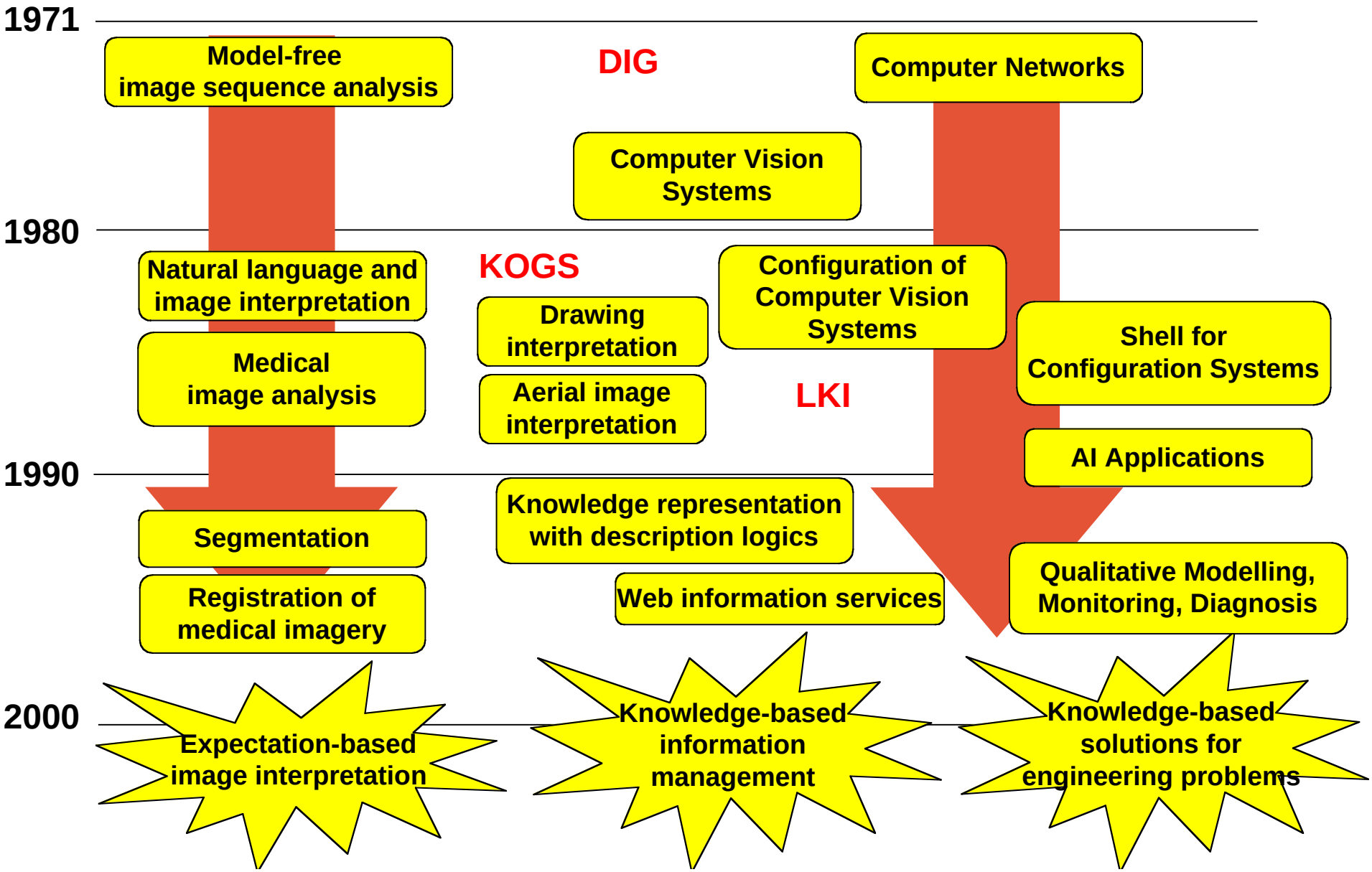
HIGH-LEVEL VISION
Introduction

Bernd Neumann

Cognitive Systems Laboratory
Hamburg University
Germany

neumann@informatik.uni-hamburg.de
http://kogs-www.informatik.uni-hamburg.de

CSL History


1971
1980
1990
2000
Model-free image sequence analysis
DIG
Computer Networks
Computer Vision Systems
Natural language and image interpretation
KOGS
Configuration of Computer Vision Systems
Drawing interpretation
Shell for Configuration Systems
Medical image analysis
Aerial image interpretation
LKI
AI Applications
Knowledge representation with description logics
Segmentation
Registration of medical imagery
Web information services
Qualitative Modelling, Monitoring, Diagnosis
Expectation-based image interpretation
Knowledge-based information management
Knowledge-based solutions for engineering problems

