Formalization and Automatic Interpretation of Constraints

Xiang Zhang & Jantien Stoter & Tinghua Ai

11th ICA Workshop on Generalisation and Multiple Representation

20th June 2008 – Montpellier, France

ITC

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION

Overview

- Motivation
- A framework of formalization and interpretation
- A special interpretation issue
- Conclusions & remarks

Introduction

- The work is initiated by the PhD research topic – Automated Evaluation of Generalized Topographic Maps.
- Generalization output specifications are defined by a set of cartographic constraints (constraint-based approach).
- The aim of the work is to improve the degree that constraints can be expressed (interpreted) by machines to facilitate this automatic process.

ITC

Problems

- Informal knowledge in constraints
  - High level concepts contained
    - Understandable by human experts
    - Not interpretable by machines
  - Some constraints lack knowledge adaptive to specific situations
    - i.e. too general to be applicable and effective
- Should be formalized (knowledge formalization)

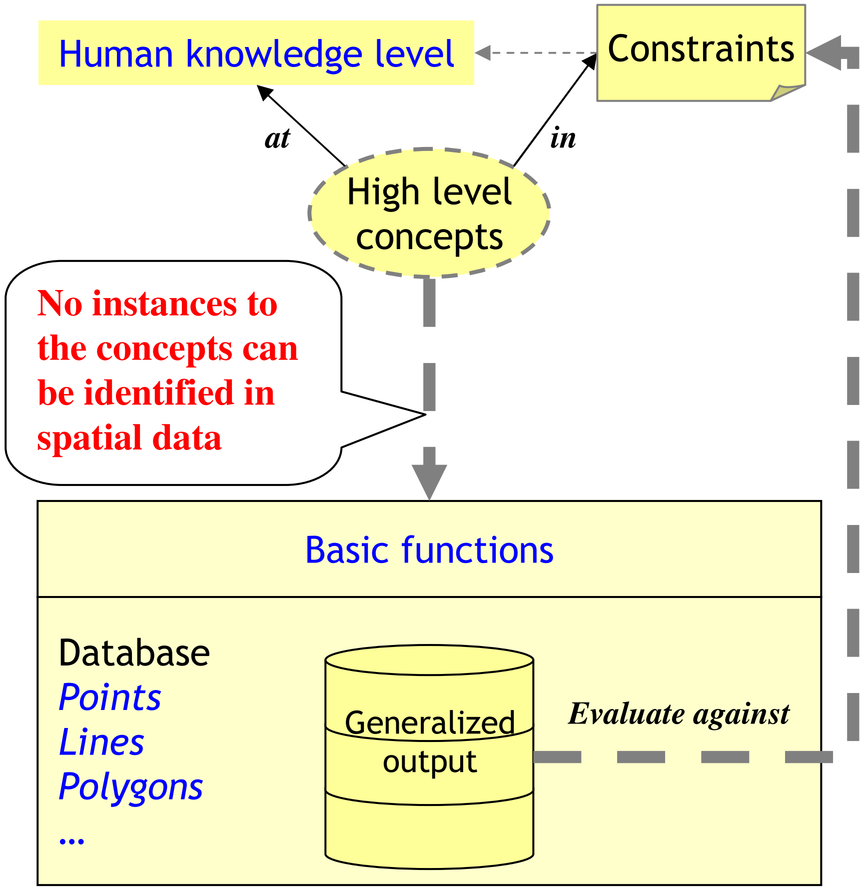
Problems (continued)

- Raw data representation
  - Basic geometry and semantic data type
    - e.g. OGC Simple feature (point, line, polygon, …)
    - Attributes are not designed for generalisation process
  - High level concepts are not explicitly represented
    - Contextual relations between proximate objects
    - Pattern (e.g. alignment of buildings)
    - Some geo-phenomena having no crisp boundary (e.g. urban area) are not modeled as ***objects*** in most databases
- Should be enhanced (data enrichment)

ITC

Problems (continued)

