Some Practical Issues in Building a Hybrid Deductive Geographic Information System with a DL-Component

KRDB 2003

Michael Wessel

University of Hamburg

“Cognitive Systems Groups”

Department of Computer Science

Project “DLS” (DFG Grant NE 279/8-1)

Contents

- Report about practical work in progress carried out in the context of the “DLS” project

Contents

- Report about <u>practical work in progress</u> carried out in the context of the "DLS" project
  - Vision of a hybrid deductive GIS

Contents

- Report about practical work in progress carried out in the context of the "DLS" project
    - Vision of a hybrid deductive GIS
    - Introduction of a first prototype

Contents

- Report about <u>practical work in progress</u> carried out in the context of the “DLS” project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar “hybrid” KR&R tasks

Contents

- Report about practical work in progress carried out in the context of the "DLS" project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar "hybrid" KR&R tasks
- The "vision" of a deductive GIS

Contents

- Report about practical work in progress carried out in the context of the “DLS” project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar “hybrid” KR&R tasks
- The “vision” of a deductive GIS
  - The domain & domain assumptions

Contents

- Report about practical work in progress carried out in the context of the “DLS” project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar “hybrid” KR&R tasks
- The “vision” of a deductive GIS
  - The domain & domain assumptions
  - Desirable reasoning tasks

Contents

- Report about <u>practical work in progress</u> carried out in the context of the "DLS" project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar "hybrid" KR&R tasks
- The "vision" of a deductive GIS
  - The domain & domain assumptions
  - Desirable reasoning tasks
- Implementation issues

Contents

- Report about practical work in progress carried out in the context of the “DLS” project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar “hybrid” KR&R tasks
- The “vision” of a deductive GIS
  - The domain & domain assumptions
  - Desirable reasoning tasks
- Implementation issues
  - Representing the data

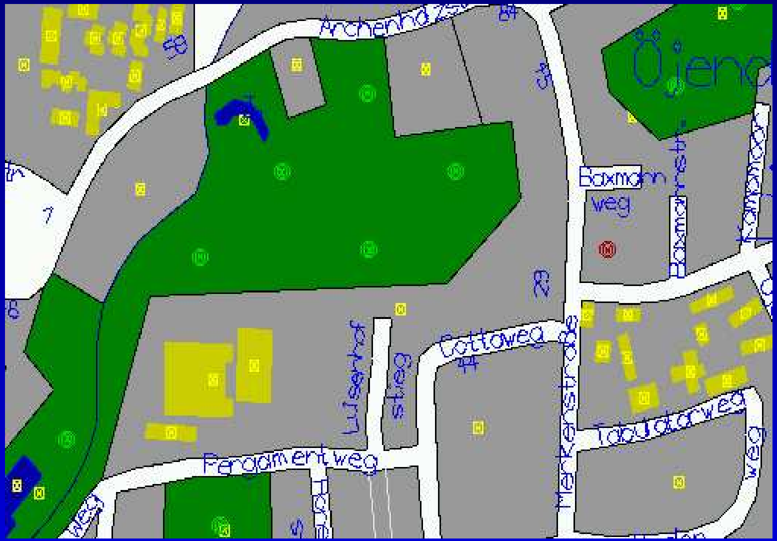
Contents

- Report about <u>practical work in progress</u> carried out in the context of the “DLS” project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar “hybrid” KR&R tasks
- The “vision” of a deductive GIS
  - The domain & domain assumptions
  - Desirable reasoning tasks
- Implementation issues
  - Representing the data
  - Value of a DL system in this scenario

Contents

- Report about <u>practical work in progress</u> carried out in the context of the "DLS" project
  - Vision of a hybrid deductive GIS
  - Introduction of a first prototype
  - A software framework for similar "hybrid" KR&R tasks
- The "vision" of a deductive GIS
  - The domain & domain assumptions
  - Desirable reasoning tasks
- Implementation issues
  - Representing the data
  - Value of a DL system in this scenario
  - "Hybrid conjunctive queries"

