

# **Konzeption und Umsetzung einer Schemaerweiterung durch Kontexte unter Berücksichtigung von Aggregierungsanfragen**

Studienarbeit von

Andreas Merkel

28.05.2004

# Gegenstand der Studienarbeit

---

- **Anwendung des Kontextmodells auf Verkehrsdaten**
  
- **Aggregierungsanfragen sollen weiterhin möglich sein**

# Übersicht

---

- **Das Kontextmodell**
  - scharfe und unscharfe Kontexte
  - Klassenbildung
- **Aggregation**
  - Aggregatfunktionen für unscharfe Klassen
  - Aggregationsbedingungen
- **Umsetzung im relationalen Datenbanksystem**
  - Repräsentation von Kontexten
- **Anwendung auf Verkehrsdaten**
  - Definition von Verkehrszuständen

# Das Kontextmodell

---

- **Erweiterung des relationalen Datenmodells**
- **Vorgestellt von Günther Schindler:**  
*Fuzzy Datenanalyse durch kontextbasierte Datenbankabfragen* (Dissertation, 1998)
- **Dient zur Klassifikation von Daten**  
Ermöglicht unscharfes Klassifizieren scharfer Daten

# scharfe Kontexte

---

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

# scharfe Kontexte

---

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

Geschwindigkeit (km/h)

0 10 20 30 40 50 60 70 80 90 100 120 130 140

# scharfe Kontexte

---

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

Geschwindigkeit (km/h)

0	10	20	30	40	50	60	70	80	90	100	120	130	140
---	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----

# scharfe Kontexte

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

		Geschwindigkeit (km/h)													
		0	10	20	30	40	50	60	70	80	90	100	120	130	140
Stillstandsanteil (%)	0														
	20														
	40														
	60														
	80														
	100														

# scharfe Kontexte

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

		Geschwindigkeit (km/h)													
		0	10	20	30	40	50	60	70	80	90	100	120	130	140
Stillstandsanteil (%)	0	Klasse 1							Klasse 2						
	20	Klasse 1							Klasse 2						
	40	Klasse 1							Klasse 2						
Stillstandsanteil (%)	60	Klasse 3							Klasse 4						
	80	Klasse 3							Klasse 4						
	100	Klasse 3							Klasse 4						

**-> Einteilung der Relation in Klassen**

# scharfe Kontexte

**scharfer Kontext =  
Einteilung des Wertebereichs eines Attributs in  
Äquivalenzklassen**

		Geschwindigkeit (km/h)													
		0	10	20	30	40	50	60	70	80	90	100	120	130	140
Stillstandsanteil (%)	0	Klasse 1							Klasse 2						
	20	zähfließender Verkehr							freier Verkehr						
Stillstandsanteil (%)	40														
	60	Klasse 3							Klasse 4						
	80	Stau							stockender Verkehr						
100															

**-> Einteilung der Relation in Klassen**

# unscharfe Kontexte

---

**Problem: oft lassen sich Objekte nicht eindeutig einer Klasse zuordnen -> unscharfe Klassifikation**

# unscharfe Kontexte

**Problem: oft lassen sich Objekte nicht eindeutig einer Klasse zuordnen -> unscharfe Klassifikation**

linguistische  
Variable

Staulänge

Terme

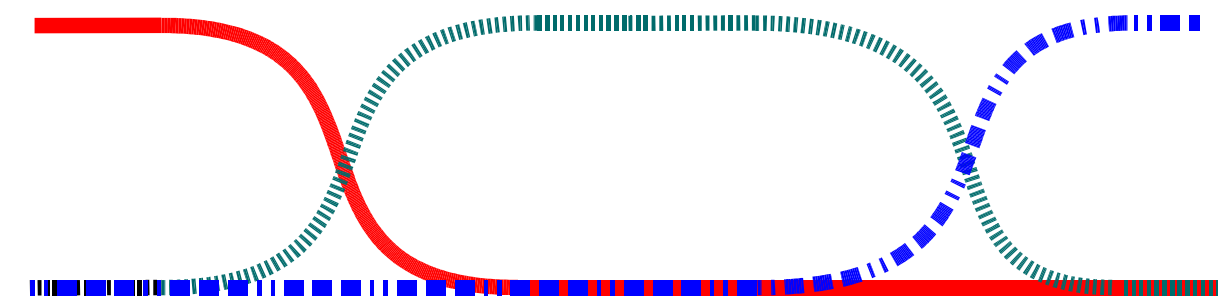
kurz

mittel

lang

Zugehörigkeits-  
funktionen

1,00  
0,75  
0,50  
0,25  
0,00



D(Staulänge)