Current Research at CSL in Hamburg

- **Computer Vision**
  High-level image interpretation
  Segmentation
  Aerial image analysis
  Medical image analysis
- **Knowledge Representation**
  Foundations of description logics
  Spatial reasoning for GIS applications
  Optimized reasoning algorithms (RACER)
- **AI applications in technical domains**
  Configuration tools
  Software configuration
  Model-based diagnosis
  Automatic computation of decision trees

Cognitive Computer Vision

http://www.ecvision.info/home/Home.htm

Cognitive computer vision is concerned with integration and control of vision systems using explicit but not necessarily symbolic models of context, situation and goal-directed behaviour. Cognitive vision implies functionalities for knowledge representation, learning, reasoning about events & structures, recognition and categorization, and goal specification, all of which are concerned with the semantics of the relationship between the visual agent and its environment.

- **integration and control**
- **explicit models**
- **not necessarily symbolic**
- **context**
- **situation**
- **goal-directed behaviour**

- **knowledge representation**
- **learning**
- **reasoning**
- **recognition**
- **categorization**
- **goal specification**
- **visual agent**

High-level Scene Interpretation

High-level scene interpretation is the task of "understanding" a scene beyond single-object recognition. Typical examples are traffic scene understanding for driver assistance, inferring user intentions in smart-room scenarios, recognizing team behavior in robocup games, discovering criminal acts in monitoring tasks.

Characteristics:

- **Interpretations involve several objects and occurrences.**
- **Interpretations depend on temporal and spatial relations between parts of a scene**
- **Interpretations describes the scene in qualitative terms, omitting geometric details.**
- **Interpretations include inferred facts, unobservable in the scene.**
- **Interpretations are based on conceptual knowledge and experience about the world.**

Examples for High-level Scene Understanding

Garbage collection in Hamburg
(1 frame of a sequence)

Note

- **parts**
- **activities**
- **intentions**
- **spatial relations**
- **temporal relations**

Buster Keaton in "The Navigator"

Note

- **episodes**
- **story**
- **emotions**
- **funnyness**

Role of viewer knowledge and reasoning

Challenge Areas of Cognitive Vision

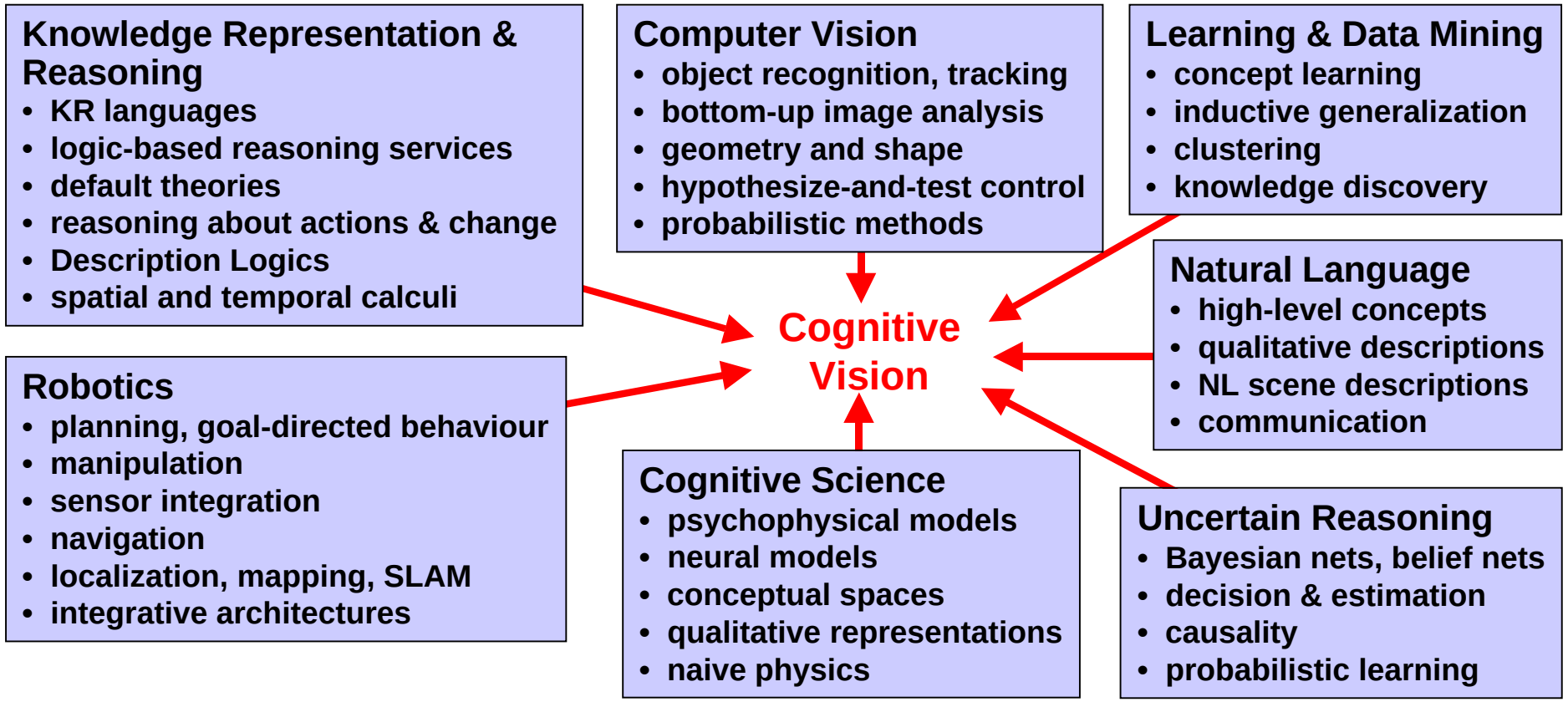
High-level scene interpretation
- conceptual descriptions of multi-object dynamic scenes
- temporal and spatial reasoning
- connecting to common-sense knowledge
- interpretation strategies