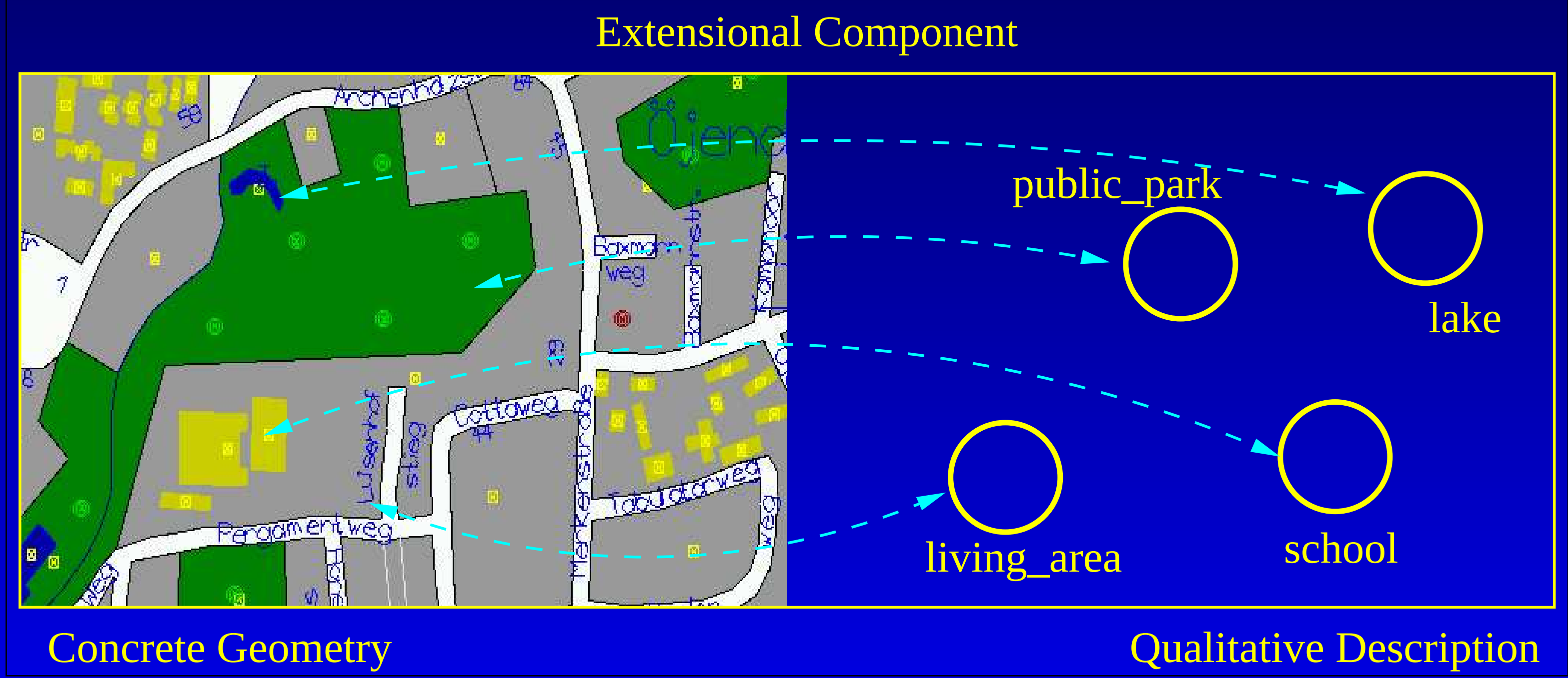
The Vision of a Deductive GIS



Concrete Geometry

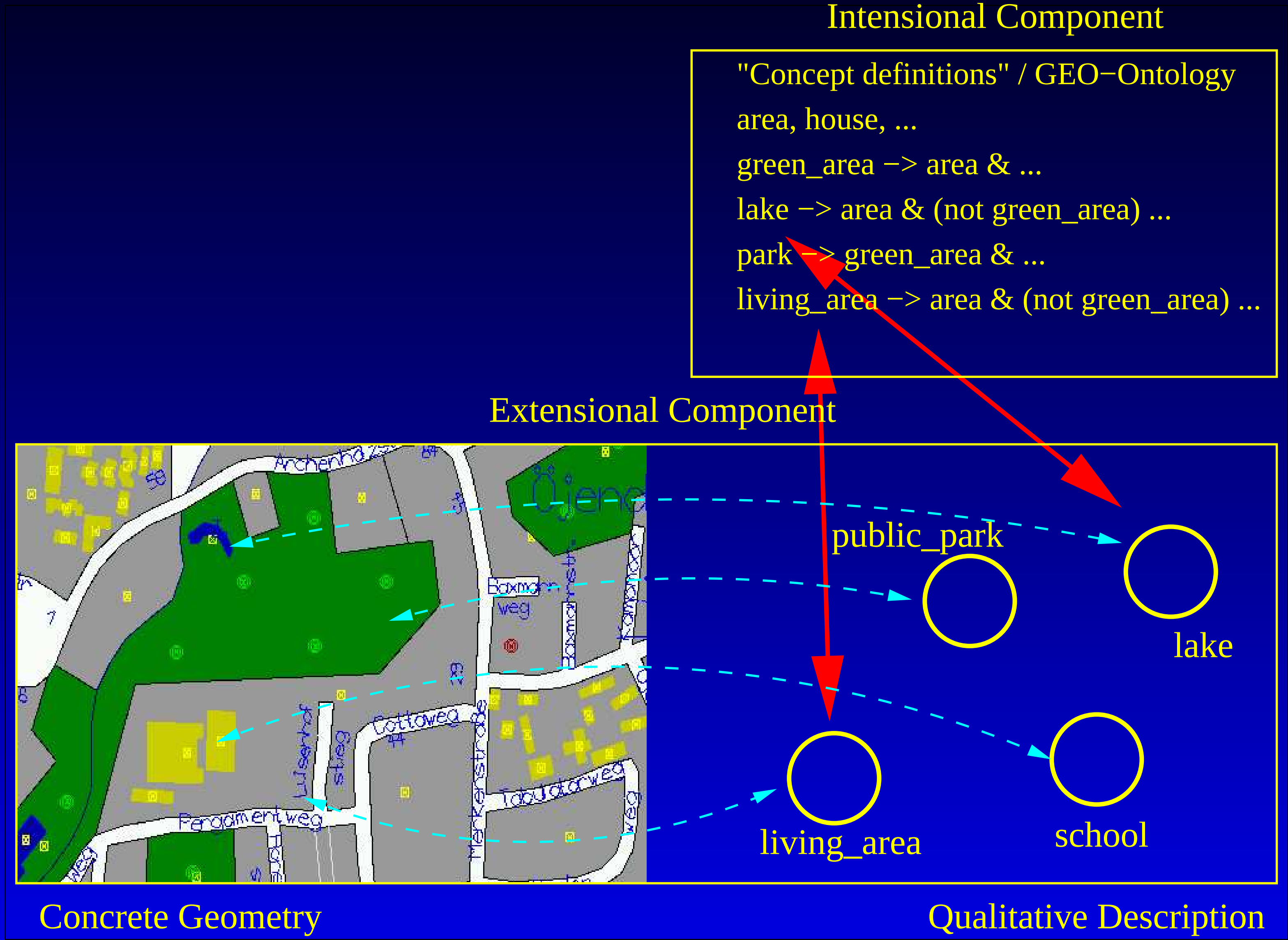
Starting point: a digital vector map

The Vision of a Deductive GIS



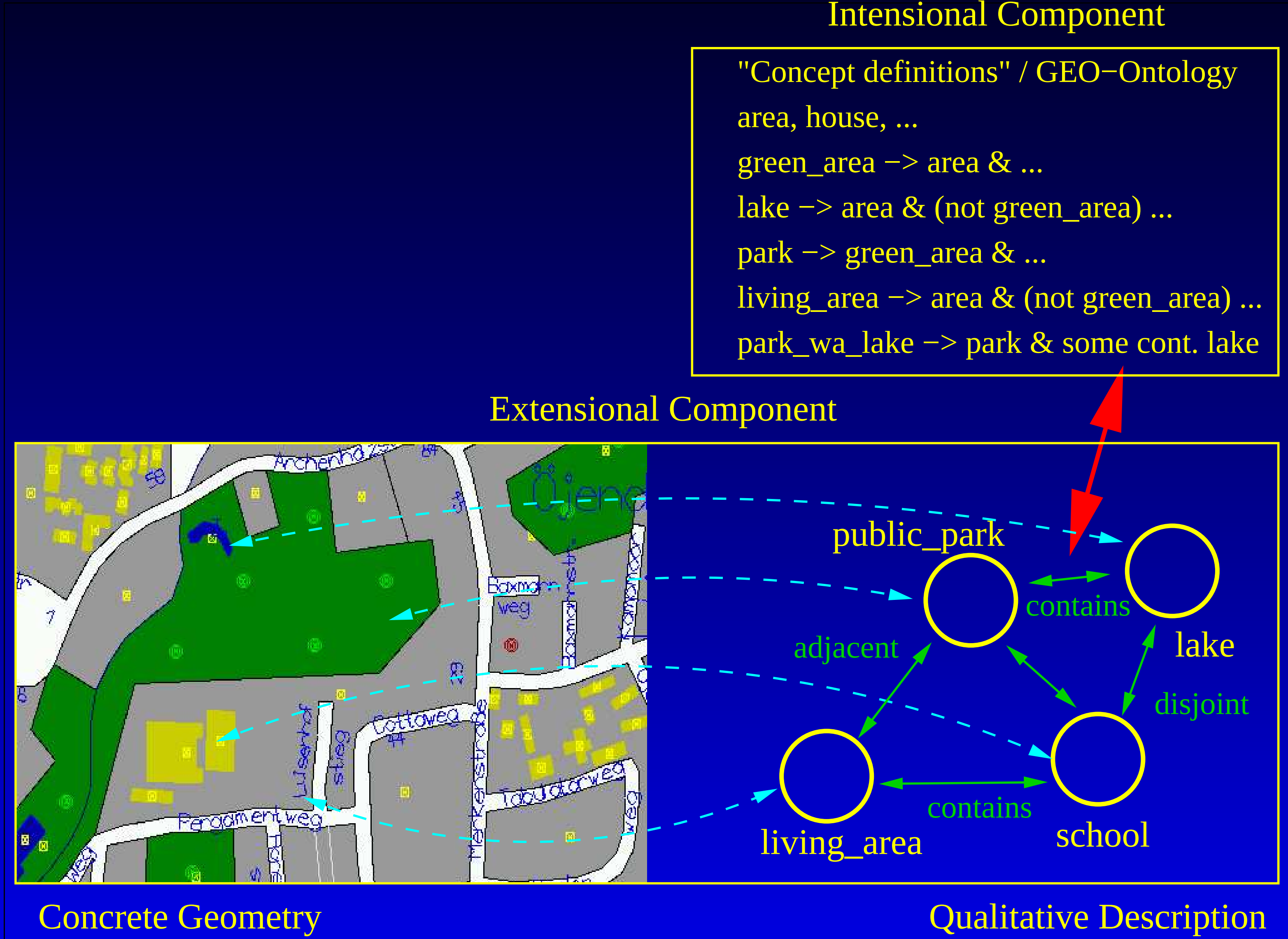
Thematic information in a map

The Vision of a Deductive GIS



Modeling of thematic concepts

The Vision of a Deductive GIS



Some concepts are really “spatio-thematic”

