.7]$ $h_6(gestaut) = [0.05, 0.15]$
$[X]$ $h_7(X) = [1, 1]$	$[22.04.05]$ $h_8(22.04.05) = [1, 1]$	$[15]$ $h_9(15) = [1, 1]$	$[frei, lebhaft, gestaut]$ $h_{10}(frei) = [0.15, 0.25]$ $h_{11}(lebhaft) = [0.65, 0.75]$ $h_{12}(gestaut) = [0.05, 0.15]$
$[X]$ $h_{13}(X) = [1, 1]$	$[22.04.05]$ $h_{14}(22.04.05) = [1, 1]$	$[16]$ $h_{15}(16) = [1, 1]$	$[frei, lebhaft, gestaut]$ $h_{16}(frei) = [0.0, 0.1]$ $h_{17}(lebhaft) = [0.55, 0.65]$ $h_{18}(gestaut) = [0.2, 0.3]$

Verkehrsprognose

Annotierte Schreibweise:

AAS	Datum	Zeitfenster	Verkehr	UG	OG	Pfad
X	22.04.05	14	frei	0,2	0,3	W ₁
X	22.04.05	14	lebhaft	0,6	0,7	W ₂
X	22.04.05	14	gestaut	0,05	0,15	W ₃
X	22.04.05	15	frei	0,15	0,25	W ₄
X	22.04.05	15	lebhaft	0,65	0,75	W ₅
X	22.04.05	15	gestaut	0,05	0,15	W ₆
X	22.04.05	16	frei	0,0	0,1	W ₇
X	22.04.05	16	lebhaft	0,55	0,65	W ₈
X	22.04.05	16	gestaut	0,2	0,3	W ₉

Modell von Lakshmanan et al.

Eigenschaften des probabilistischen relationalen Modells von Lakshmanan et al.:

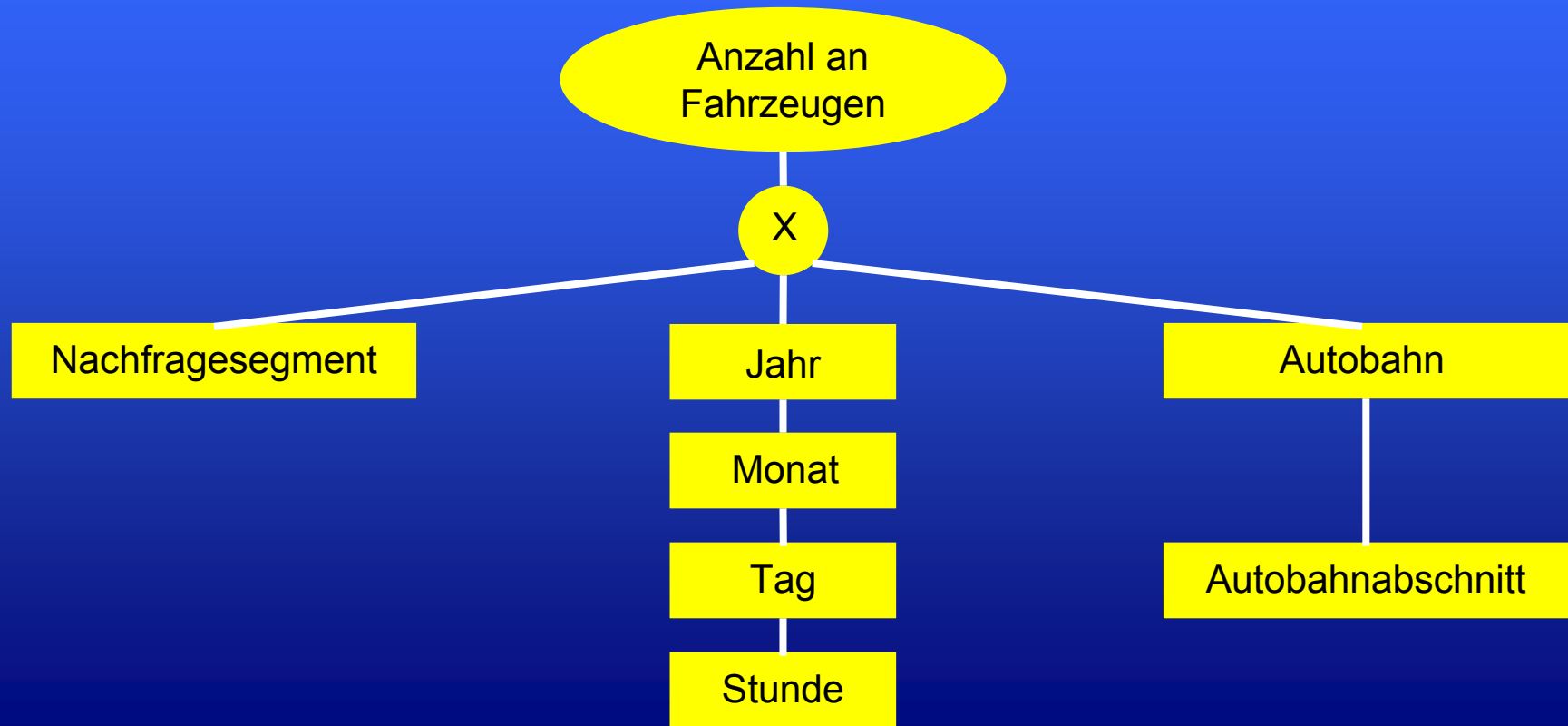
- Probabilistische Tupel aufbauend auf Modell von Barbará et al.
- Annotierte Relationen in erster Normalform
- Wahrscheinlichkeitsintervalle statt exakten Werten
- Hohe Flexibilität

Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Szenario

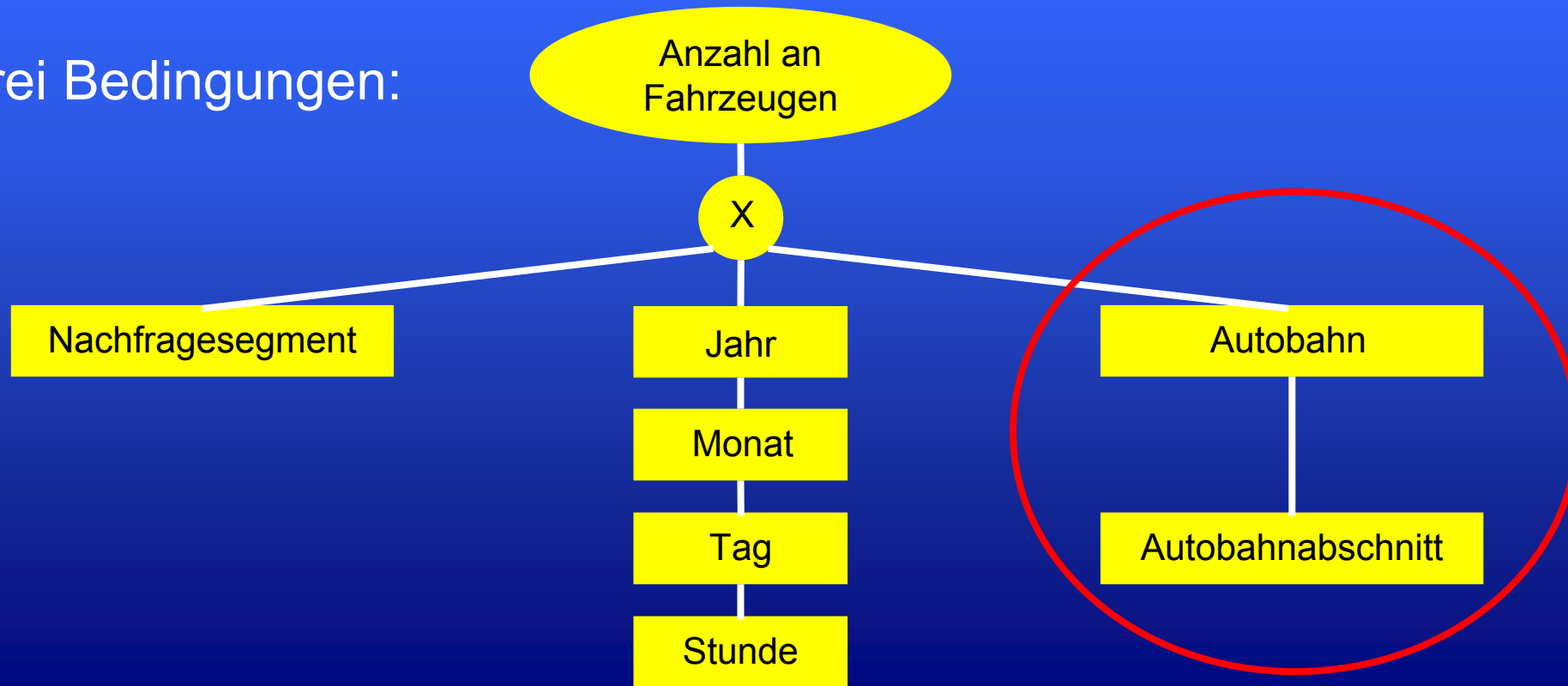
Segment	Autobahnabschnitt	Datum	Zeitraum	Fahrzeuge
PKW	X	08.04.05	15	2465