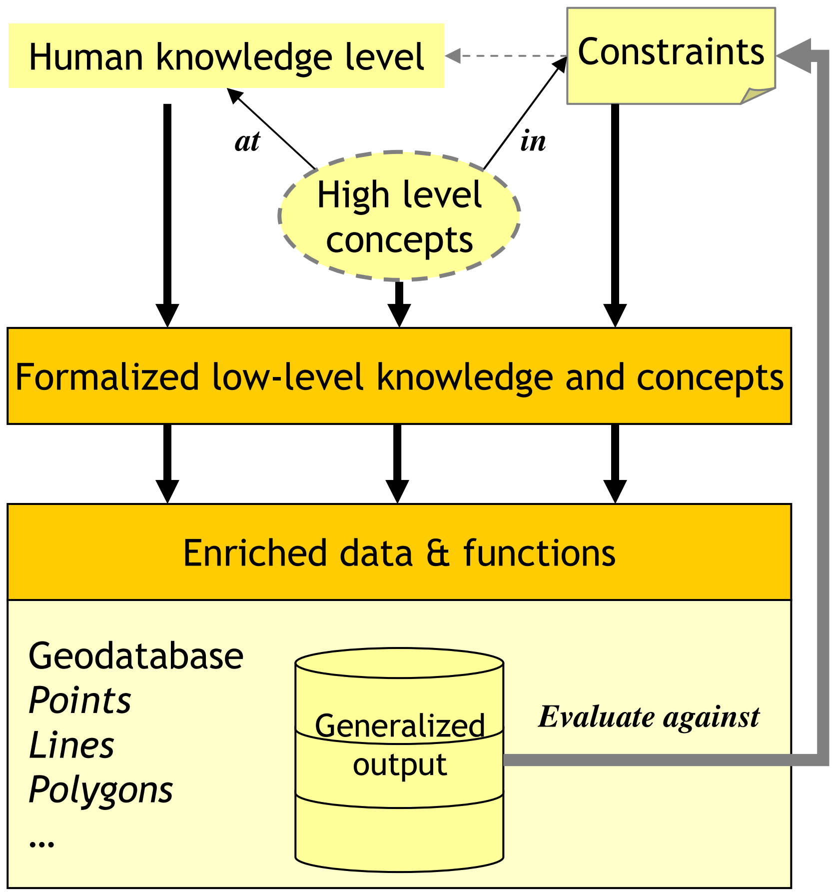
- Basic functions
  - Spatial & attribute query
  - Geometric, topological operations
  - Distance, area calculation
  - Buffering operation
  - ……