The Vision of a Deductive GIS

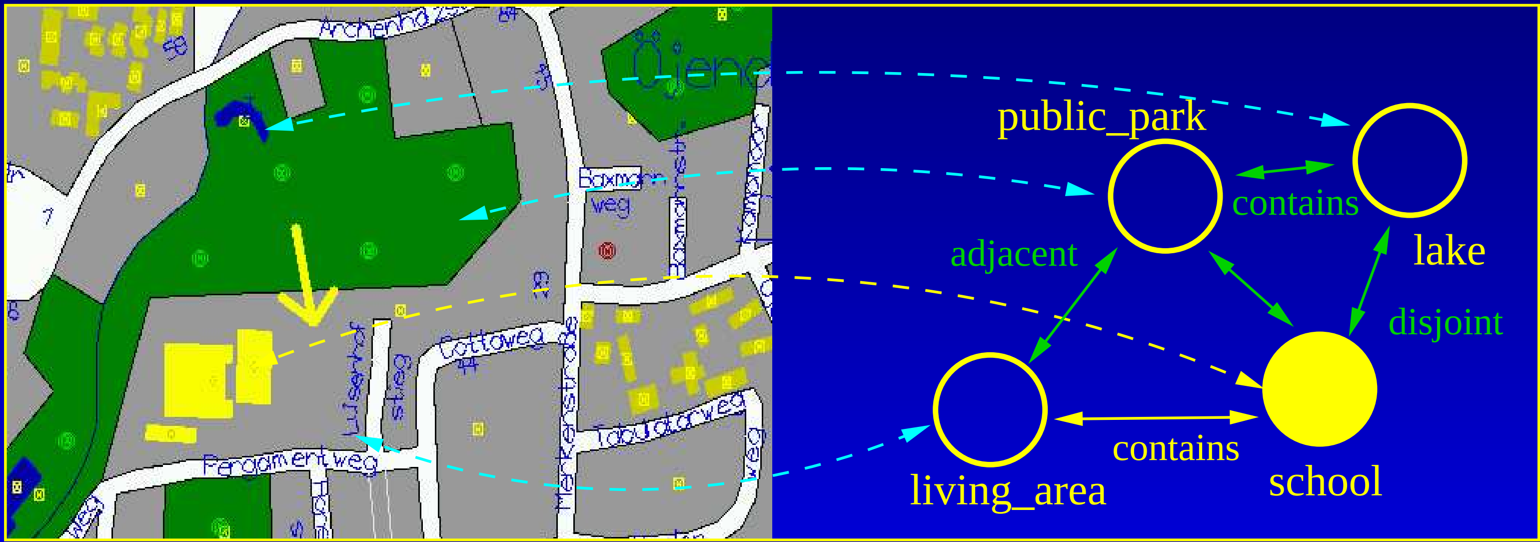
Query Component

Simple Spatial Queries:

Retrieve all areas contained within this area

Intensional Component

Extensional Component



Concrete Geometry

Qualitative Description

Purely spatial queries

The Vision of a Deductive GIS

Query Component

Simple Spatial Queries:

Retrieve all areas contained within this area

Thematic Queries:

retrieve_concept_instances(lake)

Intensional Component

"Concept definitions" / GEO−Ontology

area, house, ...

green_area −> area & ...

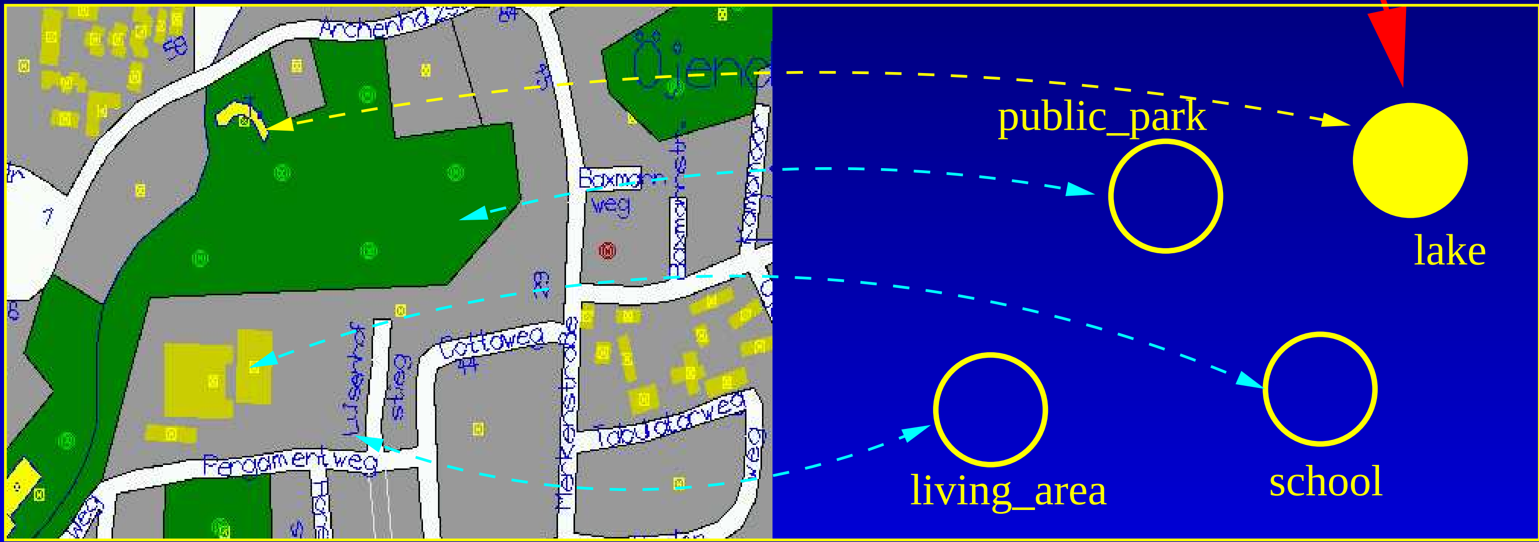
lake −> area & (not green_area) ...

park −> green_area & ...

living_area −> area & (not green_area) ...

park_wa_lake −> park & some cont. lake

Extensional Component



Concrete Geometry

Qualitative Description

Purely thematic queries

The Vision of a Deductive GIS

Query Component

Simple Spatial Queries:

- Retrieve all areas contained within this area

Thematic Queries:

- retrieve_concept_instances(lake)

Spatio−thematic Queries:

- Retrieve all parks that contain a lake

Intensional Component

"Concept definitions" / GEO−Ontology

area, house, ...

green_area −> area & ...

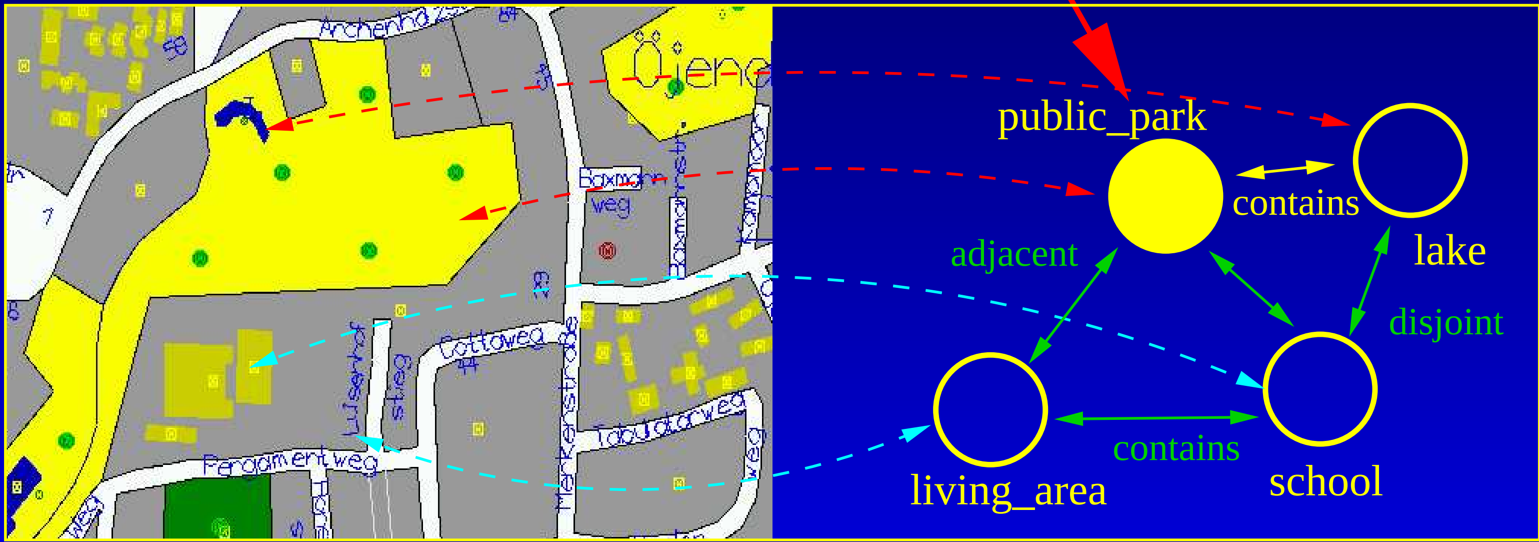
lake −> area & (not green_area) ...

park −> green_area & ...

living_area −> area & (not green_area) ...

park_wa_lake −> park & some cont. lake

Extensional Component



Concrete Geometry | Qualitative Description

“Spatio-thematic” queries

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
  - Representation of certain selected spatio-thematic aspects of a concrete map ("geographic world")
  - ? Which spatial and thematic aspects?
  - ? Data vs. information / knowledge?
  - ? Unified or hybrid representation of spatial and/or thematic aspects (different "sources")?