Active Vision
- vision and acting
- task-oriented vision
- control of attention
- embodied vision

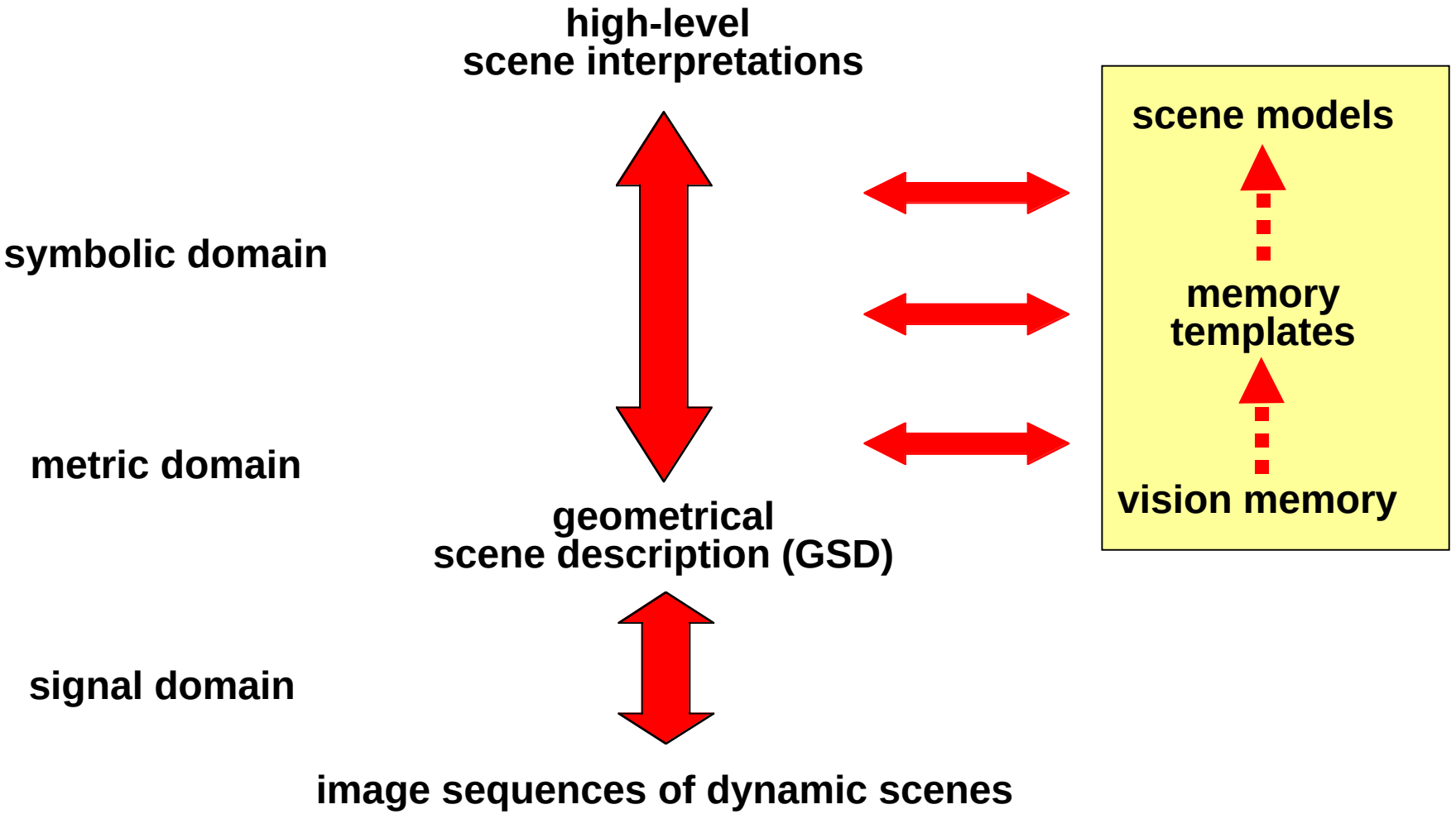
Vision and learning
- vision memory
- predicting from experience
- discovering reoccurring patterns ("suspicious coincidences")