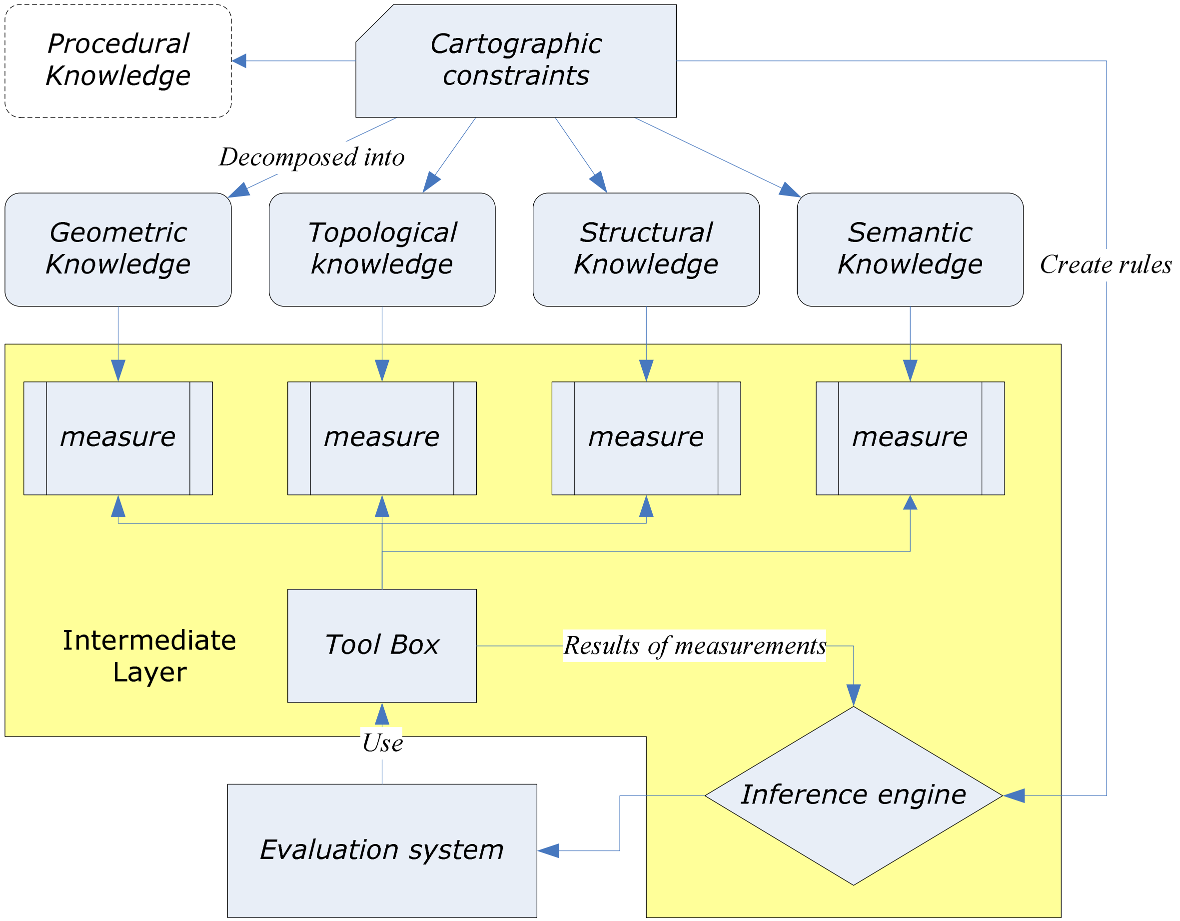
Our approach

- Decompose knowledge in constraints into low level knowledge
  - Formalize knowledge and high level concepts in constraints as much as possible

- Enriched data & functions
  - Contain basic functions
  - Identify instances for certain concepts by on demand calculation
  - ……



Proposed decomposition framework



Knowledge-based decomposition

- Assumptions:
  - Nearly infinite constraints for different generalization purposes and tasks
  - Countable kinds of low level knowledge
  - The combination of these low level knowledge can express various constraints
- Principles:
  - The decomposed low level knowledge should be atomic to be measureable.
  - Or the knowledge can be deduced from the data model (e.g. semantic info modeled as attributes).
  - High level concepts should be decoupled into low level concepts/knowledge and their relations.

Formalization of constraints: two levels

- Formalize the constraints using predicate logics (*being the rules of inference engine*)
  - **Size** (Building) > X map mm2
  - **Access** (Building, Road) = **True**
  - **Density** (target building group) = **Density** (initial building group) * X %
  - **Exist** (Building | size (Building) ∈ [a, b] ***and*** **Context** (Building) = 'rural district') = **True**
- Formalize the high level concepts using terminological reasoning (*i.e. identify instances for a certain concept*)
  - e.g. **Context** (Building) = 'rural district'
  - Decompose the high level concepts into low level concepts and their relationships (using e.g. semantic model)

A special interpretation issue

- Two constraints in coastal cartography are expressed as follows:
  - *“roads leading to a building at the end of peninsulas must not be omitted”*
  - *“buildings at the end of peninsulas should be preserved”*
- The two informal constraints can be formalized in the form as follows:
  - **Access** (Building, Road) = **true**
  - **Exist** (Building | **Context** (Building) = *‘peninsula’*) = **True**

- However, the high level concept *‘peninsula’* is not well defined.