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
  - Modeling of ontologies with “concepts” in a description language (not necessary DL)
  - ? Which spatial and thematic aspects?
  - ? Thematic, spatial, spatio-thematic concepts?
  - ? Combined or separated description languages for different aspects?
  - ? Spatio-thematic interaction?

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
  - Retrieval of interesting objects / constellations; "map analysis / reasoning"
  - ? Kind of queries
  - ? With spatial and thematic aspects are addressable?
  - ? Usage of concepts from the ontologies within queries
  - ? Evaluation of queries ("specialists" for sources)?

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
- Reasoning tasks

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
- Reasoning tasks
  - $\mathcal{E}, \mathcal{I}$: consistency checking
  - $\mathcal{I}, \mathcal{Q}$: satisfi ability and entailment of queries / concepts
  - $\mathcal{I}, \mathcal{Q}$: computation of query / concept subsumption hierarchies (“taxonomies”)
  - $\mathcal{E} \times \mathcal{I}$: instance “realization”
  - $\mathcal{Q} \times \mathcal{E} \times \mathcal{I}$: query answering using vocabulary from $\mathcal{I}$

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
- Reasoning tasks

$\Rightarrow$ Multi-dimensional space of design-decisions

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
- Reasoning tasks

⇒ Multi-dimensional space of design-decisions

? How can RACER be of value in this setting (RACER offers $\mathcal{ALCQHI}_{\mathcal{R}^+}(\mathcal{D}^-)$, but is not a "spatio-thematic" DL)

The Vision of a Deductive GIS

- System metaphor: we want a GIS similar to a DL system
- Extensional component $\mathcal{E}$
- Intensional component $\mathcal{I}$
- Query component $\mathcal{Q}$
- Reasoning tasks

$\Rightarrow$ Multi-dimensional space of design-decisions

? How can RACER be of value in this setting (RACER offers $\mathcal{ALCQHI}_{\mathcal{R}^+}(\mathcal{D}^-)$, but is not a "spatio-thematic" DL)

- Development of a flexible software OO framework allowing for experiments

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”

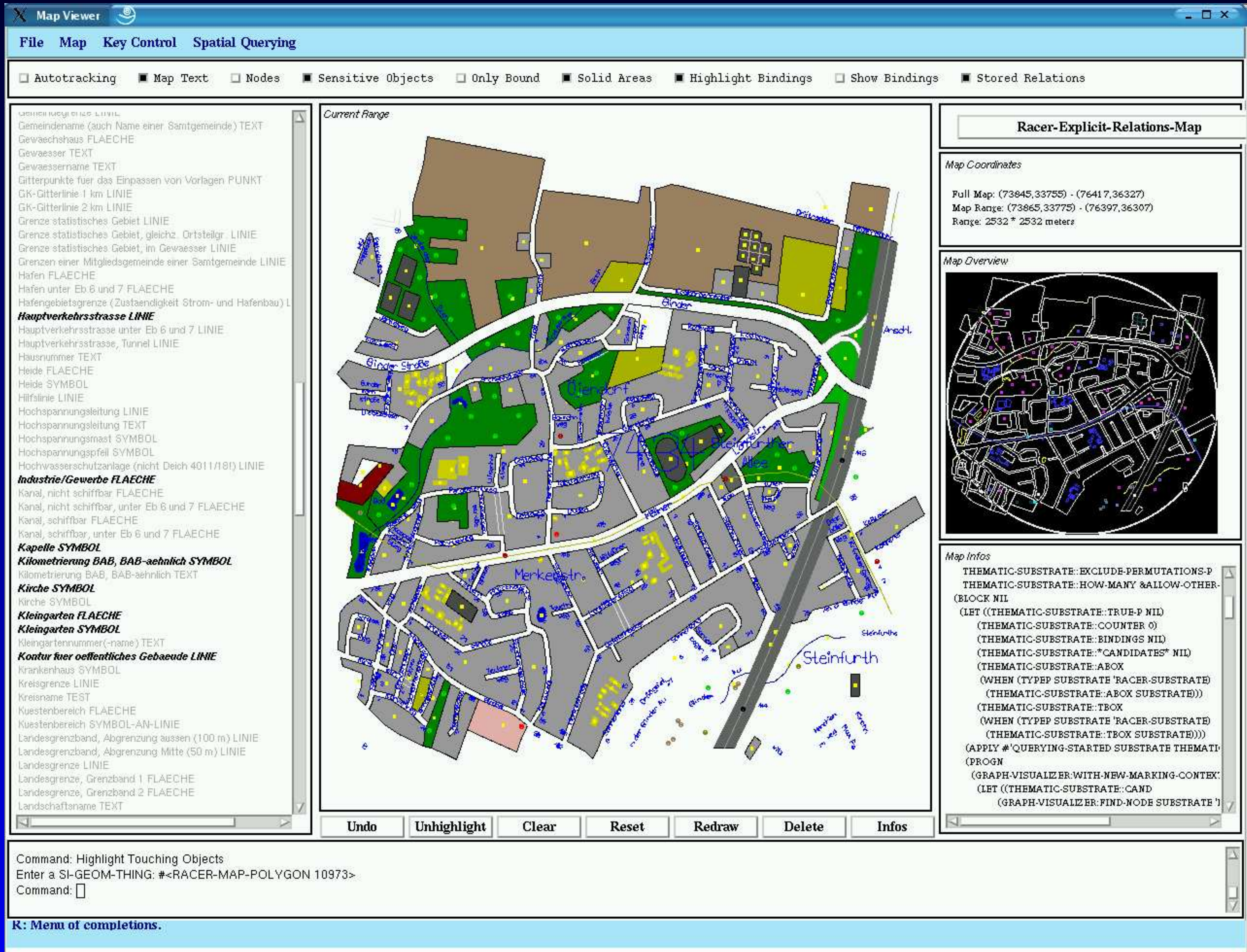
The Data

- Data from the "Amt für Vermessung und Geo-Information Hamburg"
- Two digital vector maps in the proprietary SQD format

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”
- Two digital vector maps in the proprietary SQD format
  - Map 1: 2694 objects, 361 primary objects