1 2 3 4 5 6 7 8 9 10 11 12 13 14

# Aggregation

---

**Wie kann über unscharfe Klassen aggregiert werden?**

**-> Erweiterung der Aggregatfunktionen**

**Idee: Zugehörigkeiten als Gewichtungsfaktoren nutzen**

**-> entspricht intuitivem Verständnis**

# Aggregation

---

**Wie kann über unscharfe Klassen aggregiert werden?**

**-> Erweiterung der Aggregatfunktionen**

**Idee: Zugehörigkeiten als Gewichtungsfaktoren nutzen**

- *unscharfe Anzahl*: Summe der Zugehörigkeiten
- *unscharfe Summe*: gewichtete Summe
- *unscharfer Durchschnitt*: Quotient aus unscharfer Summe und unscharfer Anzahl

**-> entspricht intuitivem Verständnis**

# Aggregation

---

## Aggregationsbedingungen

- Überlappungsfreiheit
- Vollständigkeit
- Typvereinbarkeit

# Aggregation

---

## Aggregationsbedingungen

- Überlappungsfreiheit
- Vollständigkeit
- Typvereinbarkeit

sorgen dafür, dass jeder Wert genau einmal berücksichtigt wird

# Aggregation

---

## Aggregationsbedingungen

- Überlappungsfreiheit
- Vollständigkeit
- Typvereinbarkeit

sorgen dafür, dass jeder Wert genau einmal berücksichtigt wird

**-> bei scharfen Klasseneinteilungen immer erfüllt,  
da Äquivalenzklassen**

# Aggregation

---

## Aggregationsbedingungen

- Überlappungsfreiheit
- Vollständigkeit
- Typvereinbarkeit

sorgen dafür, dass jeder Wert genau einmal berücksichtigt wird

**-> bei scharfen Klasseneinteilungen immer erfüllt, da Äquivalenzklassen**

**Für unscharfe Klassen: Modifikation der Bedingungen**  
Summe der Zugehörigkeiten eines Tupels zu allen Klassen muss eins ergeben.

# Umsetzung im relationalen Datenbanksystem

---

- durch zusätzliche Attribute

- durch eine Relation

- durch eine Funktion

# Umsetzung im relationalen Datenbanksystem

---

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	Äquivalenzklasse
5	12	A
42	97	B

## - durch eine Relation

## - durch eine Funktion

# Umsetzung im relationalen Datenbanksystem

---

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	niedrig	hoch
5	12	0.87	0.13
42	97	0.05	0.95

## - durch eine Relation

## - durch eine Funktion

# Umsetzung im relationalen Datenbanksystem

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	niedrig	hoch
5	12	0.87	0.13
42	97	0.05	0.95

## - durch eine Relation

Geschwindigkeit	Äquivalenzklasse
0	A
1	A
...	

## - durch eine Funktion

# Umsetzung im relationalen Datenbanksystem

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	niedrig	hoch
5	12	0.87	0.13
42	97	0.05	0.95

## - durch eine Relation

Geschwindigkeit	niedrig	hoch
0	1.00	0.00
1	0.99	0.01
...		

## - durch eine Funktion

# Umsetzung im relationalen Datenbanksystem

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	niedrig	hoch
5	12	0.87	0.13
42	97	0.05	0.95

## - durch eine Relation

Geschwindigkeit	niedrig	hoch
0	1.00	0.00
1	0.99	0.01
...		

## - durch eine Funktion

*äquivalenzklasse(12)=A*

# Umsetzung im relationalen Datenbanksystem

## - durch zusätzliche Attribute

Fahrzeug-ID	Geschwindigkeit	niedrig	hoch
5	12	0.87	0.13
42	97	0.05	0.95

## - durch eine Relation

Geschwindigkeit	niedrig	hoch
0	1.00	0.00
1	0.99	0.01
...		