Multidisciplinary Contributions to Cognitive Vision

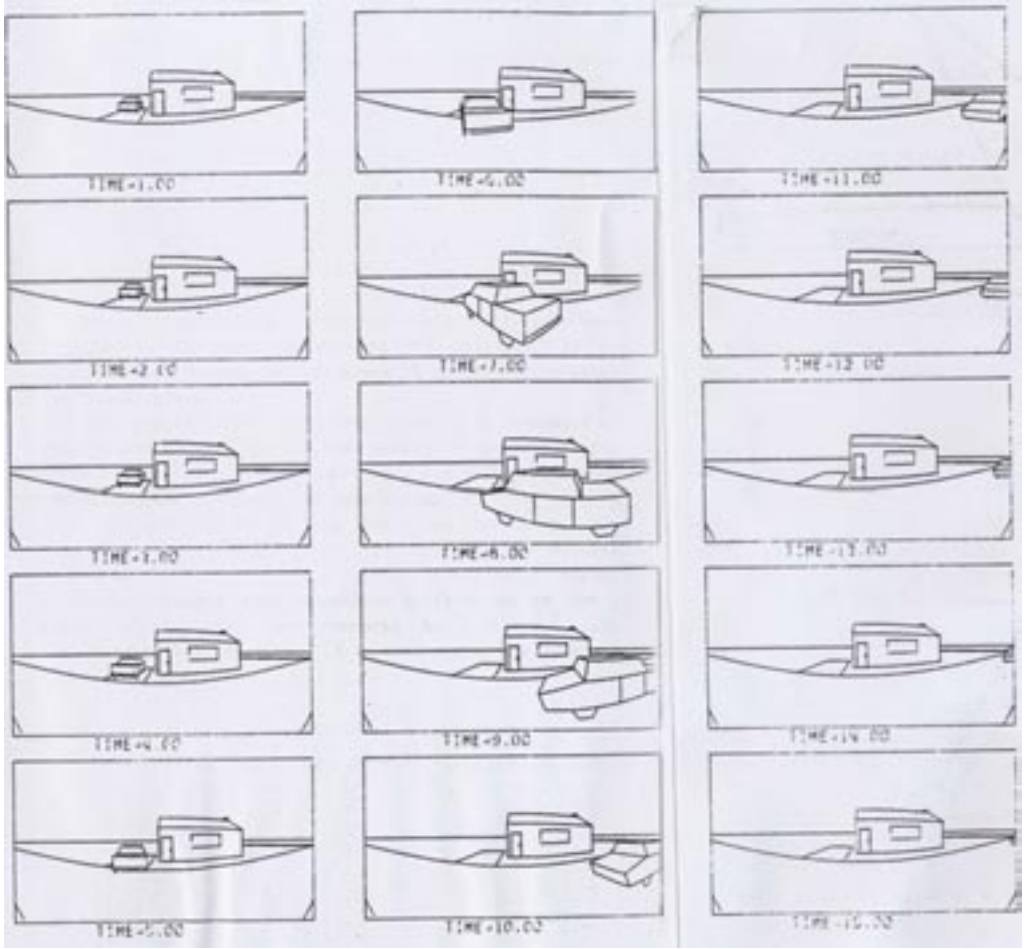
Cognitive Vision research requires multidisciplinary efforts and escape from traditional research community boundaries.