The Data

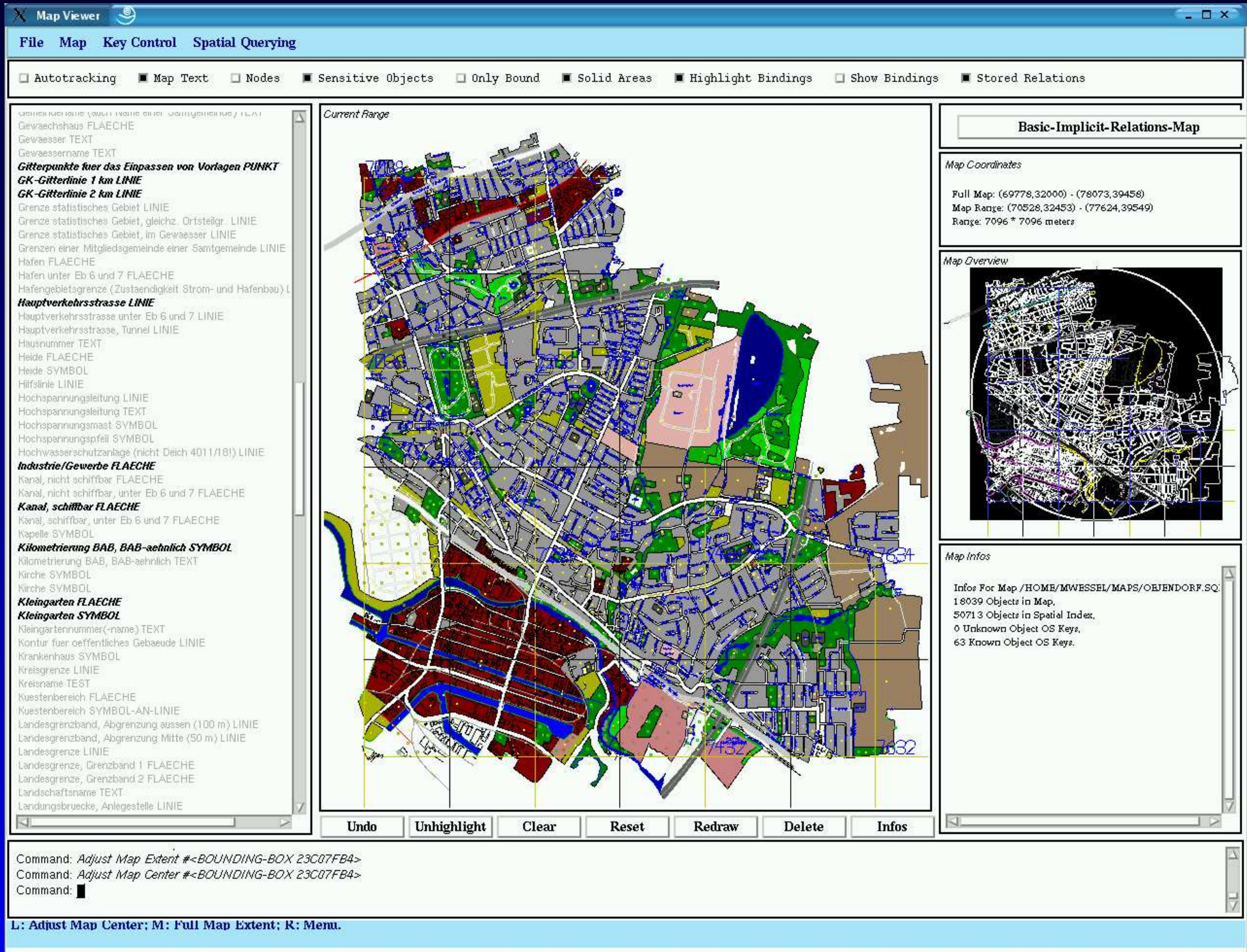


Map Viewer
File
Map
Key Control
Spatial Querying
Autotracking
Map Text
Nodes
Sensitive Objects
Only Bound
Solid Areas
Highlight Bindings
Show Bindings
Stored Relations
Gemeindename (auch Name einer Samtgemeinde) TEXT
Gewaechshaus FLAECHE
Gewaesser TEXT
Gewaessername TEXT
Gitterpunkte fuer das Einpassen von Vorlagen PUNKT
GK-Gitterlinie 1 km LINIE
GK-Gitterlinie 2 km LINIE
Grenze statistisches Gebiet LINIE
Grenze statistisches Gebiet, gleichz. Ortsteilgr. LINIE
Grenze statistisches Gebiet, im Gewaesser LINIE
Grenzen einer Mitgliedsgemeinde einer Samtgemeinde LINIE
Hafen FLAECHE
Hafen unter Eb 6 und 7 FLAECHE
Hauptverkehrsstrasse LINIE
Hauptverkehrsstrasse unter Eb 6 und 7 LINIE
Hauptverkehrsstrasse, Tunnel LINIE
Hausnummer TEXT
Heide FLAECHE
Heide SYMBOL
Hilfslinie LINIE
Hochspannungsleitung LINIE
Hochspannungsleitung TEXT
Hochspannungsmast SYMBOL
Hochspannungspfeil SYMBOL
Industrie/Gewerbe FLAECHE
Kanal, nicht schiffbar FLAECHE
Kanal, nicht schiffbar, unter Eb 6 und 7 FLAECHE
Kanal, schiffbar FLAECHE
Kanal, schiffbar, unter Eb 6 und 7 FLAECHE
Kapelle SYMBOL
Kilometrierung BAB, BAB-aehnlich SYMBOL
Kilometrierung BAB, BAB-aehnlich TEXT
Kirche SYMBOL
Kirche SYMBOL
Kleingarten FLAECHE
Kleingarten SYMBOL
Kleingartennummer(-name) TEXT
Kontur fuer oeffentliches Gebaeude LINIE
Krankenhaus SYMBOL
Kreisgrenze LINIE
Kreisname TEXT
Kuestenbereich FLAECHE
Kuestenbereich SYMBOL-AN-LINIE
Landesgrenzband, Abgrenzung aussen (100 m) LINIE
Landesgrenzband, Abgrenzung Mitte (50 m) LINIE
Landesgrenze LINIE
Landesgrenze, Grenzband 1 FLAECHE
Landesgrenze, Grenzband 2 FLAECHE
Landschaftsname TEXT
Current Range
Steinfurth
Racer-Explicit-Relations-Map
Map Coordinates
Full Map: (73845,33755) - (76417,36327)
Map Range: (73865,33775) - (76397,36307)
Range: 2532 * 2532 meters
Map Overview
Map Infos
Undo
Unhighlight
Clear
Reset
Redraw
Delete
Infos
Command: Highlight Touching Objects
Enter a SI-GEOM-THING: #<RACER-MAP-POLYGON 10973>
Command:
R: Menu of completions.

The Data

- Data from the "Amt für Vermessung und Geo-Information Hamburg"
- Two digital vector maps in the proprietary SQD format
  - Map 2: 18.039 geometric objects, 5.418 primary object

The Data


Map Viewer
File Map Key Control Spatial Querying
Autotracking Map Text Nodes Sensitive Objects Only Bound Solid Areas Highlight Bindings Show Bindings Stored Relations
Gewaechshaus FLAECHE
Gewaesser TEXT
Gewaessername TEXT
Gitterpunkte fuer das Einpassen von Vorlagen PUNKT
GK-Gitterlinie 1 km LINIE
GK-Gitterlinie 2 km LINIE
Grenze statistisches Gebiet LINIE
Grenze statistisches Gebiet, gleichz. Ortsteilgr. LINIE
Grenze statistisches Gebiet, im Gewaesser LINIE
Grenzen einer Mitgliedsgemeinde einer Samtgemeinde LINIE
Hafen FLAECHE
Hafen unter Eb 6 und 7 FLAECHE
Hauptverkehrsstrasse LINIE
Hauptverkehrsstrasse unter Eb 6 und 7 LINIE
Hauptverkehrsstrasse, Tunnel LINIE
Hausnummer TEXT
Heide FLAECHE
Heide SYMBOL
Hilfslinie LINIE
Hochspannungsleitung LINIE
Hochspannungsleitung TEXT
Hochspannungsmast SYMBOL
Hochspannungspfeil SYMBOL
Industrie/Gewerbe FLAECHE
Kanal, nicht schiffbar FLAECHE
Kanal, nicht schiffbar, unter Eb 6 und 7 FLAECHE
Kanal, schiffbar FLAECHE
Kanal, schiffbar, unter Eb 6 und 7 FLAECHE
Kapelle SYMBOL
Kilometrierung BAB, BAB-aehnlich SYMBOL
Kilometrierung BAB, BAB-aehnlich TEXT
Kirche SYMBOL
Kleingarten FLAECHE
Kleingarten SYMBOL
Kontur fuer oeffentliches Gebaeude LINIE
Krankenhaus SYMBOL
Kreisgrenze LINIE
Kreisname TEXT
Kuestenbereich FLAECHE
Landesgrenze LINIE
Landschaftsname TEXT
Landungsbruecke, Anlegestelle LINIE
Current Range
Basic-Implicit-Relations-Map
Map Coordinates
Full Map: (69778,32000) - (78073,39458)
Map Range: (70528,32453) - (77624,39549)
Range: 7096 * 7096 meters
Map Overview
Map Infos
18039 Objects in Map.
50713 Objects in Spatial Index.
0 Unknown Object OS Keys.
63 Known Object OS Keys.
Undo Unhighlight Clear Reset Redraw Delete Infos
Command: Adjust Map Extent #<BOUNDING-BOX 23C07FB4>
Command: Adjust Map Center #<BOUNDING-BOX 23C07FB4>
Command:
L: Adjust Map Center; M: Full Map Extent; R: Menu.