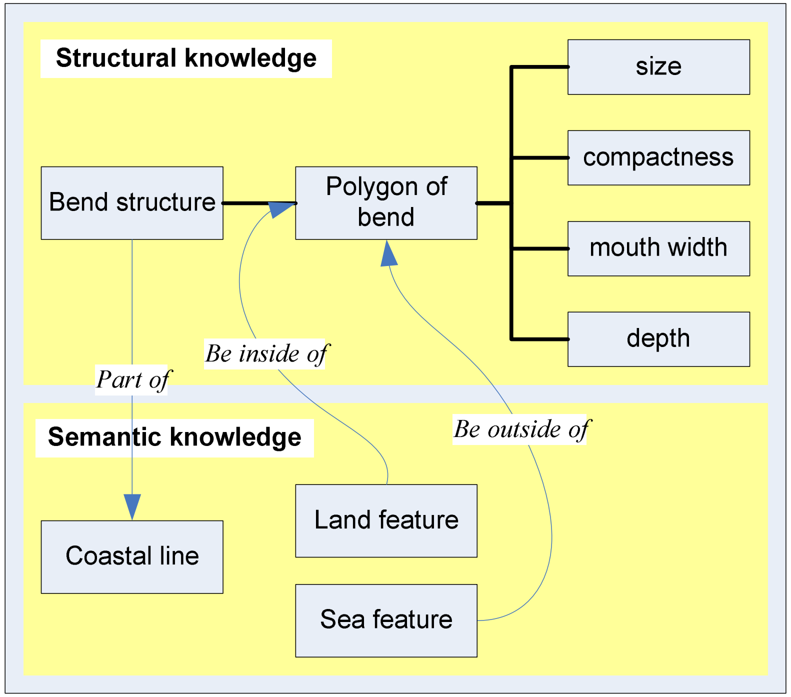
‘Peninsula’

- The geographical phenomenon *‘peninsula’* is defined in natural language as “a piece of land that is bordered on three sides by water. It can also be a headland, cape, island promontory, bill, point, or spit”. – Wikipedia

- The expression contains various synonymous and ambiguous terms, which are not machine-readable.

ITC

Semantic model of ‘Peninsula’ – concept formalization



- The model formally defines *‘peninsula’* using different types of low level knowledge and their relationships.
- *‘Peninsula’* is a *bend structure* that is a ***part of*** a *coastal line*;
- It is represented by the *polygon* of a *bend structure*, with *land feature* ***inside*** and *sea feature* ***outside***;
- The bend polygon of *‘peninsula’* is specified by some characteristics according to applications (only few of them

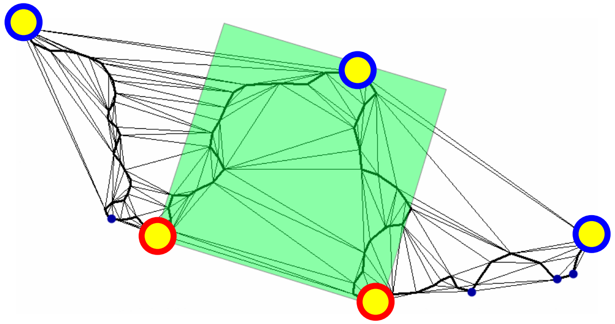
IF

meet (*bend polygon*, *sea feature*) ***and*** **coverby** (*bend polygon*, *land feature*) = **true**

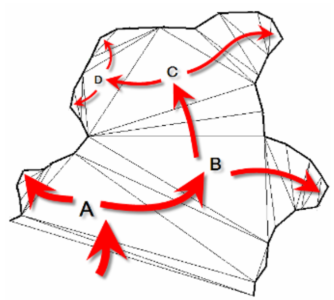
THEN

bend polygon = *‘Peninsula’*

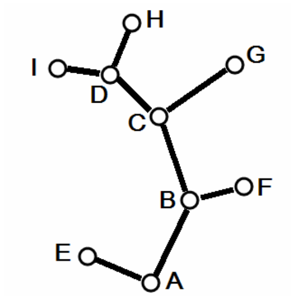
Bend structure detection – data enrichment



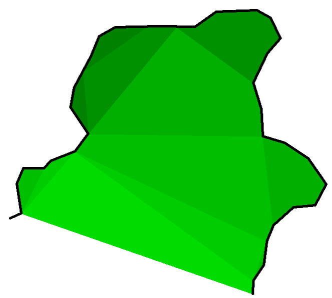
Triangulation &
Characteristic points detection