Summierbarkeit

→ Voraussetzung dafür, dass eine Aggregation bezüglich einer Kenngröße korrekt durchgeführt werden kann [Lenz&Shoshani]

Drei Bedingungen:

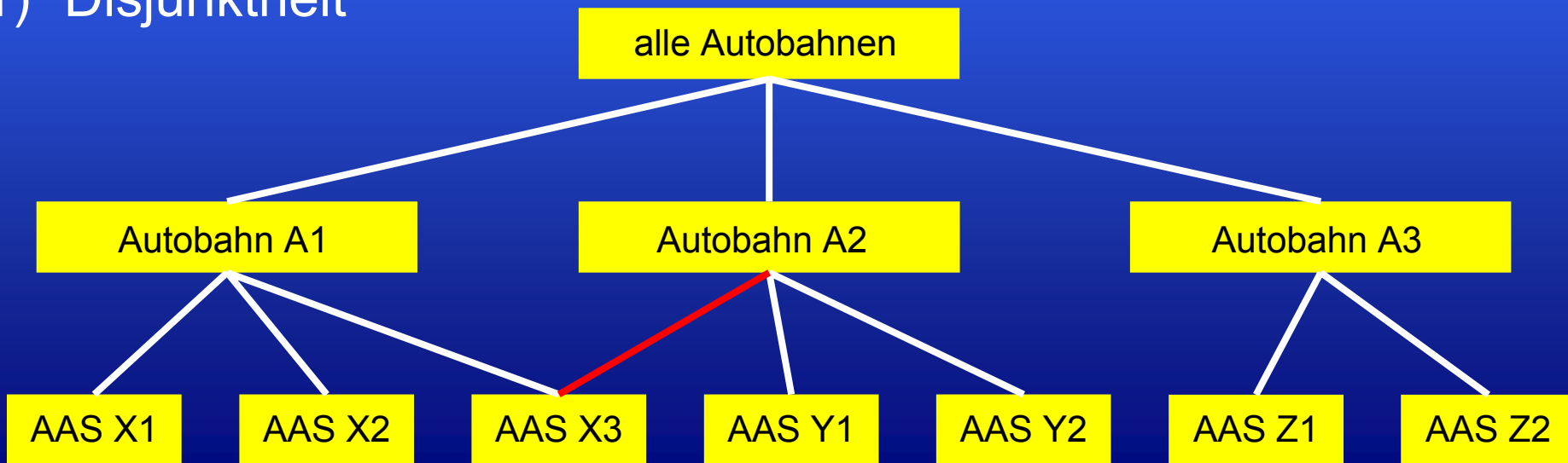


Summierbarkeit

→ Voraussetzung dafür, dass eine Aggregation bezüglich einer Kenngröße korrekt durchgeführt werden kann [Lenz&Shoshani]

Drei Bedingungen:

1) Disjunktheit



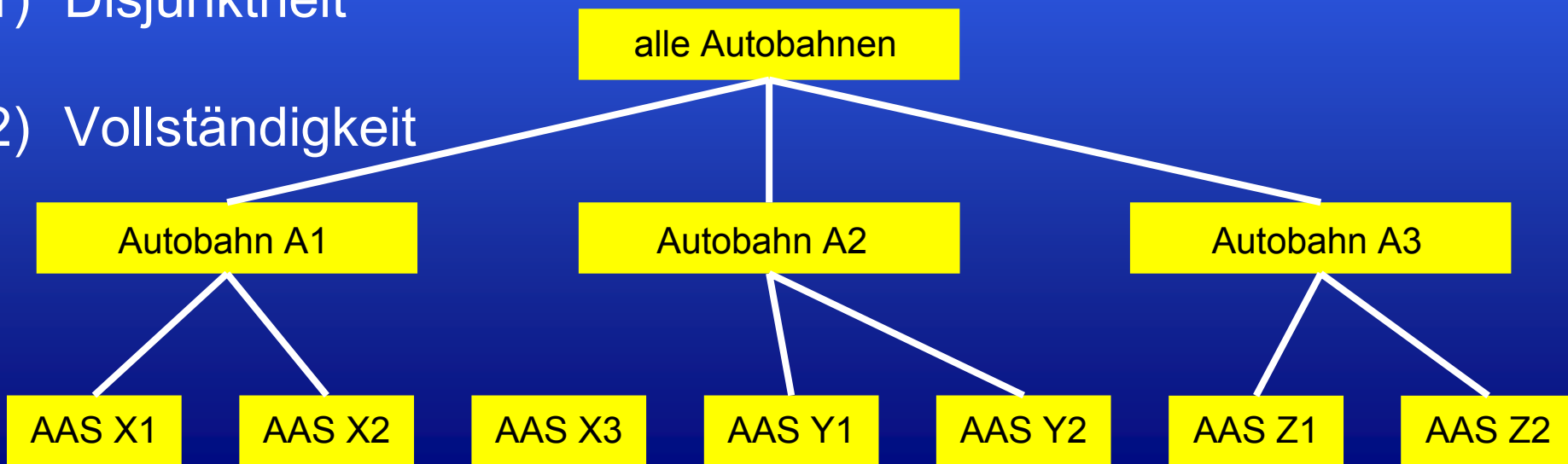
Summierbarkeit

→ Voraussetzung dafür, dass eine Aggregation bezüglich einer Kenngröße korrekt durchgeführt werden kann [Lenz&Shoshani]

Drei Bedingungen:

1) Disjunktheit

2) Vollständigkeit



Summierbarkeit

→ Voraussetzung dafür, dass eine Aggregation bezüglich einer Kenngröße korrekt durchgeführt werden kann [Lenz&Shoshani]

Drei Bedingungen:

- 1) Disjunktheit → höchstens einmal
 - 2) Vollständigkeit → mindestens einmal
- } → **genau einmal**
- 3) Typverträglichkeit

Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Szenario mit Unsicherheiten

Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
PKW	X	22.04.05	15	1789	0,15	0,25	w ₁
PKW	X	22.04.05	15	2465	0,65	0,75	w ₂
PKW	X	22.04.05	15	2982	0,05	0,15	w ₃

Szenario mit Unsicherheiten

Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
PKW	X	22.04.05	15	1789	0,15	0,25	w ₁
PKW	X	22.04.05		2465	0,65	0,75	w ₂
PKW	X	22.04.05		2982	0,05	0,15	w ₃

Anzahl an Fahrzeugen

X

Nachfragesegment

Jahr

Monat

Tag

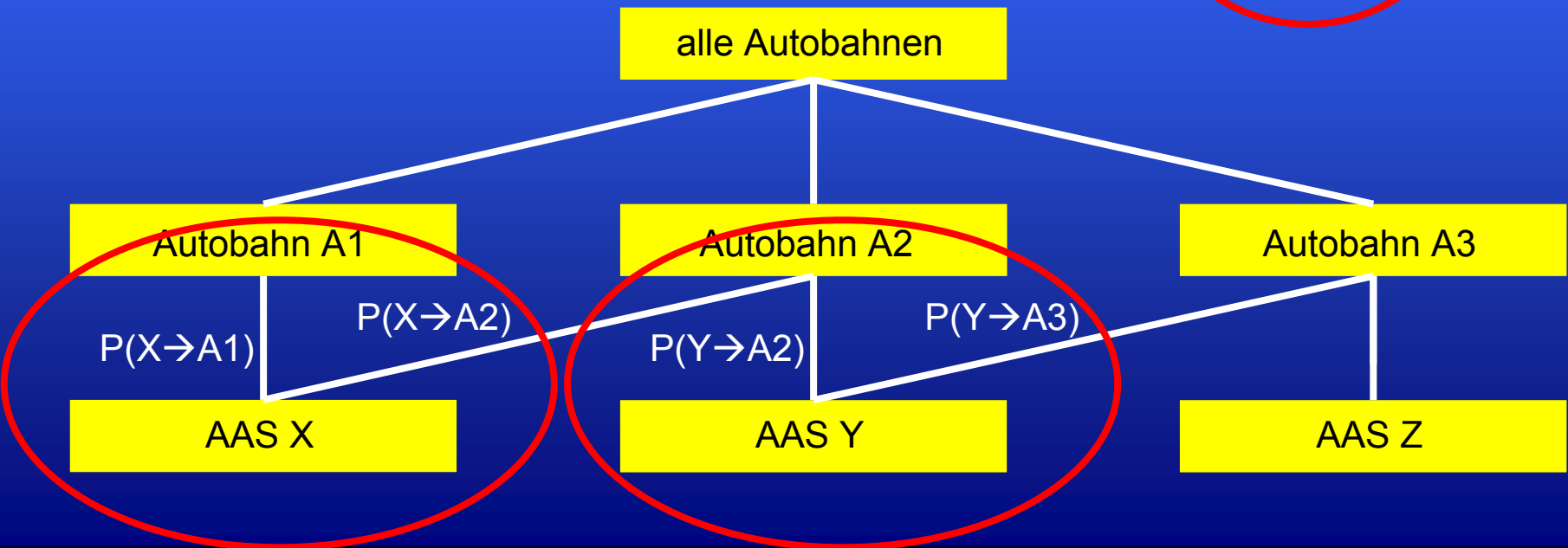
Stunde

Autobahn

Autobahnabschnitt

Szenario mit Unsicherheiten

Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
PKW	X	22.04.05	15	1789	0,15	0,25	w ₁
PKW	X	22.04.05	15	2465	0,65	0,75	w ₂
PKW	X	22.04.05	15	2982	0,05	0,15	w ₃



Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,0	0,1	w ₇
3	PKW	X	22.04.05	16	2711	0,55	0,65	w ₈
3	PKW	X	22.04.05	16	3034	0,2	0,3	w ₉

→ count((ID)(Prognose)) = 9 ↩

→ **Wahrscheinlichkeiten müssen berücksichtigt werden !**

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,0	0,1	w ₇
3	PKW	X	22.04.05	16	2711	0,55	0,65	w ₈
3	PKW	X	22.04.05	16	3034	0,2	0,3	w ₉

→ Gewichtung mit Mittelwerten der Wahrscheinlichkeitsintervalle

$$\begin{aligned} &\rightarrow \text{pcount}(\text{ID})(\text{Prognose}) = \\ &(0,25 + 0,65 + 0,1) + (0,2 + 0,7 + 0,1) + (0,05 + 0,6 + 0,25) = \\ &1 + 1 + 0,9 = 2,9 \quad \leftarrow \end{aligned}$$

Probabilistische Summierbarkeit

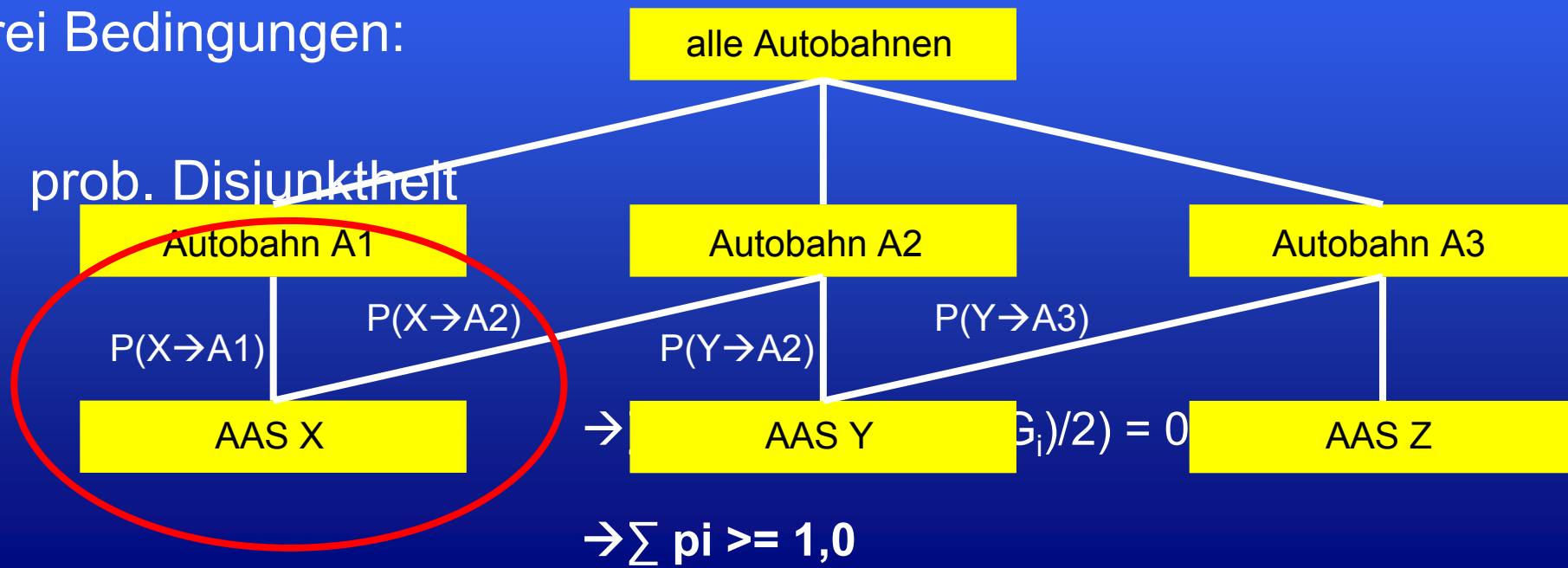
→ Voraussetzung dafür, dass eine Aggregation bezüglich einer Kenngröße korrekt durchgeführt werden kann [Lenz&Shoshani]