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”
- Two digital vector maps in the proprietary SQD format
- Objects are “classified” according to object key dictionary:

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”
- Two digital vector maps in the proprietary SQD format
- Objects are “classified” according to object key dictionary:
  - 5164 $\Rightarrow$ $\boxed{\text{lake, navigable}}$, 4128 $\Rightarrow$ $\boxed{\text{meadow}}$, 2224 $\Rightarrow$ $\boxed{\text{park}}$, 2119 $\Rightarrow$ $\boxed{\text{living area,}}$ …

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”
- Two digital vector maps in the proprietary SQD format
- Objects are “classified” according to object key dictionary:
  - 5164 $\Rightarrow$ $\boxed{\text{lake, navigable}}$, 4128 $\Rightarrow$ $\boxed{\text{meadow}}$, 2224 $\Rightarrow$ $\boxed{\text{park}}$, 2119 $\Rightarrow$ $\boxed{\text{living area,}}$ …
  - Subsumption implicitly present, but not explicitly modeled $\Rightarrow$ needs remodeling

The Data

- Data from the “Amt für Vermessung und Geo-Information Hamburg”
- Two digital vector maps in the proprietary SQD format
- Objects are “classified” according to object key dictionary:
  - 5164 ⇒ $\boxed{\text{lake, navigable}}$, 4128 ⇒ $\boxed{\text{meadow}}$, 2224 ⇒ $\boxed{\text{park}}$, 2119 ⇒ $\boxed{\text{living area,}}$ …
  - Subsumption implicitly present, but not explicitly modeled ⇒ needs remodeling
  - Some concepts really have a spatio-thematic flavor, e.g. $\boxed{\text{park with (containing) a lake}}$

Design Decisions

- Various representation possibilities for the map

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation

Design Decisions

- Various representation possibilities for the map
- Concrete spatial "data": use a spatially indexed geometric representation
  - Closed Domain Assumption (there are no other spatial objects than the present ones)

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
  - Closed Domain Assumption (there are no other spatial objects than the present ones)
  - Closed World Assumption (complete theory of a fixed single structure)

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:
  - Setting 1: Modeled as a RACER ABox with Concept Membership Assertions like
    $\boxed{area_{123} : lake \sqcup meadow \sqcup \ldots}$

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:
  - Setting 2: Like Setting 1, but additionally with Role Membership Assertions like $\boxed{(area_{123}, area_{456}) : contains}$, mirroring qualitative spatial relationships as found in the map (e.g., using RCC8 relationships)

Design Decisions

- Various representation possibilities for the map
- Concrete spatial "data": use a spatially indexed geometric representation
- Qualitative (spatio-)thematic "information":
  - Setting 3: Do not use a RACER ABox, but simply annotate map objects with RACER concept expressions (or expressions of an other reasoning engine)

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:
    - Setting 4: Do not use RACER at all, but implement your own “truly” spatio-thematic DL

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:
  - Setting 5: Don’t even use a DL

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- Qualitative (spatio-)thematic “information”:

  $\Rightarrow$ In order to allow for flexible experiments, description languages are not hard-wired into the software framework; e.g. we can use the same framework if we change the spatio-thematic description vocabulary, e.g. switching from RCC8 relationships to qualitative <u>distance</u> relationships

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic “information”: use your favorite description language

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic “information”: use your favorite description language
- Notion of a “reasoning substrate”:

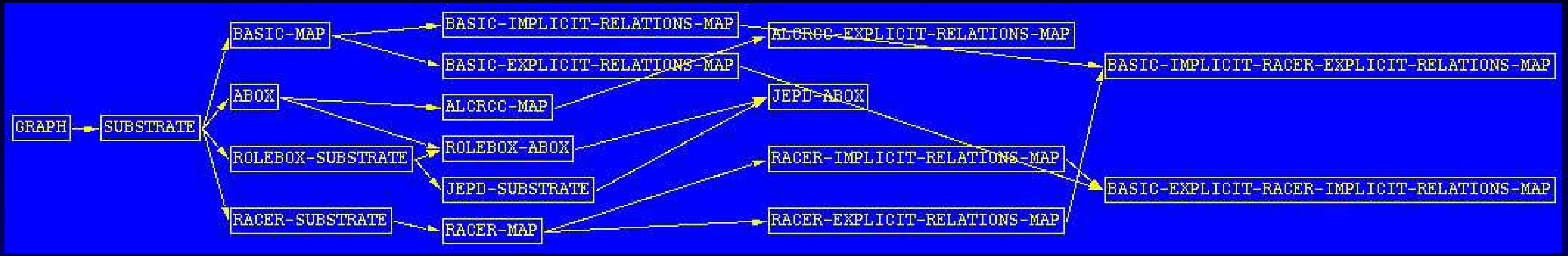
Design Decisions

- Various representation possibilities for the map
- Concrete spatial "data": use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic "information": use your favorite description language
- Notion of a "reasoning substrate":
  - General-purpose "labeled graph"-like notion with exchangeable node and edge labeling languages

Design Decisions

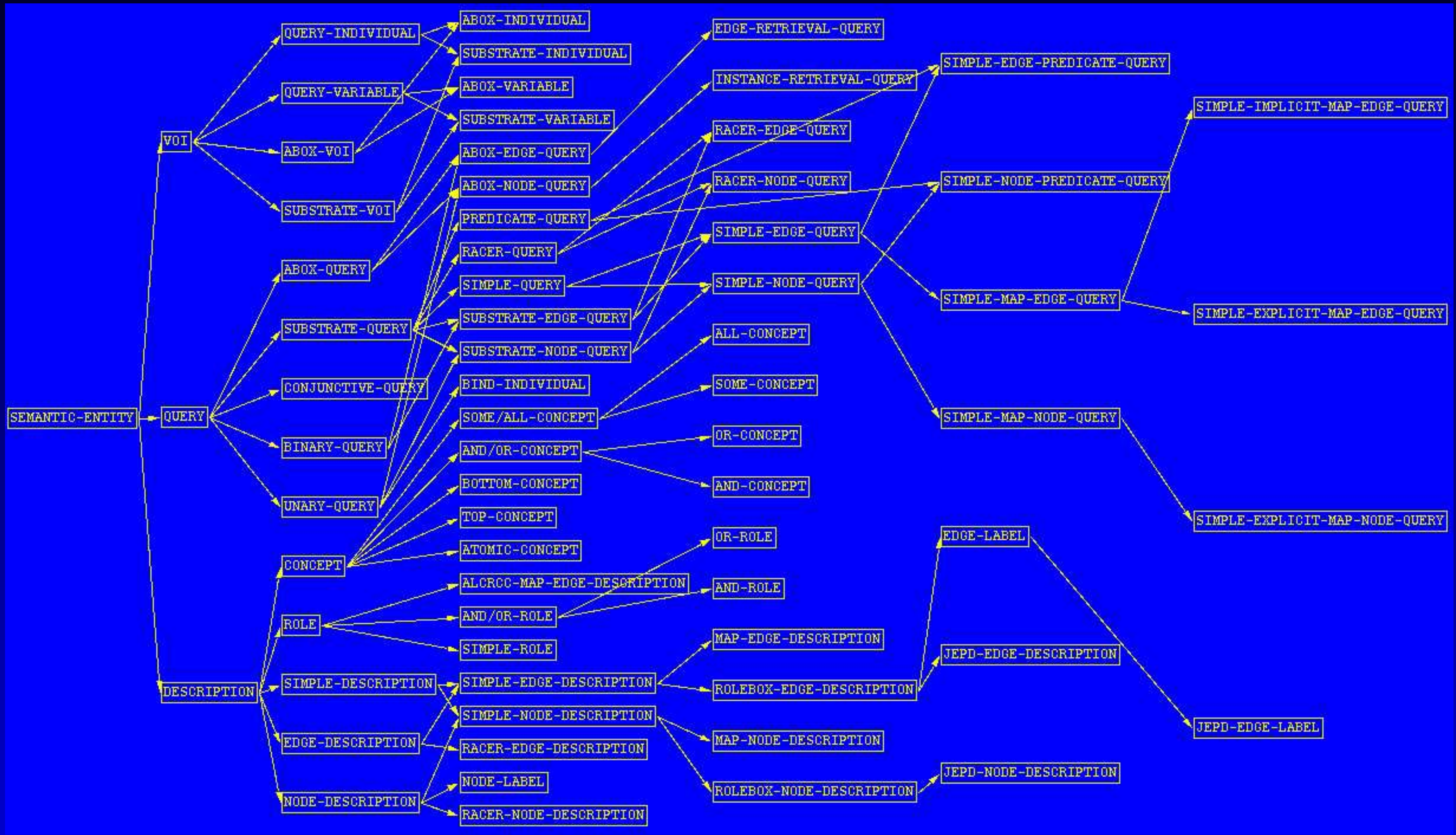
- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic “information”: use your favorite description language
- Notion of a “reasoning substrate”:
  - General-purpose “labeled graph”-like notion with exchangeable node and edge labeling languages
  - Use inheritance to get specialized substrate classes, languages and reasoners

