Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

MIN-Fakultät
Fachbereich Informatik
Arbeitsbereich SAV/BV (KOGS)

# IP2: Image Processing in Remote Sensing
# 11. Knowledge Based Interpretation and Exam Preparation

Summer Semester 2014

Benjamin Seppke

# Agenda

- Kinds of knowledge
  - Further data and scene knowledge
  - Common Sense and expert knowledge
- Knowledge representation
  - Implicit
  - Explicit
- Description Logics as a common modeling base
  - Exemplary models
  - Reasoning
  - Challenges
- Open challenges
- Exam preparation

# Knowledge? Knowledge!

- Knowledge may show up in different ways, compare with IP2 – Part 1!
  - Geometric knowledge
  - Topologic knowledge
  - Temporal knowledge
- Knowledge may be given by
  - (potentially) incomplete facts
  - Common sense knowledge
  - Domain expert knowledge
- Usually more factual data than expert knowledge!

# Sources of Remote Sensing Knowledge

- Image analysis/interpretation
- Cartographic data (e.g. provided by local agencies)
  - Street maps
  - Land-use maps
  - Oceanographic data: nautical, tidal, current charts etc.
- Derived data:
  - Sea surface (anomaly) height
  - Earth gravity field
  - Atmospheric data
- Expert interviews
- Physics (e.g. of fluid dynamics)
- Common sense (e.g. spatio-temporal) knowledge

# Challenges for the Use of Remote Sensing Knowledge

- Each data source may introduce another data format, standardized (but) only to its domain!
- Usually large amount of data – too large for most reasoning systems!
- Highly heterogeneous data:
  - Dense knowledge vs. sparse (point-like) knowledge
  - Raster vs. polygonal knowledge
  - General knowledge between facts
- Derived data may not be appropriate and erroneous!

→ Development of Black (processing) Box systems in former times!