Segmentation &
Organizing structure



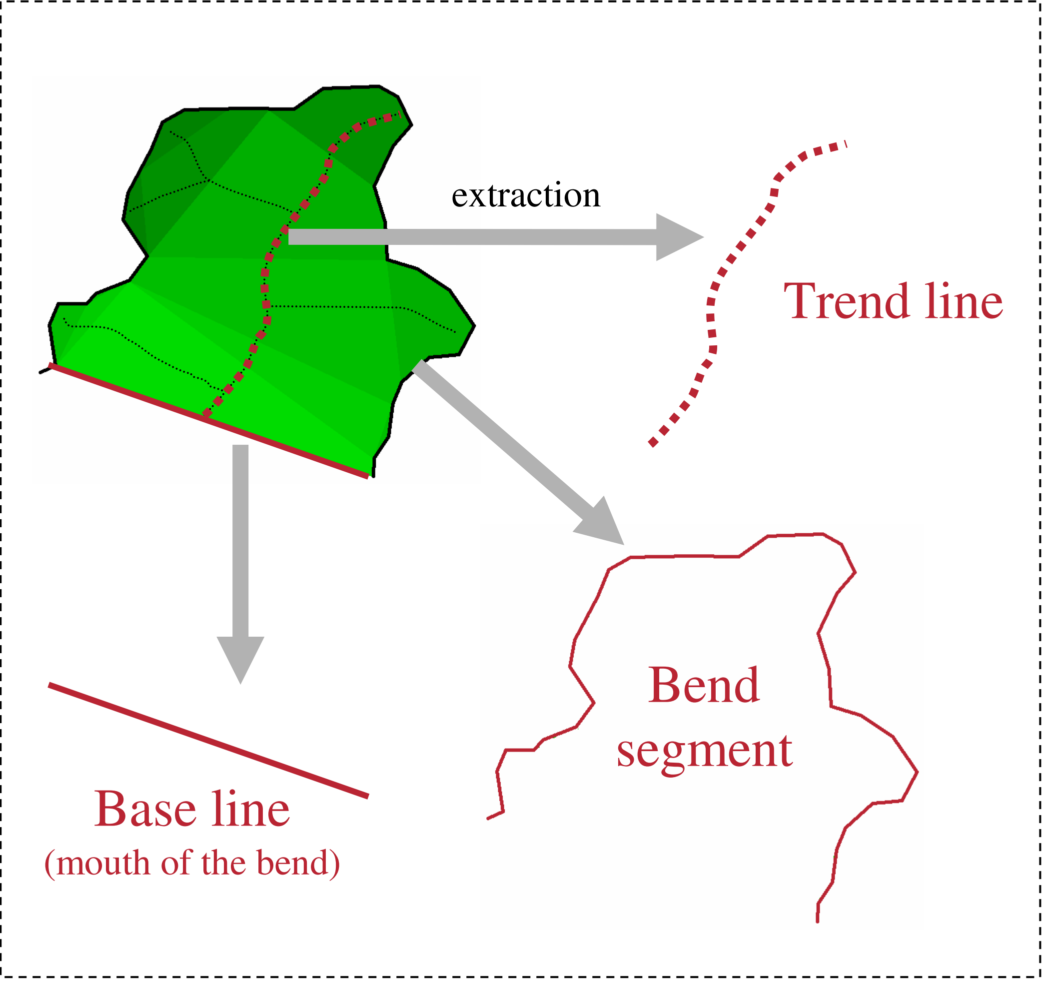
Binary tree of the hierarchical structure



Hierarchical bend structure &
Bend polygon (in green)