Basic building blocks for high-level scene interpretation

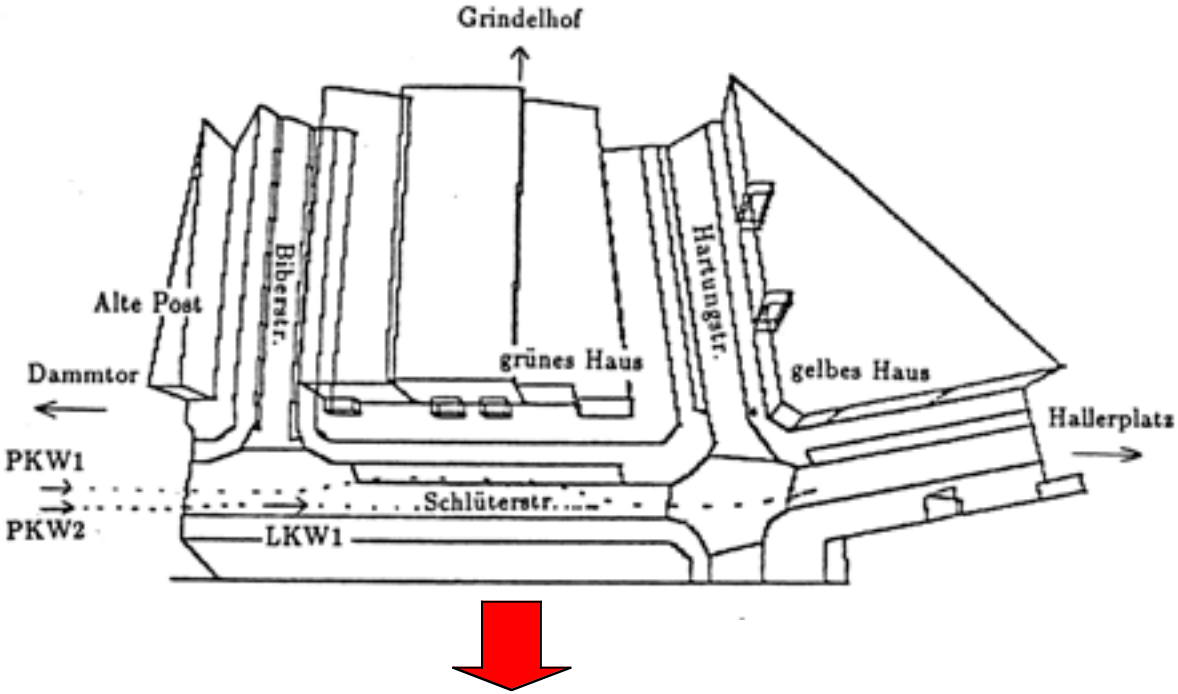


Early Traffic Scene Analysis (Badler 75)



motion description in terms of qualitative spatial relations, e.g. "moving across"

Natural Language Description of Traffic Scenes (Neumann & Novak 86)



DIE SZENE ENTHAELT DREI BEWEGTE OBJEKTE: ZWEI PKWS UND EINEN LKW.

EIN GELBER PKW FAEHRT IN RICHTUNG HALLERPLATZ. DABEI UEBERHOLT ER DEN LKW AUF DER SCHLUETERSTRASSE. DER GELBE PKW RAST VON DER ALTEN POST VOR DAS GELBE HAUS. ER ERREICHT DIE HARTUNGSTRASSE. ER HAELT AN. ER HAELT.

EIN SCHWARZER PKW ERREICHT DIE SCHLUETERSTRASSE. ER NAEHERT SICH DEM LKW VON DER ALTEN POST. DER SCHWARZE PKW FAEHRT IN RICHTUNG HALLERSTRASSE.

DER LKW FAEHRT VON DER ALTEN POST VOR DAS GRUENE HAUS. DABEI STOPPT ER VOR IHM. ER HAELT. ER FAEHRT IN RICHTUNG DAMMTOR WEITER. ER ENTFERNT SICH VON DEM GELBEN PKW. DER LKW HAELT AN. ER HAELT.