Design Decisions



Subclasses of class `substrate`

Design Decisions



Subclasses of class `semantic entity`

Design Decisions

- Various representation possibilities for the map
- Concrete spatial "data": use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic "information": use your favorite description language
- Notion of a "reasoning substrate":
  - General-purpose "labeled graph"-like notion with exchangeable node and edge labeling languages
  - Use inheritance to get specialized substrate classes, languages and reasoners
  - Special-purpose index structures

Design Decisions

- Various representation possibilities for the map
- Concrete spatial “data”: use a spatially indexed geometric representation
- (Qualitative) (spatio-)thematic “information”: use your favorite description language
- Notion of a “reasoning substrate”:
  - General-purpose “labeled graph”-like notion with exchangeable node and edge labeling languages
  - Use inheritance to get specialized substrate classes, languages and reasoners
  - Special-purpose index structures
  - DL-system inspired protocols (interfaces)

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; “object keys” are remodeled as (quite simple) RACER TBox

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox

  $\Rightarrow$ even a simple ontology is of great value; the query

  $\boxed{\text{retrieve\_concept\_instances(green\_area)}}$

  would not return instances of the (intuitive) sub-concepts $meadow$ and $park$ otherwise

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox,"Map Substrate") = "Racer Map Substrate": exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; “object keys” are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox,“Map Substrate”) = “Racer Map Substrate”: exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
  - Map 1: 130.321 RMAs if we represent the disconnected relationship DC, 1804 without DC

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; “object keys” are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox, “Map Substrate”) = “Racer Map Substrate”: exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
  - Map 2: 29.354.724 with DC, 19.988 without DC

An Experiment with RACER

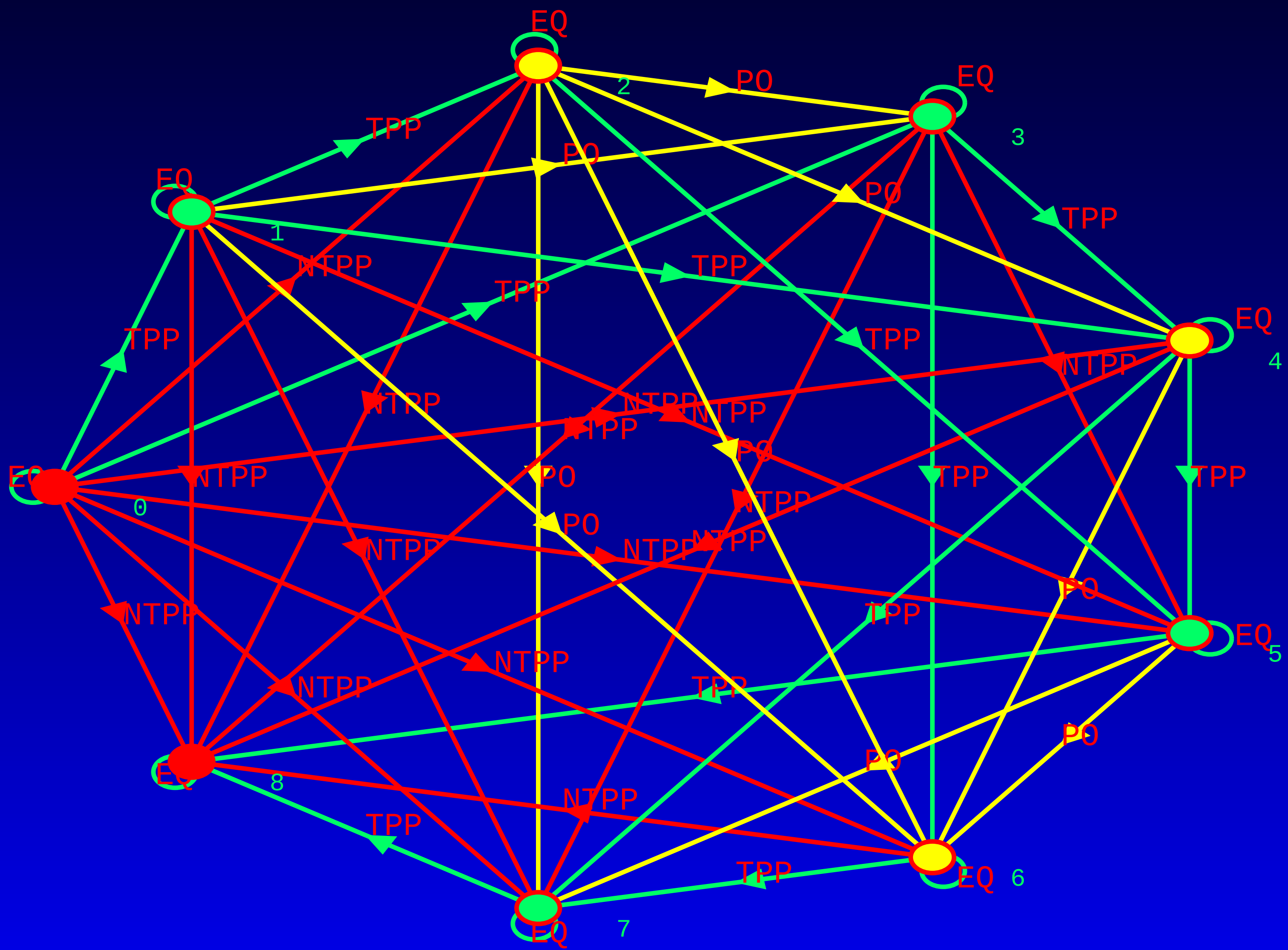
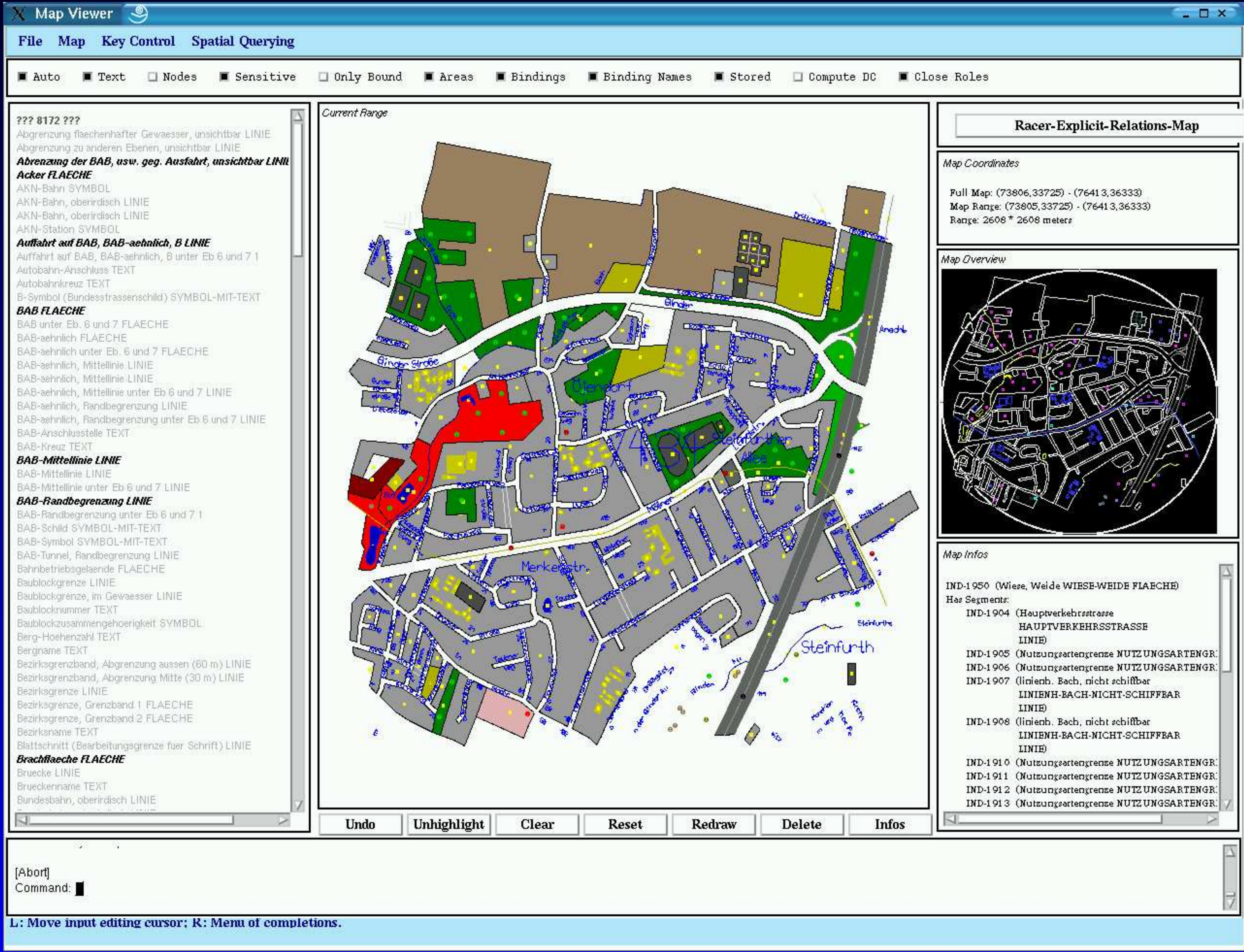