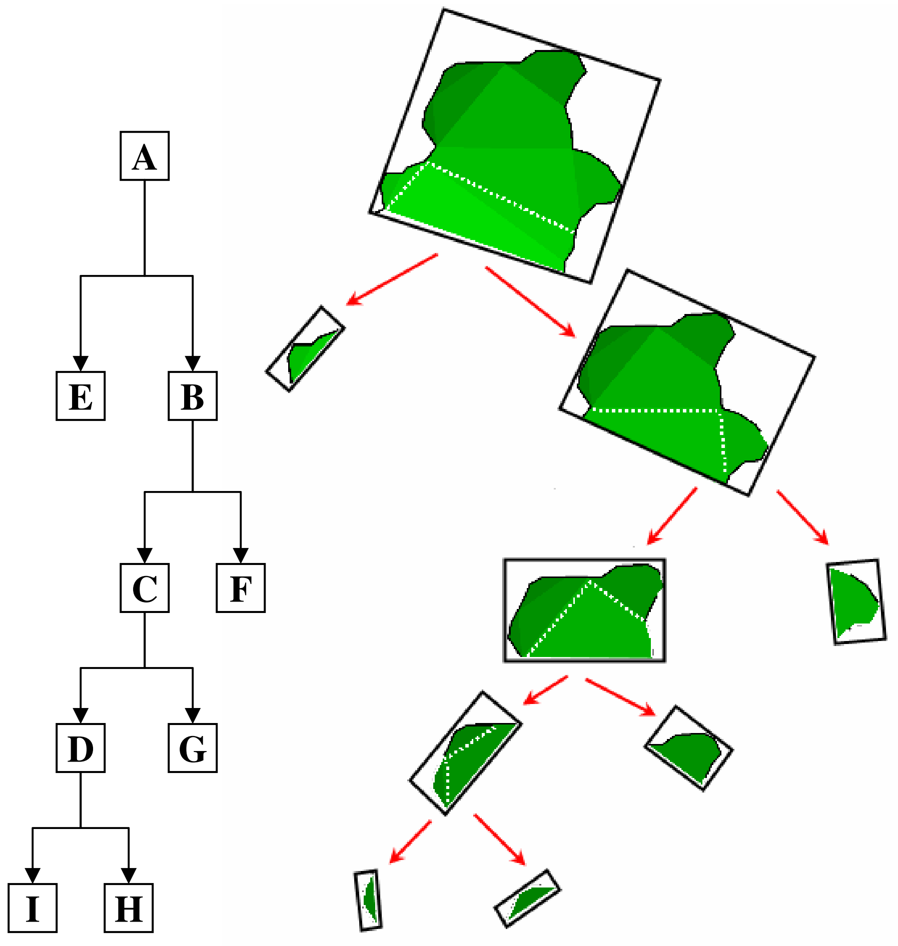
Bend structure descriptions – data enrichment

- Skeletonization by tracing from entrance triangle to end triangles
- The number of end triangles determine the number of skeletons
- Extraction of trend line based on the skeletons
- Definition of base line and bend segment



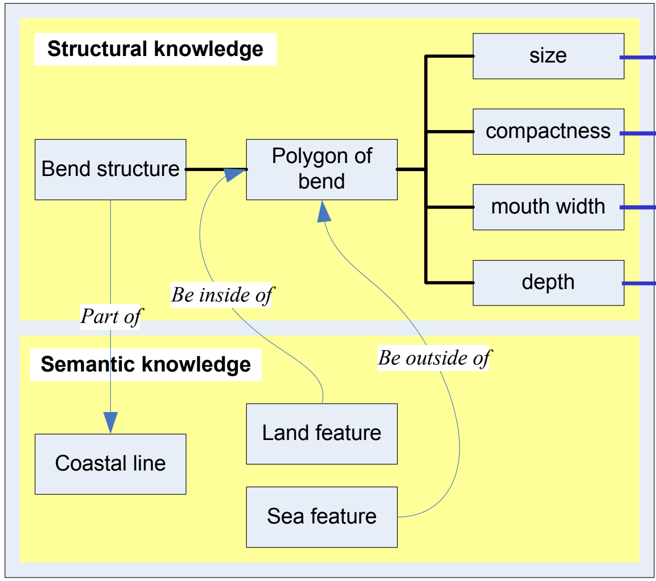
Level of Details of bend structure – data enrichment

- Visiting all the nodes of the binary tree will access all bends embedded
- These micro bend structures are different in size and other descriptions
- Descriptions like base line, bend segment, and trend line can be defined at different *LoDs*



Computational model for ‘Peninsula’

Semantic Model of ‘Peninsula’