# Implicit Knowledge Use

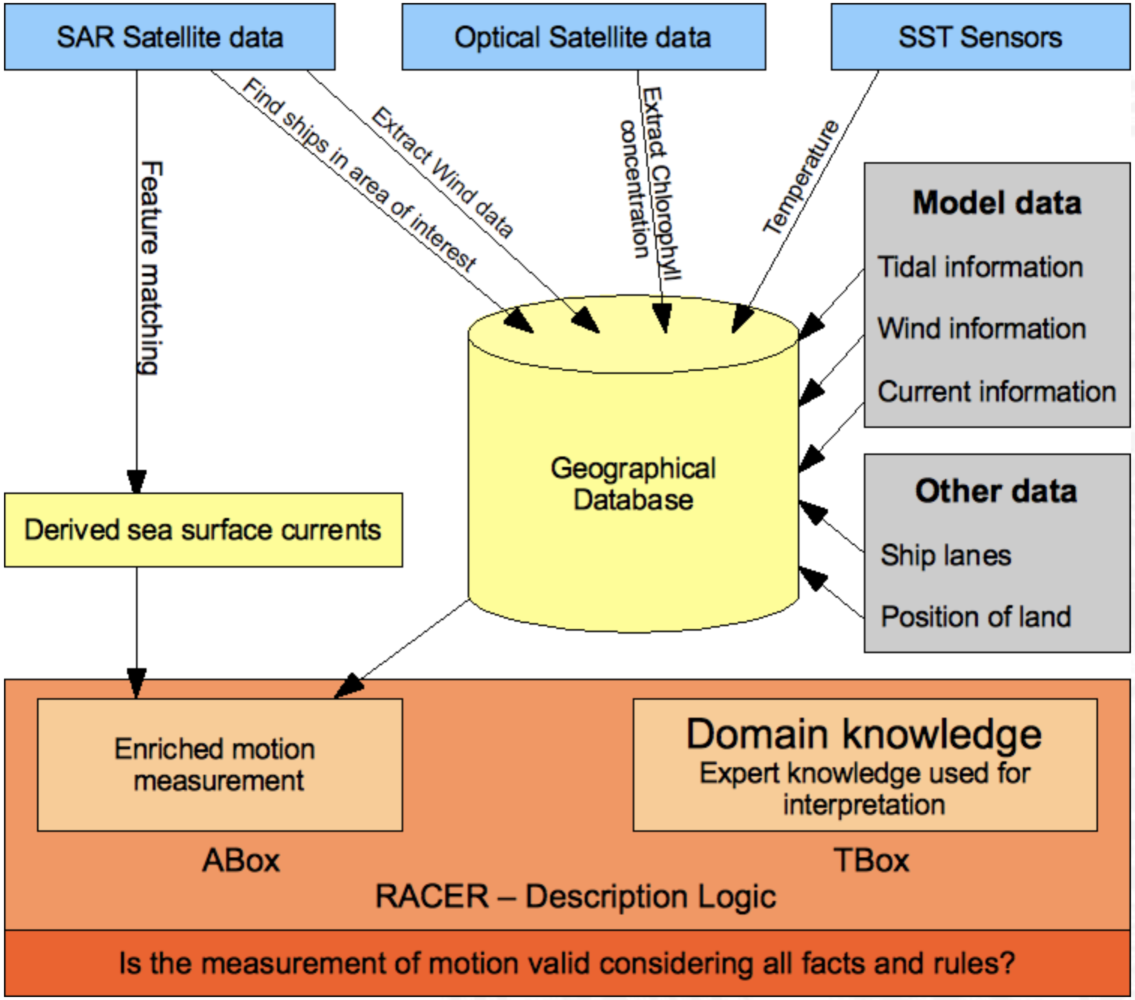
- Some expert knowledge is "hard wired" into an usually black box system
- Scene analysts use the system by
  1. Importing an image
  2. Set the appropriate thresholds / other values
  3. Let the systems produce (1-click) results
  4. Interpret the results
- Drawbacks
  – Implicitly modeled knowledge neither visible nor interchangeable
  – Thresholds may follow expert knowledge → not modeled!
  – Interpretation step mostly unassisted / purely manual
- Example for such a system: ENVI

# Explicit Knowledge Use

- Idea: **All sources of knowledge** shall be combined with respect to:
  - Expert knowledge
  - Common sense knowledge

  to derive automated scene interpretations!
- Need for:
  - Good knowledge modeling (large factual database!)
  - Appropriate knowledge representation and reasoning systems:
    - Expressiveness
    - Speed
    - Logical correctness
- Currently no "Remote Sensing"-tailored system, but eCognition® provides at least:
  - Image segmentation + decision tree semantics
  - Called: "Semantic classification"
- But: Many logic systems with reasoning service available!

# Example: Explicit Knowledge Modeling
## Revisiting the Derivation of Sea surface currents



# Logics for Knowledge Representation

- Propositional Logic *Aussagenlogik*
  - Low expressivity
  - Easy understandable (in form of truth tables)
  - Decidable
- Predicate Logic *Prädikatenlogik*
  - High expressivity
  - May be hard to understand (models need to be fulfilled)
  - In general: not decidable!
- Description Logic *Beschreibungslogik*
  - Decidable subsets of Predicate Logics
  - Highly optimized reasoners
  - Modeling tools available (e.g. Protégé)

# Description Logics I

- Family of knowledge representation languages:
  - originated from early attempts in the 1970s to model knowledge with concept-or class-based structures.
  - Nowadays: the semantic basis for the Semantic Web (e.g.,OWL DL is basically a description logic).
- Most contemporary DLs can be considered as subsets of first-order logic
  - → The inference services are well defined here!
  - → (comparably) fast reasoning algorithms exist.