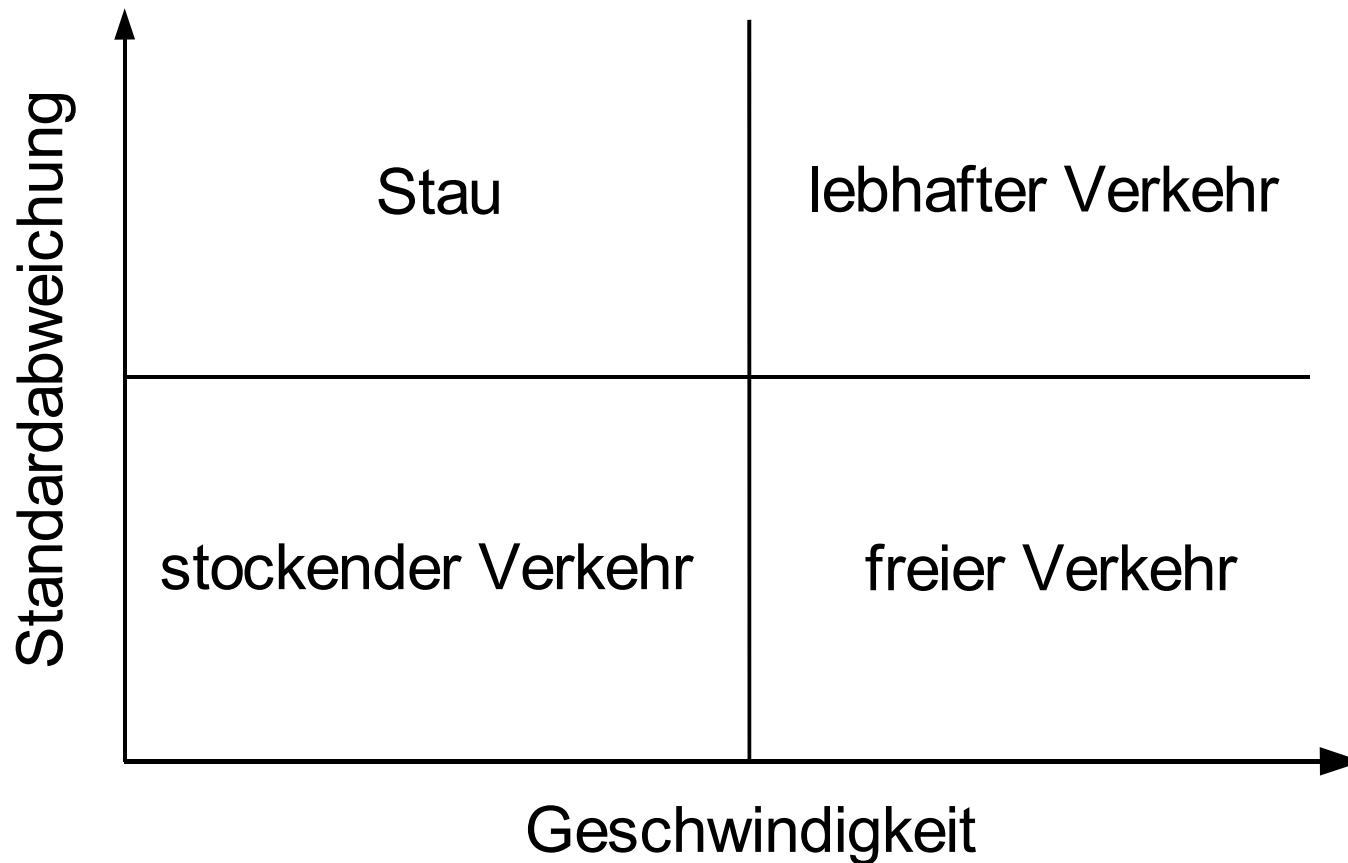
## - durch eine Funktion

*niedrig(12)=0.87*

*hoch(12)=0.13*

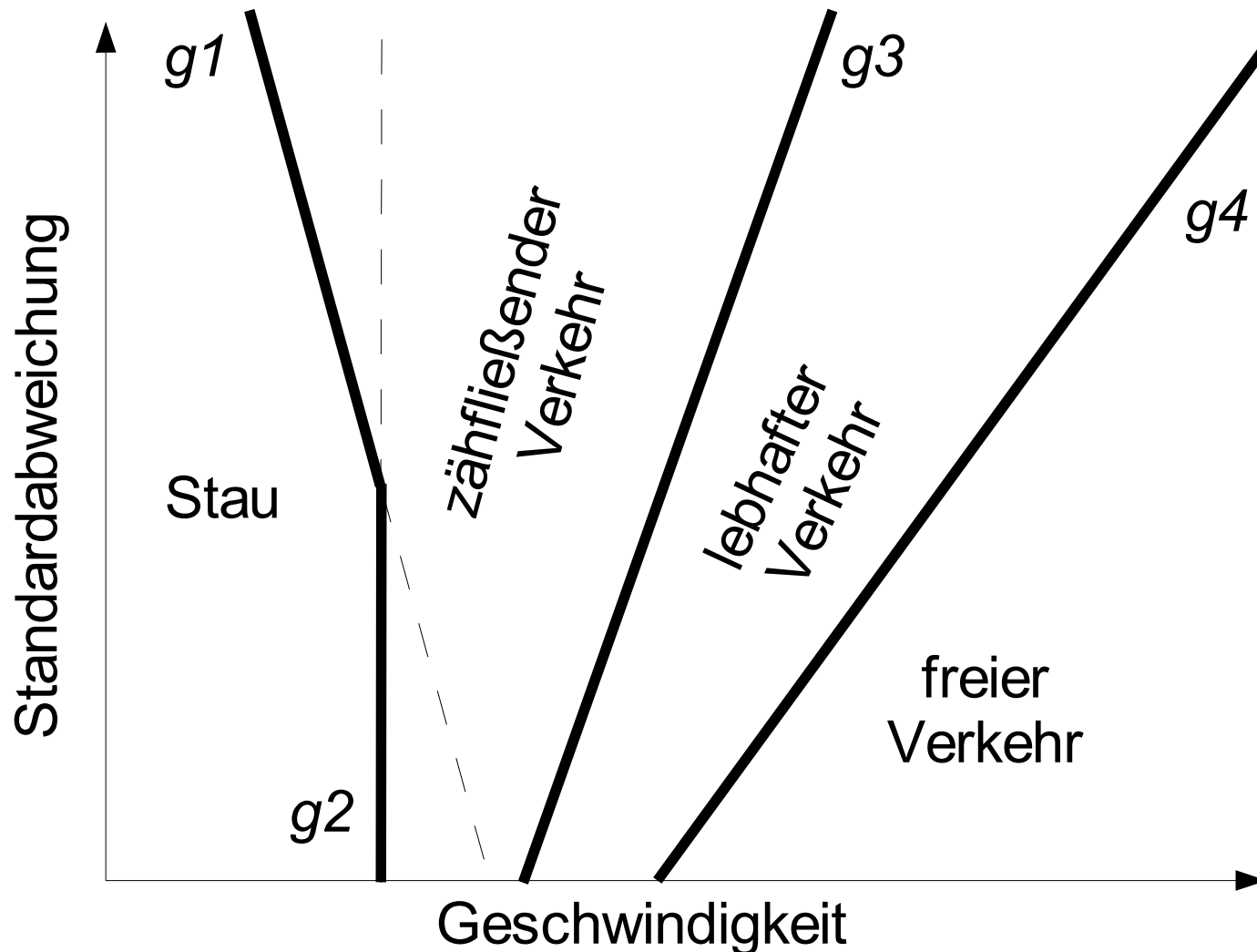