Probabilistische Summierbarkeit

→ Voraussetzung dafür, dass eine **probabilistische** Aggregation bezüglich einer Kenngröße korrekt **über unsicheren Daten** durchgeführt werden kann

Drei Bedingungen:

1) **prob. Disjunktheit**



Probabilistische Summierbarkeit

→ Voraussetzung dafür, dass eine **probabilistische** Aggregation
 bezüglich einer Kenngröße korrekt **über unsicheren Daten**
 durchgeführt werden kann

Drei Bedingungen:

- 1) prob. Disjunktheit
- 2) prob. Vollständigkeit

ID	Zeit	Fahrzeuge	UG	OG	Pfad
2	15	1789	0,15	0,25	w ₄
2	15	2465	0,65	0,75	w ₅
2	15	2982	0,05	0,15	w ₆

$$\rightarrow \sum p_i = \sum ((UG_i + OG_i)/2) = 0,2 + 0,7 + 0,1 = 1,0$$

$$\rightarrow \sum p_i \geq 1,0$$

Probabilistische Summierbarkeit

→ Voraussetzung dafür, dass eine **probabilistische** Aggregation bezüglich einer Kenngröße korrekt **über unsicheren Daten** durchgeführt werden kann

Drei Bedingungen:

- | | | |
|--------------------------|---------------------|---|
| 1) prob. Disjunktheit | → höchstens einmal | } → genau einmal
→ $\sum p_i = 1$ |
| | → $\sum p_i \leq 1$ | |
| 2) prob. Vollständigkeit | → mindestens einmal | |
| | → $\sum p_i \geq 1$ | |

Probabilistische Summierbarkeit

→ Voraussetzung dafür, dass eine **probabilistische** Aggregation bezüglich einer Kenngröße korrekt **über unsicheren Daten** durchgeführt werden kann

Drei Bedingungen:

- | | | | |
|--------------------------|---------------------|---|---|
| 1) prob. Disjunktheit | → höchstens einmal | } | → genau einmal
→ $\sum p_i = 1$ |
| | → $\sum p_i \leq 1$ | | |
| 2) prob. Vollständigkeit | → mindestens einmal | } | |
| | → $\sum p_i \geq 1$ | | |
| 3) Typverträglichkeit | | | |

Mittelwertbedingung

Die Summe der Mittelwerte der Wahrscheinlichkeitsintervalle aller Mengen von annotierten Tupeln einer Relation, die jeweils eine Entität repräsentieren, muss 1,0 ergeben!

- Voraussetzung für probabilistische Summierbarkeit
- Voraussetzung für korrekte Aggregation von unsicheren Daten

Auswirkungen auf das probabilistische Modell von Lakshmanan:

- Mittelwertbedingung muss gelten:
 - Einschränkung bei den Wahrscheinlichkeitsintervallen
- Schlüsselattribute müssen deterministisch sein
 - Zur Identifikation der Tupel, die eine Entität repräsentieren

Summierbarkeit im Beispiel

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,0	0,1	w ₇
3	PKW	X	22.04.05	16	2711	0,55	0,65	w ₈
3	PKW	X	22.04.05	16	3034	0,2	0,3	w ₉

→ ID = 1: $0,25 + 0,65 + 0,1 = 1,0$ ✓

→ ID = 2: $0,2 + 0,7 + 0,1 = 1,0$ ✓

→ ID = 3: $0,05 + 0,6 + 0,25 = 0,9$ ↯ → Normierung

Summierbarkeit im Beispiel

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

→ ID = 1: $0,25 + 0,65 + 0,1 = 1,0$ ✓

→ ID = 2: $0,2 + 0,7 + 0,1 = 1,0$ ✓

→ ID = 3: $0,055 + 0,667 + 0,278 = 1,0$ ✓

→ Mittelwertbedingung ist erfüllt

→ Summierbarkeit ist gegeben

Modifizierte Operatoren

- `pcount((ID)(Prognose)) =`

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

- $\text{pcount}(\text{ ID })(\text{ Prognose }) = (0,25+0,65+0,1) + (0,2+0,7+0,1) + (0,055+0,667+0,278) = 1 + 1 + 1 = 3$
- $\text{psum}(\text{ Fahrzeuge })(\text{ Prognose }) =$

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

- $$\text{pcount}(\text{ ID })(\text{ Prognose }) = (0,25+0,65+0,1) + (0,2+0,7+0,1) + (0,055+0,667+0,278) = 1 + 1 + 1 = 3$$
- $$\text{psum}(\text{ Fahrzeuge })(\text{ Prognose }) = (0,25*1554 + 0,65*2321 + 0,1*2658) + (0,2*1789 + 0,7*2465 + 0,1*2982) + (0,055*1589 + 0,667*2711 + 0,278*3034) = 388,5 + 1508,65 + 265,8 + 357,8 + 1725,5 + 298,2 + 87,395 + 1808,237 + 843,452 = 7283,434$$
- $$\text{max}(\text{ Fahrzeuge })(\text{ Prognose }) =$$