# Description Logics II

- Knowledge in DL systems comes in two disguises:
  - class- or concept-based knowledge, → Stored in T-Box
  - individual-specific knowledge → Stored in A-Box
- Binary relations for the Individuals (called roles).
  If R is such a role, and $C$ and $D$ are concept descriptions:

*concept ::=* *atomic-concept | top | bottom*

*concept ::=* *(and C D) | (or C D) |*
*(some R C) | (all R D) |*
*(not C)*

# Description Logics III

- The denoted set of individuals in the domain of discourse is specified inductively by means of a so-called Tarski-style interpretation function $I$:

$$I((and\ C\ D)) = I(C) \cap I(D)$$
$$I((or\ C\ D)) = I(C) \cup I(D)$$
$$I((some\ R\ C)) = \{\, i \mid \exists\, j \in \Delta : j \in I(C), (i, j) \in I(R) \,\}$$
$$I((all\ R\ C)) = \{\, i \mid \forall\, j \in \Delta : (i, j) \in I(R) \Rightarrow j \in I(C) \,\}$$
$$I((not\ C)) = \Delta \setminus I(C)$$

- $I(C)$ is also called the extension of $C$ (w.r.t. an interpretation).

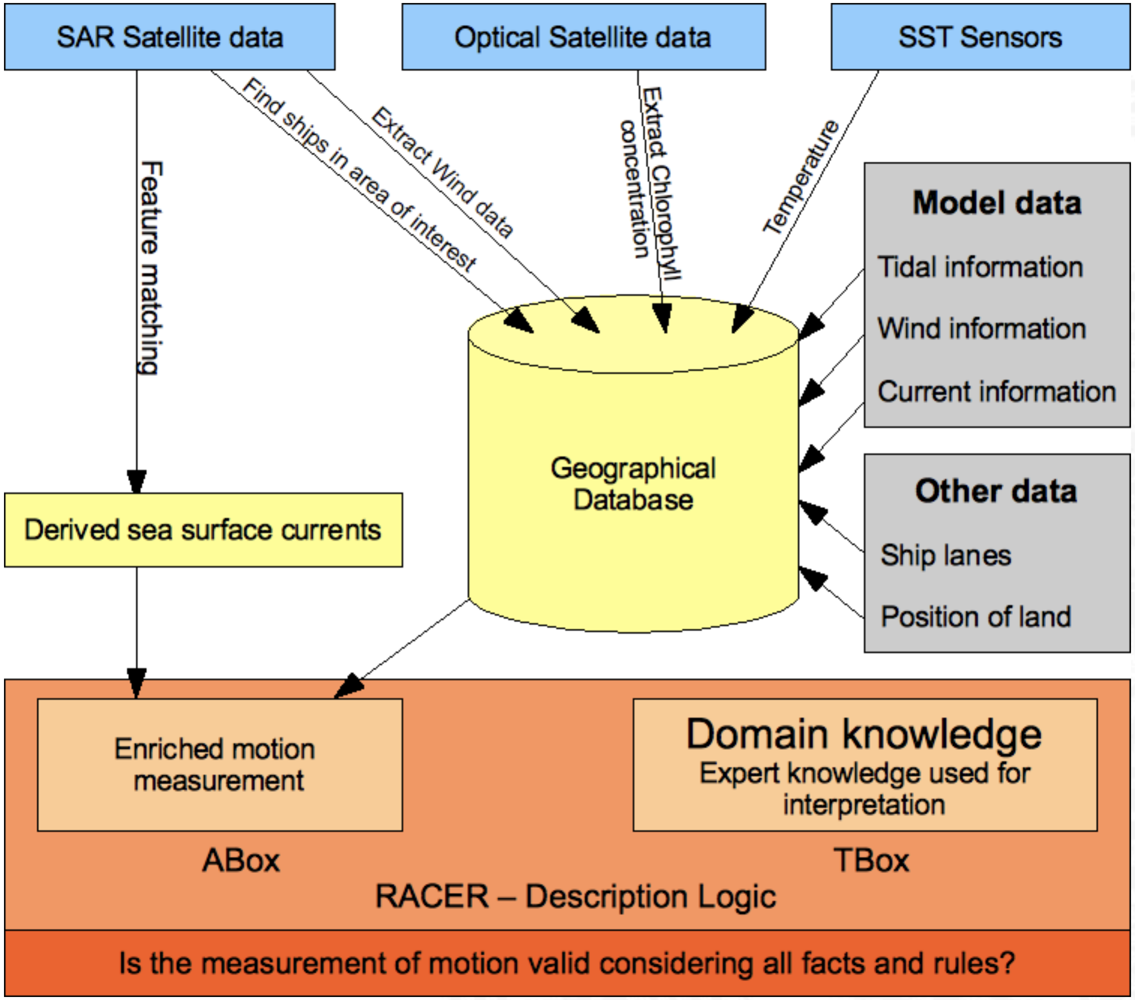
# Description Logics IV

- A concept $C$ is said to be satisfiable or consistent iff there is at least one interpretation function and non-empty domain $\Delta$ such that I maps $C$ to a non-empty subset of $\Delta$ ;
- An interpretation, which satisfies $C$ is also called a model of $C$.
- An important relationship between concepts is the subsumption relationship:
  - It is said that $C$ is subsumed by $D$ if the extension of $C$ is a subset of the extension of $D$ in all models of $C$ and D.
  - Then, $C$ is called the more specific, subsumee, and $D$ the more general concept or subsumer