Occurrence Recognition (Hongeng 02)

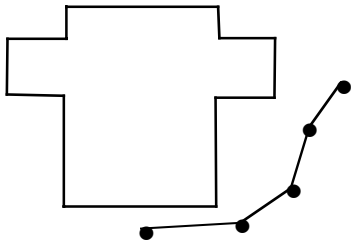
recognizing assaults

recognizing thefts at a phonebooth

Signal-Symbol Problems (1)

Mapping from quantitative into qualitative representations

Example:

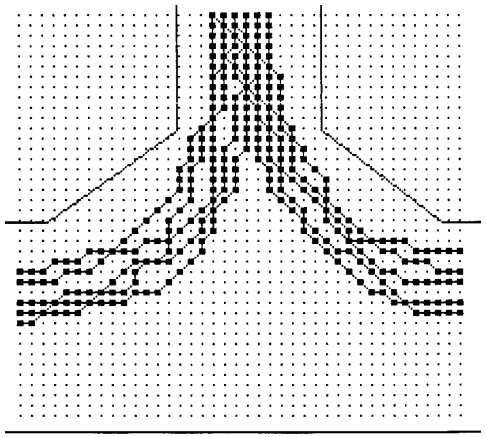


"along"

"Citytour" (Wahlster 87)

Mapping from qualitative into quantitative representations

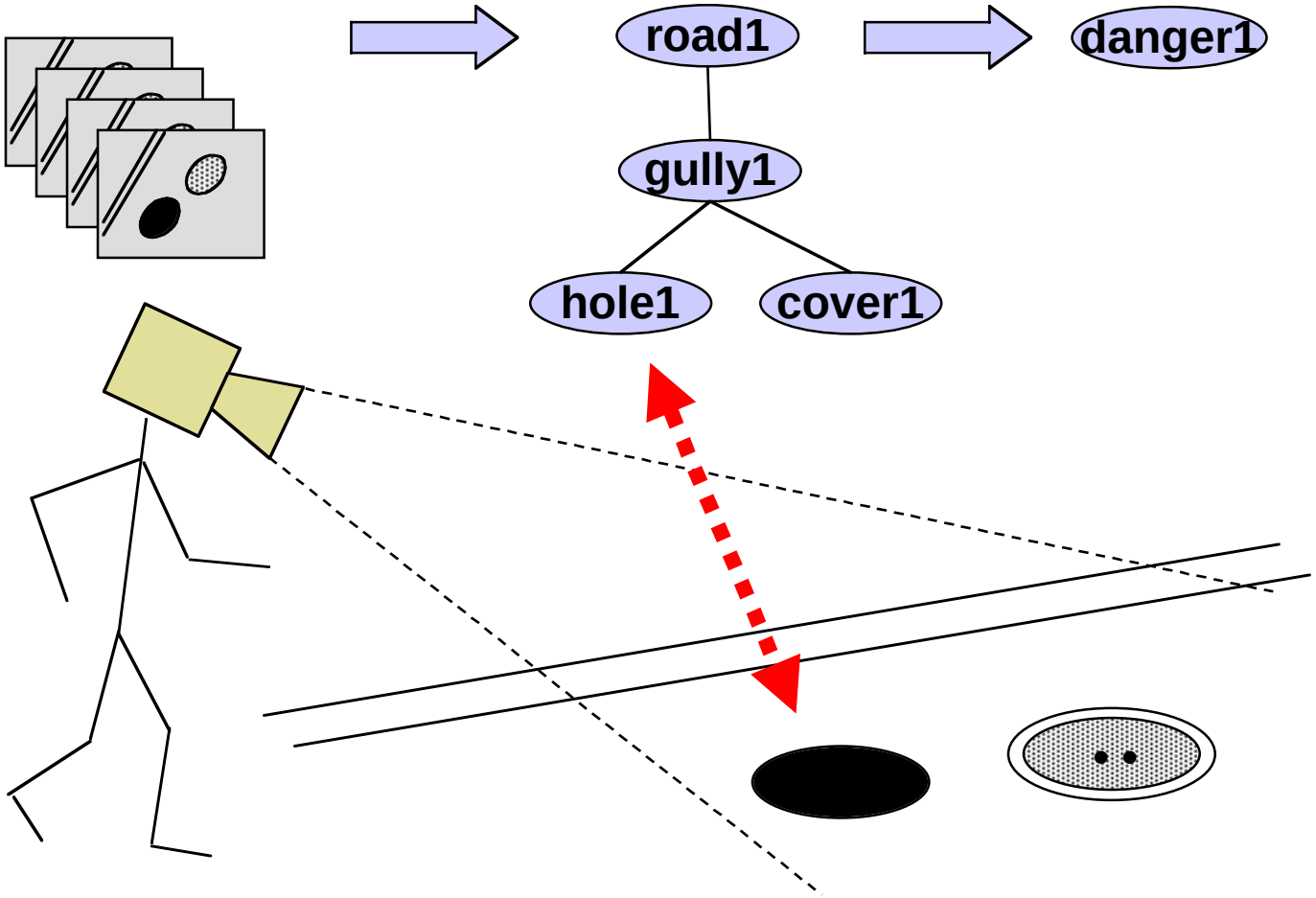
Example: ***"abbiegen"***



"typicality fields" (Mohnhaupt & Neumann 91)