# Anwendung auf Verkehrsdaten

## Floating Car Daten: Definition von Verkehrszuständen (Beispiel: Autobahn)



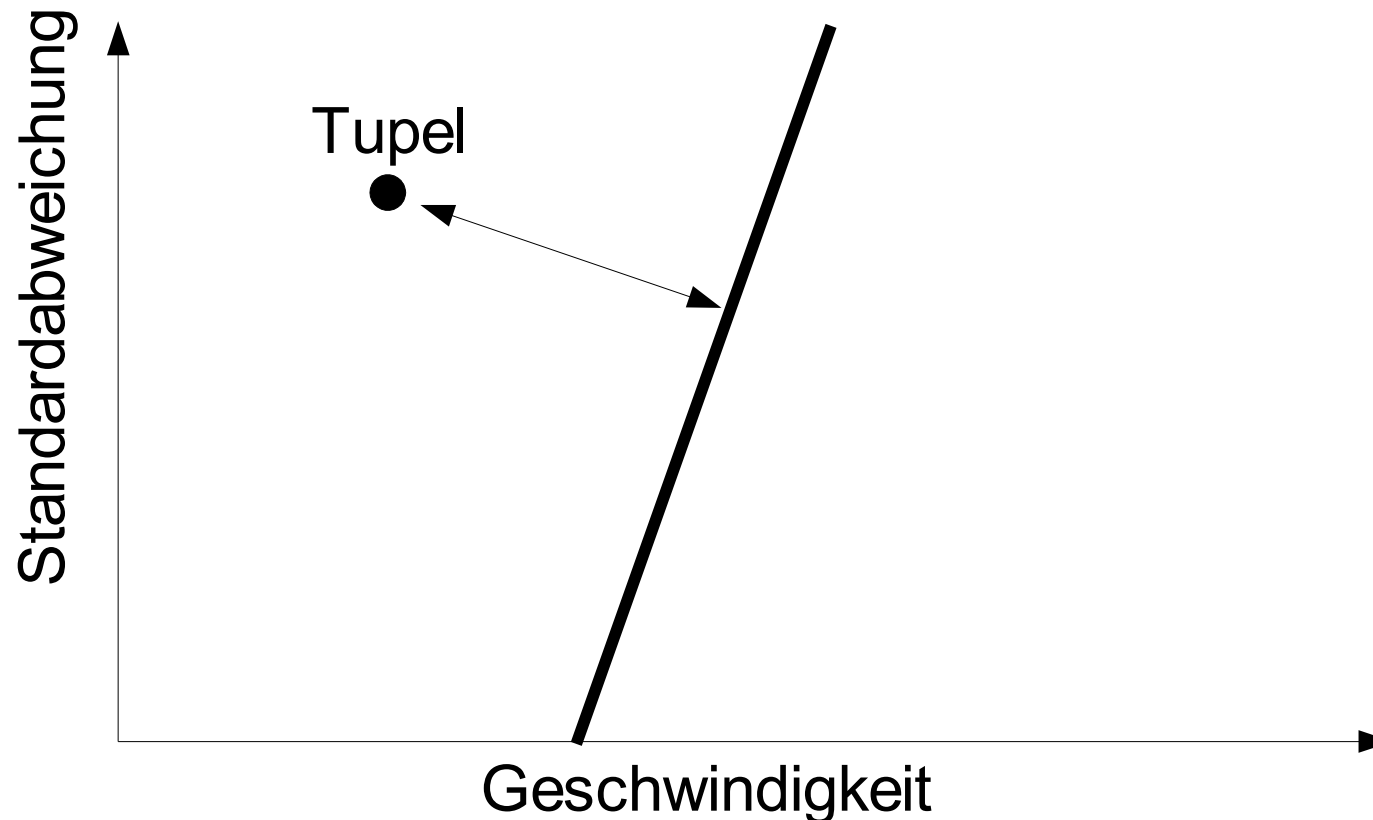
# Anwendung auf Verkehrsdaten

Flexiblere Einteilung durch beliebig verlaufende Geraden:



# Anwendung auf Verkehrsdaten

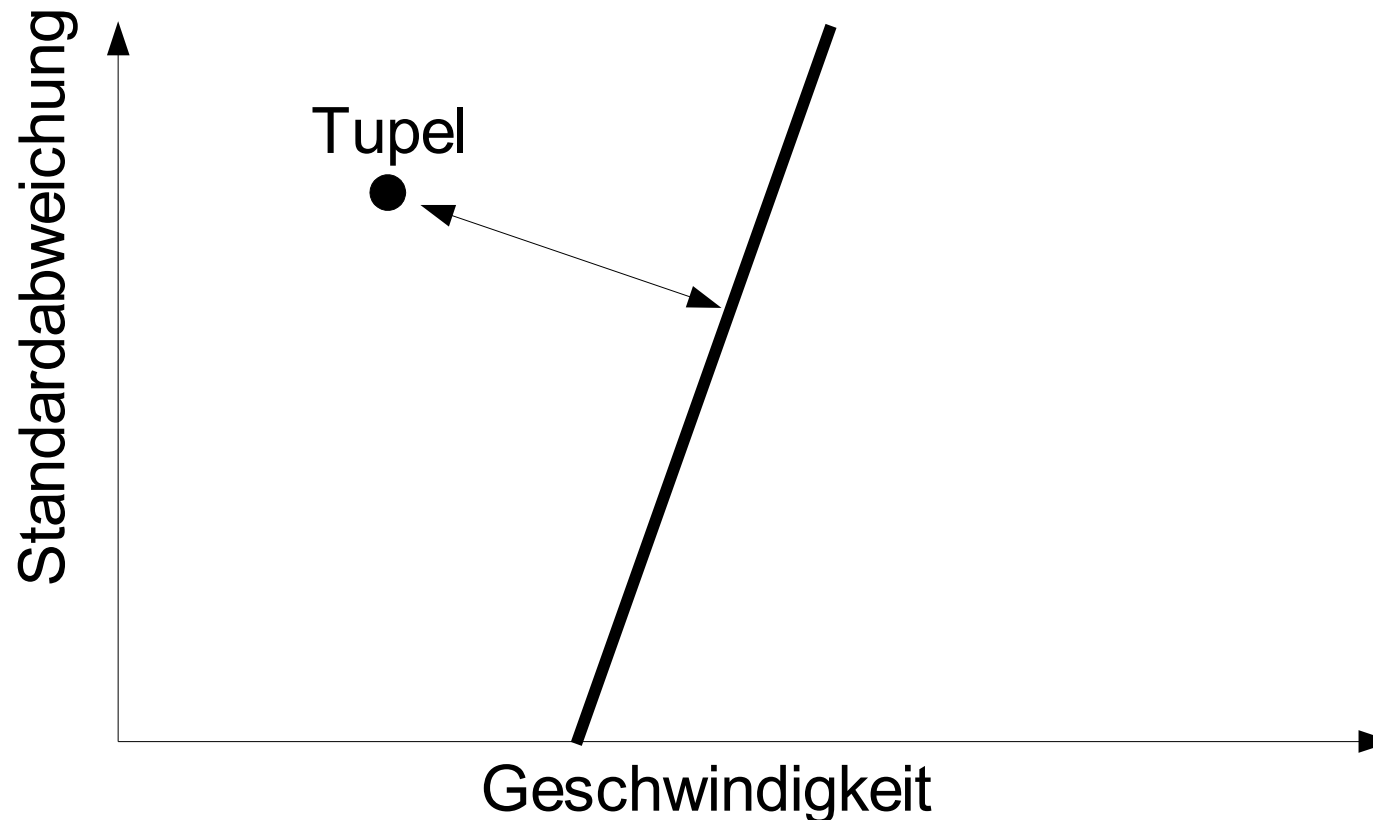
## Zwischenschritt: Einführung zusätzlicher Abstandsattribute