# The Description Logic System Racer Pro

- Implements the expressive description logic *SHIQ(Dn)*:
  - Transitive, functional and inverse roles,
  - Role specialization hierarchies,
  - Reasoning with data types (e.g., strings, reals, integers, Booleans), and
  - Some additional concept constructors (e.g. qualified number restrictions of OWL2)
- Offers many advanced proprietary features, such as:
  - (Grounded) First-order queries, rules,
  - Programmatic "server-sided" scripting,
  - extensibility, and
  - some innovative inference services (such as abductive query answering).
- More than 10 years of continuous improvements → one of the fastest A-Box reasoning system
- An ideal basis for knowledge-intensive applications which require A-Box reasoning and A-Box query answering
- Has already proven to fit well for reasoning by means of computer vision scene interpretation [4].

# Case Study: Description Logic Modeling for the validation of derived sea surface currents



# Qualitative vs. Quantitative Modeling

- The high expressivity of Racer allows quantitative modeling of a motion vector FV due to the closed domain of numbers by:

| | |
|---|---|
| *(constrained FV V alx FVx)* | *(constraints (= V alx x))* |
| *(constrained FV V aly FVy)* | *(constraints (= V aly y))* |
| *(constrained FV V alu FVu)* | *(constraints (= V alu u))* |
| *(constrained FV V alv FVv)* | *(constraints (= V alv v))* |

- Classical approach: Define qualitative measures for knowledge:

*(instance FV (and measuredcurrent*
*(some has-direction west)*
*(some has-velocity moderate)))*

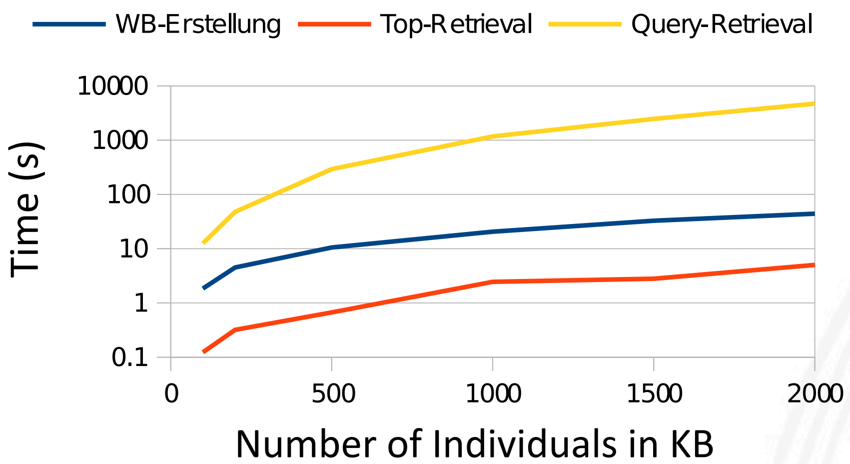
With:

- Direction concepts: *northwest, north, northeast, east, southeast, south, southwest, west*
- Velocity concepts: *slow, moderate, high*
- Localization roles: *touches, is-next-to, is-far-away-from*

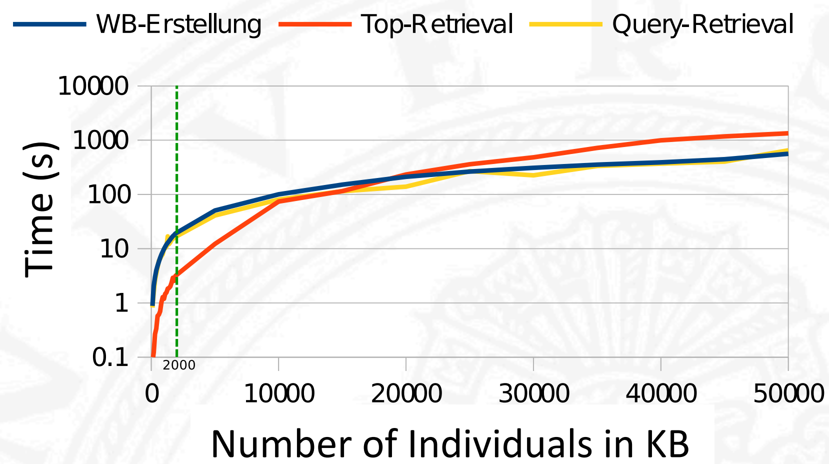
- Which kind of representation shall be used?

# Time Comparison of Reasoning Tasks

## Quantitative Representation




## Qualitative Representation




- Favoring the qualitative approach for large number of individuals
- Need to define the ontology appropriate to the task!
- Automatic creation must know the semantics of the qualities to map from real values to qualitative values

# Manual Modeling of the T-Box Smoothness Constraints

- Propagating localization roles to ensure smoothness:

*(implies (some has-velocity slow) (all touches (all has-velocity slow-tp)))*
*(implies (some has-velocity slow) (all is-next-to (all has-velocity slow-np)))*
*(implies (some has-velocity slow) (all is-far-away-from (all has-velocity slow-fp)))*

- Use propagates rules to define erroneous concepts
  - For velocities and the *touching* localization rule:

*(equivalent touches-with-wrong-velocity (or (some has-velocity (and slow high-tp))*
*(some has-velocity (and slow moderate-tp))*
*(some has-velocity (and moderate slow-tp))*
*(some has-velocity (and moderate high-tp))*
*(some has-velocity (and high slow-tp))*
*(some has-velocity (and high moderate-tp))))*

  - Combine for all localization rules:

*(equivalent spatially-related-with-wrong-velocity*
*(or touches-with-wrong-velocity*
*is-next-to-wrong-velocity*
*is-far-away-from-wrong-velocity))*

# Manual Modeling of the T-Box
# Look-Alike Detection

- Ships on ship routes may yield to look-alikes

*(equivalent shiproute-problem (and measuredcurrent*
*(or (some touches shiproute)*
*(some is-next-to shiproute))))*

- Motion estimates may be unreliable close to coastal areas:

*(equivalent coastal-problem (and measuredcurrent*
*(or (some touches land)*
*(some is-next-to land))))*

- Combine both as look-alike problems

*(equivalent lookalike-problem (or shiproute-problem coastal-problem))*

# Manual Modeling of the T-Box
# Other error concepts

- For the research on sea surface current estimation, the following errors have been modeled by means of the T-Box:
  - Intra-smoothness errors
  - Look-Alike errors
  - Deviation from (numeric) model results of the sea surface currents
  - Deviation from wind measurements
- Each error concept is defined by its unique set of role / concept relationships
- The most general problem concept is then defined by:

*(equivalent problem (or lookalike-problem*
*modelledcurrent-problem*
*currentsmoothness-problem)*
*wind-problem)*

- More precise description may be derived using subsumption!