Signal-Symbol Problems (2)

Symbol grounding



Signal-Symbol Problems (3)

Grounded symbolic reasoning

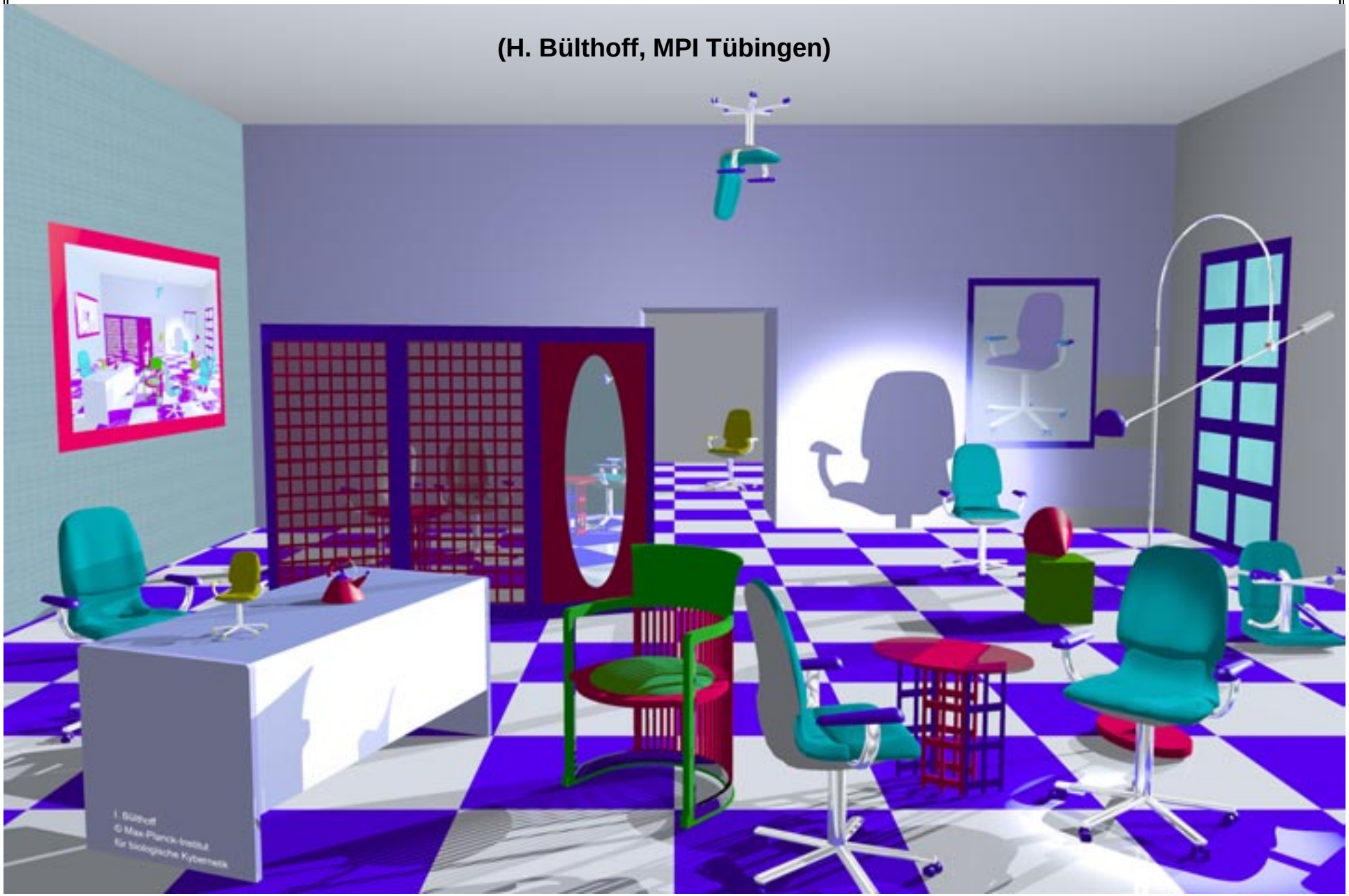
Deductions from symbolic knowledge about a scene should not only be correct wrt to the underlying logic but also wrt to common sense.

Examples:

(implies (and house (some near lake)) mosquito-house)
(instance house1 house)
(instance lake1 lake)
(related house1 lake1 near)
=> (instance house1 mosquito-house)

(instance house1 house)
(instance cup1 cup)
(related house1 cup1 inside)
=> inconsistent???