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

- $$\text{pcount}(\text{ID})(\text{Prognose}) = (0,25+0,65+0,1) + (0,2+0,7+0,1) + (0,055+0,667+0,278) = 1 + 1 + 1 = 3$$
- $$\text{psum}(\text{Fahrzeuge})(\text{Prognose}) = (0,25 \cdot 1554 + 0,65 \cdot 2321 + 0,1 \cdot 2658) + (0,2 \cdot 1789 + 0,7 \cdot 2465 + 0,1 \cdot 2982) + (0,055 \cdot 1589 + 0,667 \cdot 2711 + 0,278 \cdot 3034) = 388,5 + 1508,65 + 265,8 + 357,8 + 1725,5 + 298,2 + 87,395 + 1808,237 + 843,452 = 7283,434$$
- $$\text{max}(\text{Fahrzeuge})(\text{Prognose}) = 3034$$
- $$\text{max}_{\alpha=0,5}(\text{Fahrzeuge})(\text{Prognose}) =$$

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

- $\text{pcount}(\text{ID})(\text{Prognose}) = (0,25+0,65+0,1) + (0,2+0,7+0,1) + (0,055+0,667+0,278) = 1 + 1 + 1 = 3$
- $\text{psum}(\text{Fahrzeuge})(\text{Prognose}) = (0,25*1554 + 0,65*2321 + 0,1*2658) + (0,2*1789 + 0,7*2465 + 0,1*2982) + (0,055*1589 + 0,667*2711 + 0,278*3034) = 388,5 + 1508,65 + 265,8 + 357,8 + 1725,5 + 298,2 + 87,395 + 1808,237 + 843,452 = 7283,434$
- $\text{max}(\text{Fahrzeuge})(\text{Prognose}) = 3034$
- $\text{max}_{\alpha=0,5}(\text{Fahrzeuge})(\text{Prognose}) = 2711$
- $\text{pavg}(\text{Fahrzeuge})(\text{Prognose}) = \text{psum}(\text{Fahrzeuge})(\text{Prognose}) / \text{pcount}(\text{ID})(\text{Prognose}) = 7283,434 / 3 = 2427,81$

Modifizierte Operatoren

Statt Gewichtung mit Wahrscheinlichkeiten bei Summe:
→ Kombination aller möglichen Summen in mehreren Bereichen

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

ID	Segment	AAS	Datum	Zeit	Fahrzeuge	UG	OG	Pfad
1	PKW	X	22.04.05	14	1554	0,2	0,3	w ₁
1	PKW	X	22.04.05	14	2321	0,6	0,7	w ₂
1	PKW	X	22.04.05	14	2658	0,05	0,15	w ₃
2	PKW	X	22.04.05	15	1789	0,15	0,25	w ₄
2	PKW	X	22.04.05	15	2465	0,65	0,75	w ₅
2	PKW	X	22.04.05	15	2982	0,05	0,15	w ₆
3	PKW	X	22.04.05	16	1589	0,005	0,105	w ₇
3	PKW	X	22.04.05	16	2711	0,617	0,717	w ₈
3	PKW	X	22.04.05	16	3034	0,228	0,328	w ₉

Modifizierte Operatoren

Σ (Fahrzeuge)	P	Pfad
4932	0,00275	$w_1 + w_4 + w_7$
6054	0,03335	$w_1 + w_4 + w_8$
6377	0,0139	$w_1 + w_4 + w_9$
5608	0,009625	$w_1 + w_5 + w_7$
6730	0,116725	$w_1 + w_5 + w_8$
7053	0,04865	$w_1 + w_5 + w_9$
6125	0,001375	$w_1 + w_6 + w_7$
7247	0,016675	$w_1 + w_6 + w_8$
7570	0,00695	$w_1 + w_6 + w_9$

Modifizierte Operatoren

Statt Gewichtung mit Wahrscheinlichkeiten bei Summe:

- Kombination aller möglichen Summen in mehreren Bereichen
- hier drei Bereiche, Einteilung frei wählbar
- Ergebnisse weiterhin mit Wahrscheinlichkeitswert
- kann auch für weitere Aggregationen verwendet werden
 - Mittelwertbedingung ist erfüllt