# Automatic A-Box Creation

- Number of vectors is reduced by clustering algorithm
- For each vector:
  - Create an instance of the vector in the A-Box
  - Map from real to symbolic values and combine via roles with the created vector instance:
    - Vector velocity
    - Vector direction
    - Cluster standard deviation
      → vector smoothness constraint
  - Query the geographical database for additional facts located within max (is-far-away) range.
  - If found create concepts and connect to the vector via appropriate roles

# Example: A-Box for Sea Surface Currents

- Consider the following A-Box:

*(instance i (and measuredcurrent*
*(some has-direction southwest)*
*(some has-velocity moderate)))*

*(instance j (and measuredcurrent*
*(some has-direction south)*
*(some has-velocity high)))*

*(instance l land)*

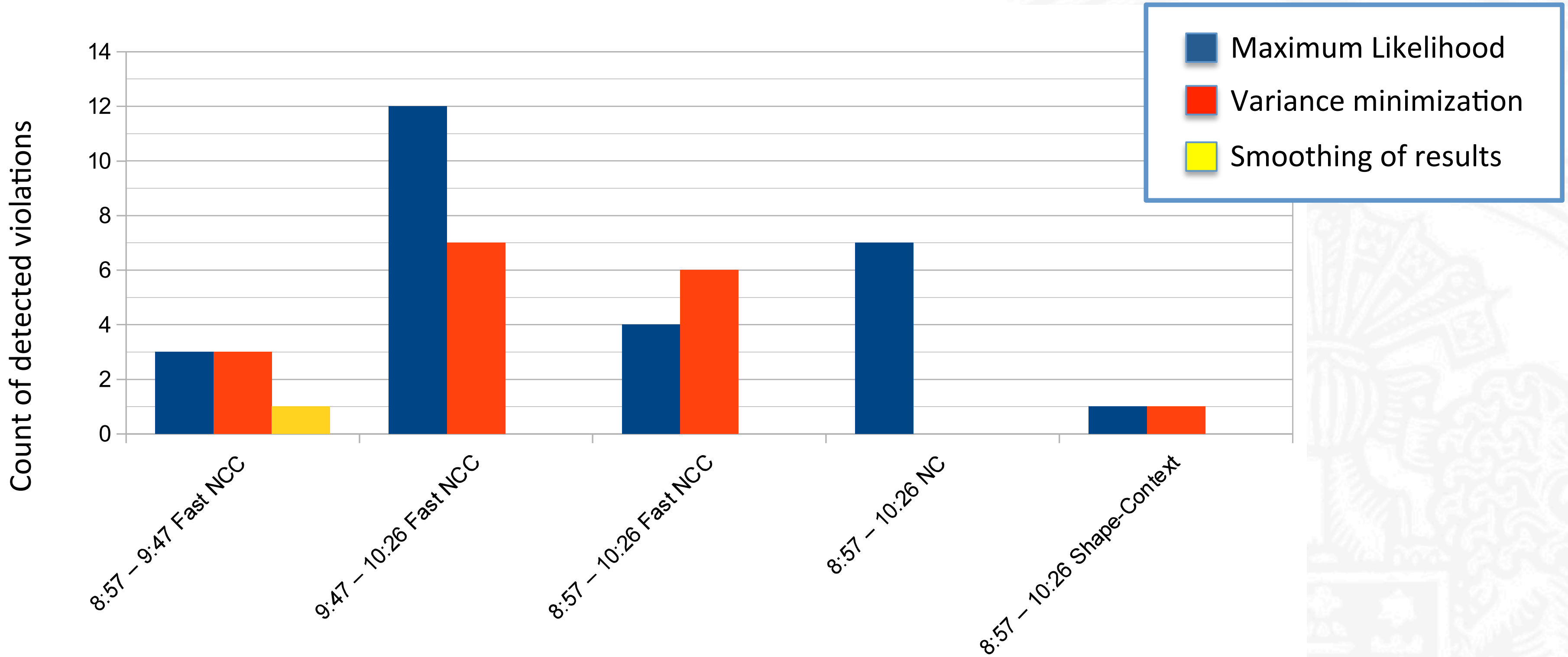
*(instance m (and modelledcurrent*
*(some has-direction west)*
*(some has-velocity high)))*