What is a "chair"?



Uncertainty Problems (1)

Fuzzyness of concepts

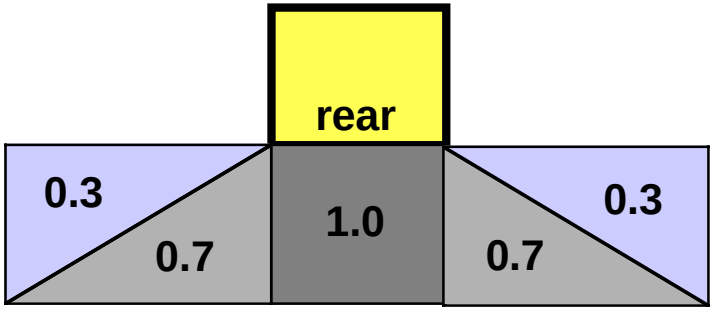
Many high-level concepts have unsharp boundaries.

"behind" "overtake" "meet"

=> mapping into logical propositions may be problematic

- **Fuzzy set theory offers "degree of applicability"**

- **Probability theory offers statistical measures for language use**

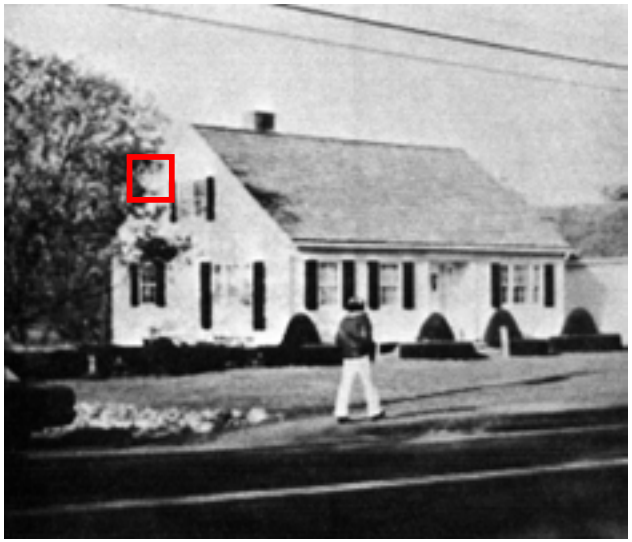


Fuzzy definition of behind

Uncertainty Problems (2)

Uncertainty of data

Example: Object boundaries



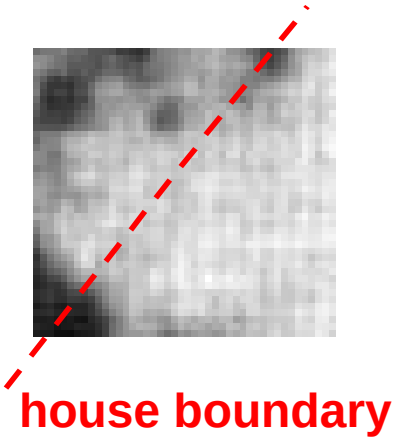


Image interpretation is fundamentally ill-defined

Uncertainty Problems (3)

Exploring multiple hypotheses

Answers from several disciplines:

- **graph matching**
- **heuristic search**
- **optimization theory**
- **logic theories**
- **probability & utility theory**
- **case-based reasoning**
- **neural networks**
- **particle physics**

(and others)

Mixed bottom-up and top-down interpretation strategies have been rarely explored

Uncertainty Problems (4)

Cultural clash betweeen logical and probabilistic reasoning

Probabilistic methods are nor yet seamlessly integrated with logical calculi

<u>Interesting recent developments</u>:

- **First-order probabilistic inference (Poole 03)**
- **Probabilistic relational models (http://dags.stanford.edu/PRMs/)**

Example for reasoning in image interpretation:
(from Kanade´s invited lecture at IJCAI-03:
"Computer Vision: AI or Non-AI Problem?")

car on left side of street
(uncertain orientation of car)

japanese signs => left-hand traffic

} **orientation of car resolved**

Topics of Tuesday

What is High-level Scene Interpretation?	3
Historical Examples	8
Relational Matching	27
Rule-Based Interpretation	42
Description Logics	47
Image Interpretation as Deduction	61
Logics of Image Interpretation	66
Image Interpretation as Configuration	72
Signal-symbol Interface	75
Occurrence Models	87
Navigating in Hallucination Space	96
Recognizing Intentions	101
Bayesian Nets	105
Summary	117