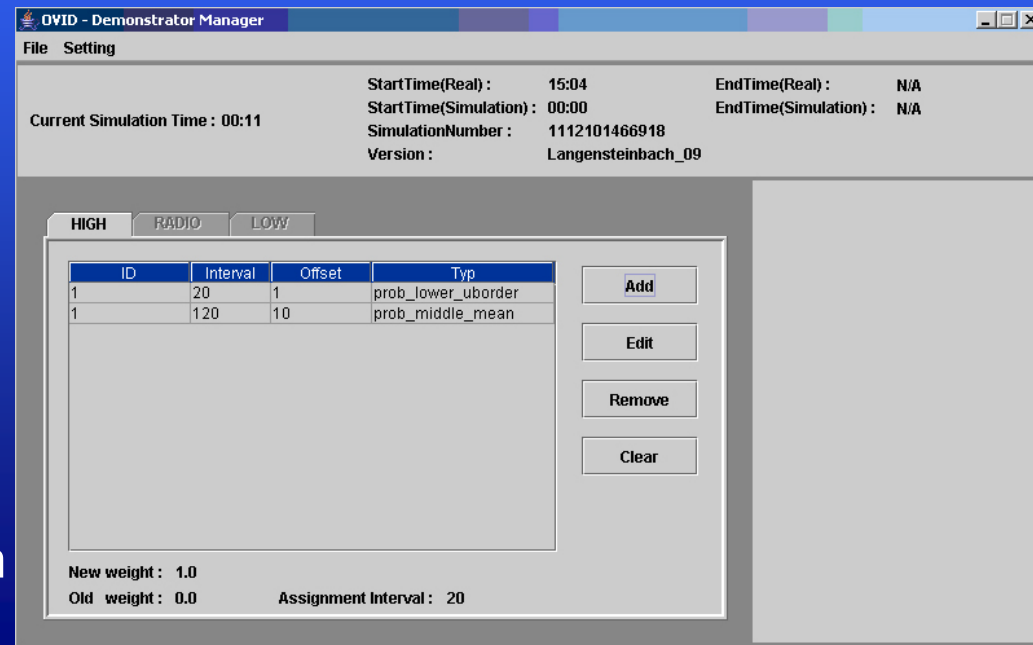
$\Sigma(\text{Fahrzeuge})$	B-UG	B-OG	P
5901	0	6500	0,094275
7128	6501	7600	0,64221
8170	7601	∞	0,263515

Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Einbettung in den Demonstrator

- IPD für Verwaltung der simulierten Verkehrsdaten verantwortlich
→ Verkehrsprognosen in Form von O-D-Matrizen
- Neue Prognosealgorithmen: O-D-Matrizen mit global zugeordneten Wahrscheinlichkeitsintervallen
- komplett in bestehende Software-Architektur integriert
- Problem, dass andere Module noch keine probabilistischen O-D-Matrizen verarbeiten können
- langfristige Prognosen auf Basis von historischen Messdaten
→ können für kurzfristige Prognosen herangezogen werden



OVID - Demonstrator Manager
 File Setting

Current Simulation Time : 00:11
 StartTime(Real) : 15:04 EndTime(Real) : N/A
 StartTime(Simulation) : 00:00 EndTime(Simulation) : N/A
 SimulationNumber : 1112101466918
 Version : Langensteinbach_09

ID	Interval	Offset	Typ
1	20	1	prob_lower_uborder
1	120	10	prob_middle_mean

Add
 Edit
 Remove
 Clear

New weight : 1.0
 Old weight : 0.0 Assignment Interval : 20

Prototypische Implementierung

- modifizierte Operatoren prototypisch implementiert
 - für die Aggregationen: Anzahl, Summe, Minimum, Maximum, Durchschnitt
- vor dem Hintergrund des OVID-Szenarios
 - Verkehrsdaten aus mehreren Simulationsläufen können aggregiert werden
- prototypische Implementierung und Integration in Demonstrator in Java

Überblick

- **Unsichere Verkehrsdaten**
- **Modelle für unsichere Daten**
- **Aggregationen im
multidimensionalen Modell**
- **Aggregationsoperatoren für
unsichere Daten**
- **Umsetzung in die Praxis**
- **Zusammenfassung**

Zusammenfassung

- Untersuchung und Bewertung verschiedener probabilistischer relationaler Modelle
- Summierbarkeit auf probabilistischen Fall erweitert
→ Voraussetzung für Aggregation von unsicheren Daten
- Aggregationsoperatoren an unsichere Daten angepasst
→ prototypische Implementierung
- probabilistische Prognose in OVID-Demonstrator von IPD-Seite integriert
→ Verarbeitung in anderen Modulen noch nicht möglich

Vielen Dank für Ihre Aufmerksamkeit!

Fragen?