*(instance s shiproute)*

*(instance w (and windcurrent*
*(some has-direction southeast)*
*(some has-velocity high)))*

*(related i j touches)*
*(related i m touches)*
*(related i s is-next-to)*
*(related i w is-far-away-from)*
*(related j l is-next-to)*

# Detected Smoothness Violations of Feature based Results w.r.t the KB (Baltic Sea)



# Example: Questioning Racer Pro

- Using the new Racer Pro Query Language (nRQL):

```
? (retrieve (?x) (?x problem))
> (((?x j)) ((?x i)))

? (retrieve (?x) (?x lookalike-problem))
> (((?x j)) ((?x i)))
? (retrieve (?x) (?x shiproute-problem))
> (((?x i)))
? (retrieve (?x) (?x coastal-problem))
> (((?x j)))

? (retrieve (?x) (?x modelledcurrent-velocity-problem))
> (((?x i)))
? (retrieve (?x) (?x currentsmoothness-velocity-problem))
> (((?x j)) ((?x i)))
? (retrieve (?x) (?x wind-problem))
> (((?x i)))

? (individual-direct-types i)
> ((currentsmoothness-velocity-problem) (modelledcurrent-velocity-problem)
   (wind-problem) (shiproute-problem))
```

# Knowledge Based Analysis: Open Research Challenges

- Knowledge modeling (in Remote Sensing) is still under research!
- Main questions:
  - How to get from data- to knowledge base?
  - How to perform reasoning in small/appropriate time?
  - How to model geographical dependencies efficiently for large (e.g. raster) data
  - What about missing information?
    - Concepts
    - Dependencies between concepts
    - Missing roles?
  - Sensor-Fusion or knowledge fusion
  - How to deal with different granularities?
- And many many more...

# Knowledge Based Analysis: Starting Points for Further Research

- J. Roddick, M. Egenhofer, E. Hoel, D. Papadias, and B. Salzberg:

  Spatial, Temporal and Spatio-Temporal Databases—Hot Issues and Directions for PhD Research

  SIGMOD Record
- Mehul Bhatt:

  Reasoning about Space, Actions and Change – A Paradigm for Applications of Spatial Reasoning
- R. Möller, V. Haarslev and C. Lutz:

  Spatiotemporal Reasoning Based on Inferences: The ALCRP(D)-Approach

  *… and much more work from Ralf Möller!*

# Exam Preparation

- The oral exam (up to 30 minutes, overall time including IP2- High level vision) may cover all the fields of the lecture series:
  - Gravitational Astronomy
  - Orbits, Acquisition Constraints and Missions
  - Fundamentals of EM-Radiation
  - Remote Sensing Sensors
  - Image Processing:
    - Image Characteristics and Preprocessing
    - Classification and Segmentation
    - Edge Detection and Motion Estimation
  - Knowledge based Image Interpretation
- The next slides collect some possible questions…

# Introduction and Gravitational Astronomy

- What were the main discoveries, which formed the basics for Remote Sensing?
- Give a definition of Remote Sensing!
- Give a typical workflow for Remote Sensing processing!
- Describe Kepler's laws and their importance to space-based earth observation!
- Why are reference planes and epochs of importance?
- Explain the effect of perturbation and how it can help with orbit design!

# Orbits and Acquisition Constraints

- Give at least four different types of orbits and explain them!
- Where on a Molniya orbit is the observing time high, where low?
- Which orbits do weather satellite typically have and why?
- Which orbits do EO satellites usually have?
- Give advantages of sun-synchronous orbits?
- What has to be considered if one wants to take space-borne stereo images?
- How much of a rocket's mass has to be fuel?