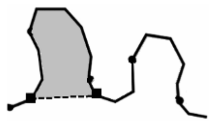
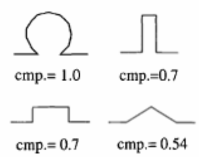
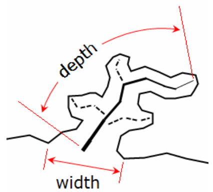
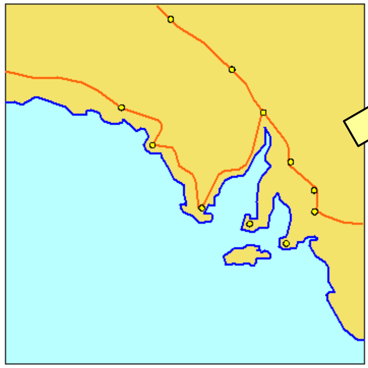
Characteristics	Measures	Example
Size	size =area (polygon); polygon is enclosed by *bend segment* and *base line*	
Compactness Index	“ the ratio of the area of the polygon over the circle whose circumference length is the same as the length of the circumference of the polygon” (*Wang and Müller, 1998*)	
Mouth width	width = length (*base line*)	
Depth	depth = length (*trend line*) (*Ai, 2007*)	

Table 1: Measures for characterizing bend structure (enriched functions)

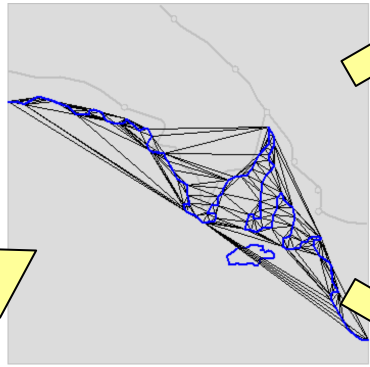
Test case

meet (*<u>bend polygon, sea feature</u>*) ***and*** **coverby** (*<u>bend polygon, land feature</u>*) = **true**

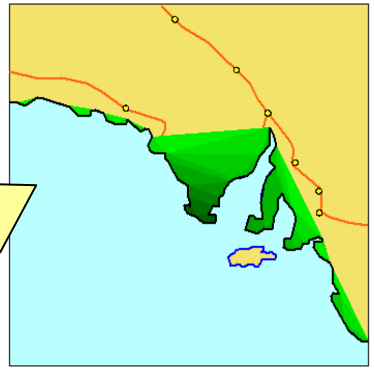
feature classes in data model



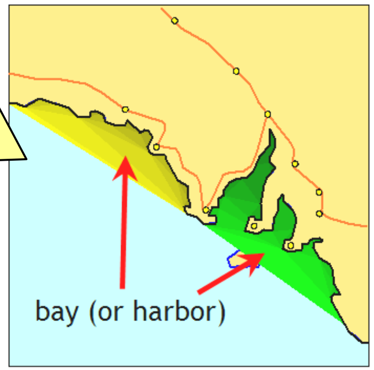
coastal area



Bend detection



'peninsula' identified



'bay' identified

coverby (*<u>bend polygon, sea feature</u>*) ***and*** **meet** (*<u>bend polygon, land feature</u>*) = **tru**

Conclusions & remarks

- More constraints can be expressed (interpreted) by machines if knowledge involved can be formalized.
- Decomposition of constraints provides a flexible way for machines to 'understand' the constraints, which are then possible to be evaluated automatically.
- In the peninsula interpretation issue, the structural knowledge is derived independent of semantic knowledge. This provides the flexibility for the reuse of the functionalities.

Conclusions & remarks

- The semantic model of ‘peninsula’ plays the role of formalizing and representing the high level concepts, and the bend structure detection, characterization and measure play the role of enrichment of data and functions.

- The semantic model is at conceptual level, certain formalization concepts should be employed as further implementation. The semantic model has to be linked with the underlying spatial data model. Once the data model changes, the semantic model is invalid. The predicate logics should also be implemented with suitable formal concepts.

ITC

Thanks!
Any questions?

Thanks to the Dutch program RGI for supporting my travelling here!

xzhang@itc.nl