# Anwendung auf Verkehrsdaten

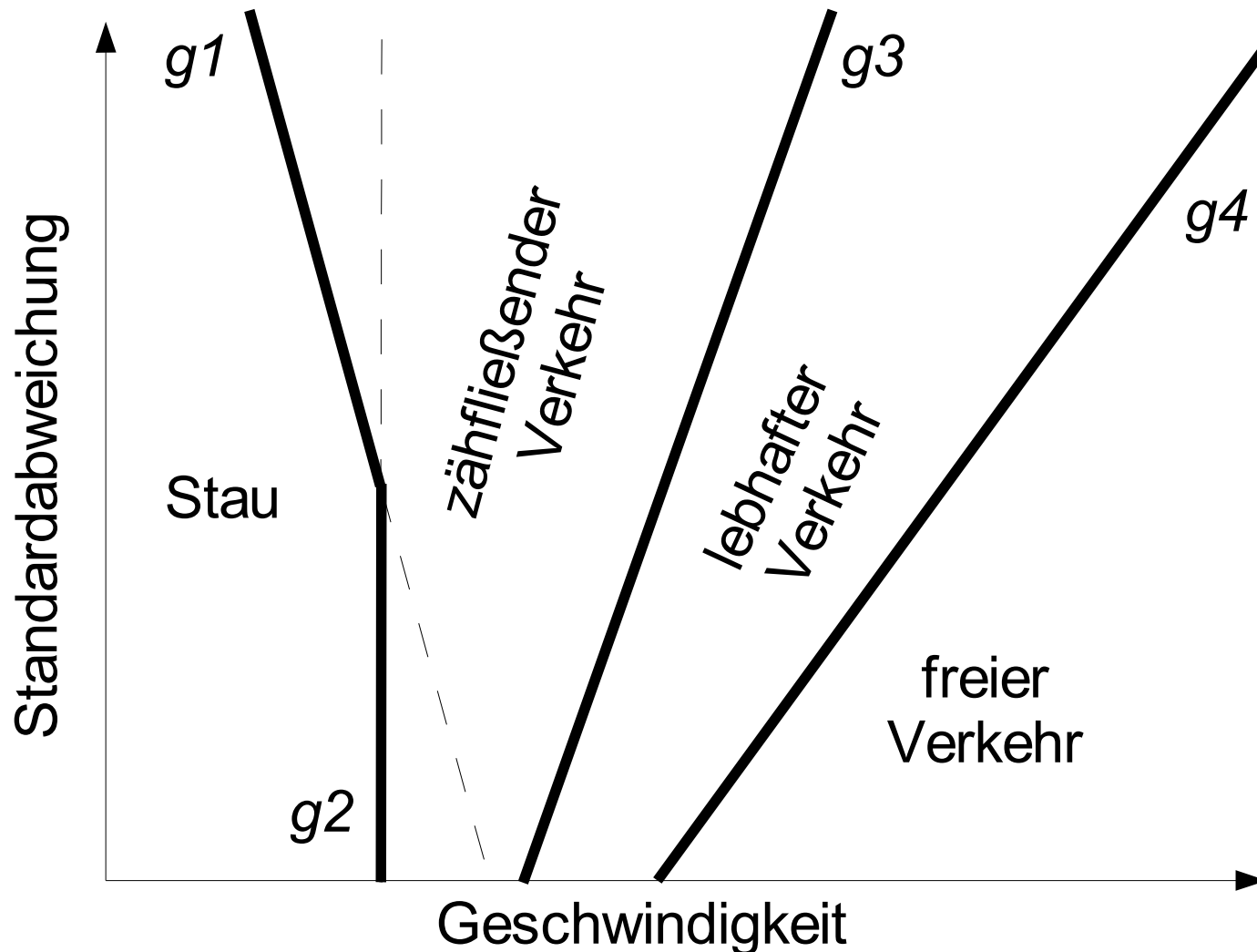
## Zwischenschritt: Einführung zusätzlicher Abstandsattribute

-> linguistische Variable mit den Termen *links* und *rechts*



# Anwendung auf Verkehrsdaten

**Flexiblere Einteilung durch beliebig verlaufende Geraden:**



# Anwendung auf Verkehrsdaten

---

- **Floating Car Daten** ✓

# Anwendung auf Verkehrsdaten

---

- **Floating Car Daten** ✓
- **Simulationsdaten für Verkehrsmessstationen**  
Querschnittsmessungen  
-> Bestimmung des Verkehrszustandes ähnlich wie bei  
Floating Car Daten

# Anwendung auf Verkehrsdaten

---

- **Floating Car Daten** ✓
- **Simulationsdaten für Verkehrsmessstationen**  
Querschnittsmessungen  
-> Bestimmung des Verkehrszustandes ähnlich wie bei Floating Car Daten
- **Güterverkehrsdaten**  
Statistik über den gesamten Güterverkehr in Deutschland gemeldeter Fahrzeuge  
-> Nutzung von Kontexten zur gröberen Strukturierung  
-> Berechnung von Aggregaten

# Zusammenfassung

---

## - Kontextmodell

scharfer Kontext: Einteilung in Äquivalenzklassen  
unscharfer Kontext: linguistische Variable

## - Aggregation

Erweiterung der Aggregatfunktionen für unscharfe Klassen

Aggregationsbedingungen:

- für scharfe Einteilungen immer erfüllt
- für unscharfe Einteilungen muss Summe der Zugehörigkeiten eins ergeben

# Zusammenfassung

---

## - **Umsetzung im relationalen Datenbanksystem**

- durch zusätzliche Attribute
- durch eine Relation
- durch Funktionen

## - **Anwendung auf Verkehrsdaten**

- Floating Car Daten
- Simulationsdaten für Verkehrsmessstationen
- Güterverkehrsdaten