# EM-Radiation: Waves and Basic Principles

- How can EM waves be created?
- Explain the correspondence between magnetic and electric field w.r.t. EM waves!
- How fast are EM waves at vacuum?
- Explain the influence of the dielectric constant e.g. for waves which hit water!
- What does “polarization” mean?
- What are wave packages and coherence?
- Explain the Doppler effect!

# EM-Radiation: Interaction with Matter

- Distinguish the radiometric from the spectrometric and photometric system!
- What are black bodies and what may be described by their radiation?
- Explain emissivity of a material at the example of snow!
- Explain why EM radiation has to be considered as a particle flow sometimes!
- Explain the photo-electric effect!
- What are the requirements for reflection?
- Explain the absorption index of a material at the example of sliver or gold!

# The Atmosphere of the Earth

- Explain the dependency between pressure and temperature!
- What is the Barometric Scale Factor?
- Describe the temperature distribution inside the atmosphere!
- How do we chemically characterize the atmosphere?
- What is the magnetosphere, and why do we need it?
- Explain the kinds of scattering of EM waves by the atmosphere!
- Explain, why the sky is blue!
- Describe the trajectories of EM waves from sun through the atmosphere!

# Optical Sensors

- What constraints the EO by means of optical sensors?
- Briefly explain the differences between photographic sensors, framing and scanning optical sensors!
- Give two advantages of Along-Track scanners over Rotating scanners
- What are NADIR- and Off-NADIR modes?
- What is the azimuth axis?
- What is the revisit time?
- Can a single high-resolution optical satellite be used to monitor a single people over hours, like it is shown in Hollywood movies?

# Microwave and SAR-Sensors

- Distinguish between active and passive MW Systems!
- What are the main advantages when using microwaves?
- Briefly explain satellite based altimetry!
- Why is altimetry of such importance, e.g. in climate research?
- What is mainly measured by satellite scatterometers?
- Explain the difference between real and synthetic Aperture!
- Which are the important steps and advantages of SAR systems?
- What is Speckle Noise?
- What is Interferometric SAR?

# Image Characteristics and Preprocessing

- Name at least two sources of geometric distortions!
- What does "panorama distortion" mean?
- Explain the effect of relief displacement!
  - Why may it be useful sometimes?
  - What are the corresponding effects for imaging radar?
- What are the important steps in sensor normalization?
- Why do we need good atmospheric correction?
- What does registration mean w.r.t. Remote Sensing?
  - Which different registration tasks may occur?
  - What has to be considered in mountainous areas?
  - What registration methods do you know for point to point correspondences?

# Image Classification

- Describe the Classification task!
- What distinguishes photo interpretation from quantitative analysis?
- Explain the supervised classification approach in general!
- What does “Maximum Likelihood Classification” mean?
- What are discriminant functions and Thresholds?
- When may the minimum distance classification be applied?
- What is “Context Classification? Give an Example!
- What is Clustering? Where is it needed?

# Edge Detection

- Describe the correspondence of the image gradient and edges in images!
- What is the main difficulty for the definition of multispectral edges?
- Explain the differences of the mean, max and multi-spectral gradient for edge detection!
- What is the main advantage of the multi-spectral gradient measure?

# Motion Derivation

- Name and explain the two algorithmic classes for motion detection!
- What are the advantages of feature-based methods?
- When are differential methods applicable?
- Explain the problem of large spatiotemporal distances!
- Explain the concept of focused search!

# Knowledge Based Interpretation

- Give examples for different kinds of knowledge in the context of Remote Sensing!
- Explain the difference between implicit and explicit knowledge use!
- How do Description Logics store knowledge?
- What kind of knowledge box may be filled automatically (e.g. by means of derived results)
- What are the main challenges in knowledge representation and reasoning w.r.t. Remote Sensing applications?