Illustration of a typical ABox

An Experiment with RACER

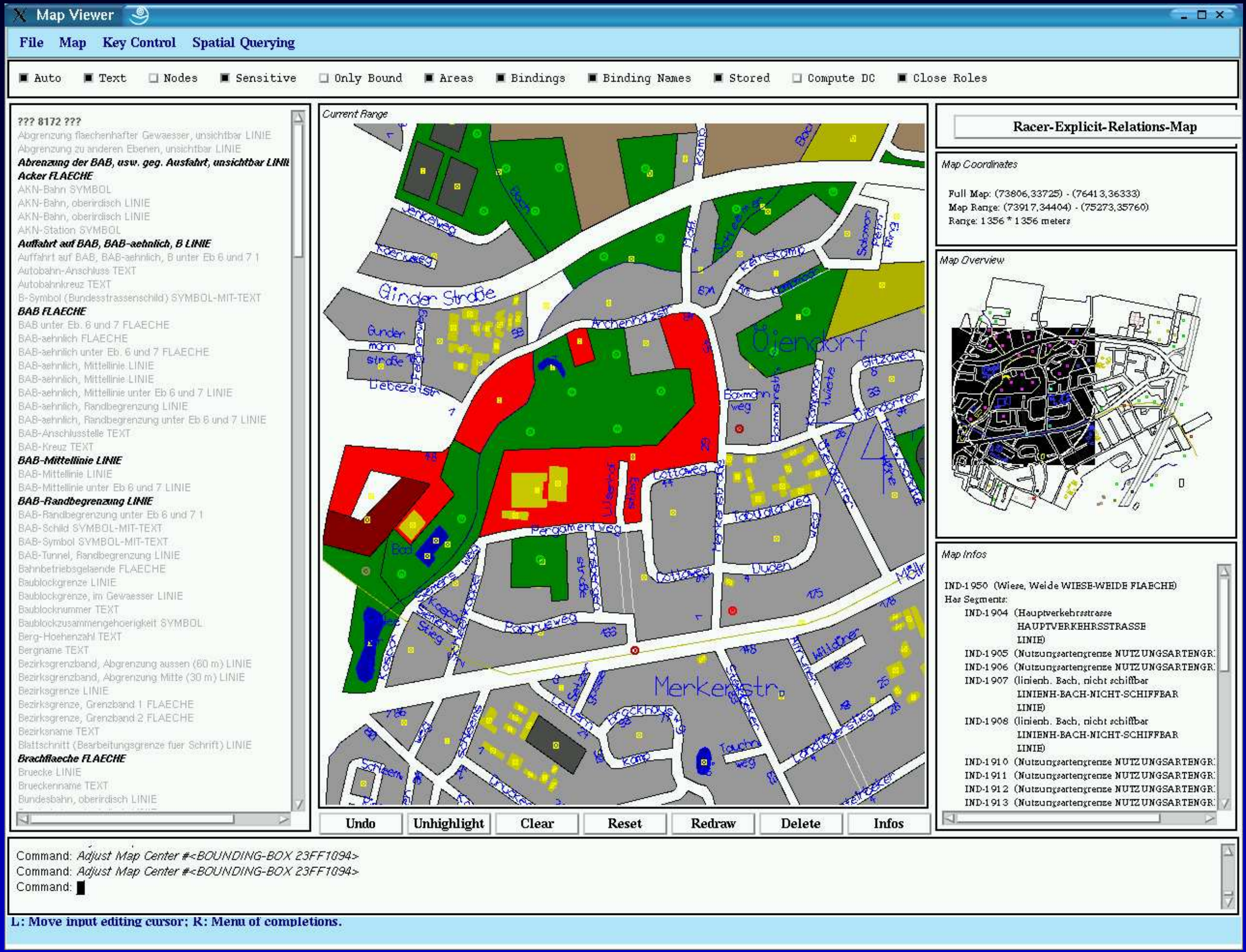
- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox, "Map Substrate") = "Racer Map Substrate": exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
- Pose simple instance retrieval queries to RACER

An Experiment with RACER



$$green\_area \sqcap \exists contains.lake$$

An Experiment with RACER



$$living\_area \sqcap$$
$$\forall adjacent.(green\_area \sqcup parking\_lot) \sqcap$$
$$\exists adjacent.(park \sqcap \exists contains.lake)$$

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox, "Map Substrate") = "Racer Map Substrate": exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
- Pose simple instance retrieval queries to RACER
- Closing of spatial roles (R) required to realize "spatial closed domain assumption" in order to answer $\forall R.C$ queries correctly

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox, "Map Substrate") = "Racer Map Substrate": exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
- Pose simple instance retrieval queries to RACER
- Closing of spatial roles (R) required to realize "spatial closed domain assumption" in order to answer $\forall R.C$ queries correctly

$\Rightarrow$ add $i : (\leq\ n\ R) \sqcap (\geq\ n\ R)$ to individual i, where $n =_{def} |\{\, j \mid (i, j) : R \in \mathfrak{A}\,\}|$

An Experiment with RACER

- $\mathcal{I} = TBox$: modeling of purely thematic concepts; "object keys" are remodeled as (quite simple) RACER TBox
- $\mathcal{E}$ = (ABox, "Map Substrate") = "Racer Map Substrate": exhaustively add RCC8 role membership assertions, computed from the geometry, and thematic descriptions to the ABox
- Pose simple instance retrieval queries to RACER
- Closing of spatial roles (R) required to realize "spatial closed domain assumption" in order to answer $\forall R.C$ queries correctly
- RACER 1.7.7 performs much better than RACER 1.7.6, but only until we add DC and close the roles

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
  - Inherently unsatisfiable queries will not be recognized as such

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
  - Inherently unsatisfiable queries will not be recognized as such
  - TBox (resp. $\mathcal{I}$): might become inconsistent without being noticed, missing subsumption relationships etc. $\Rightarrow$ model only purely thematic concepts

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
  - Inherently unsatisfiable queries will not be recognized as such
  - TBox (resp. $\mathcal{I}$): might become inconsistent without being noticed, missing subsumption relationships etc. $\Rightarrow$ model only purely thematic concepts
  - "Query subsumption" incomplete, but okay for optimization purposes (caching/reusing of answer sets)

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)
  - RACER cannot handle even small ABoxes with DC and closed roles

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)
  - RACER cannot handle even small ABoxes with DC and closed roles
  - Explicit representation of 29 million role membership assertions is not a good idea

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)
  - RACER cannot handle even small ABoxes with DC and closed roles
  - Explicit representation of 29 million role membership assertions is not a good idea
  - "Specialized reasoners" will perform much better

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)

$\Rightarrow$ Using RACER in this way seems to be inappropriate

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)

$\Rightarrow$ Using RACER in this way seems to be inappropriate

$\Rightarrow$ Implementation of special-purpose reasoners

Problems with the Approach

- Incompleteness of reasoning in $\mathcal{I}$, $\mathcal{Q}$
- Since the ABox is "correctly closed", query answering is complete (assuming an unfoldable TBox) $\Rightarrow$ RACER performs "spatial closed domain reasoning"
- RACER has problems with the specific structure of the ABoxes (probably "worst case" for a tableaux-based system!)

$\Rightarrow$ Using RACER in this way seems to be inappropriate

$\Rightarrow$ Implementation of special-purpose reasoners

$\Rightarrow$ We can still use RACER for "sub-reasoning" tasks

More Expressive Queries

- Simple non-recursive conjunctive queries

More Expressive Queries

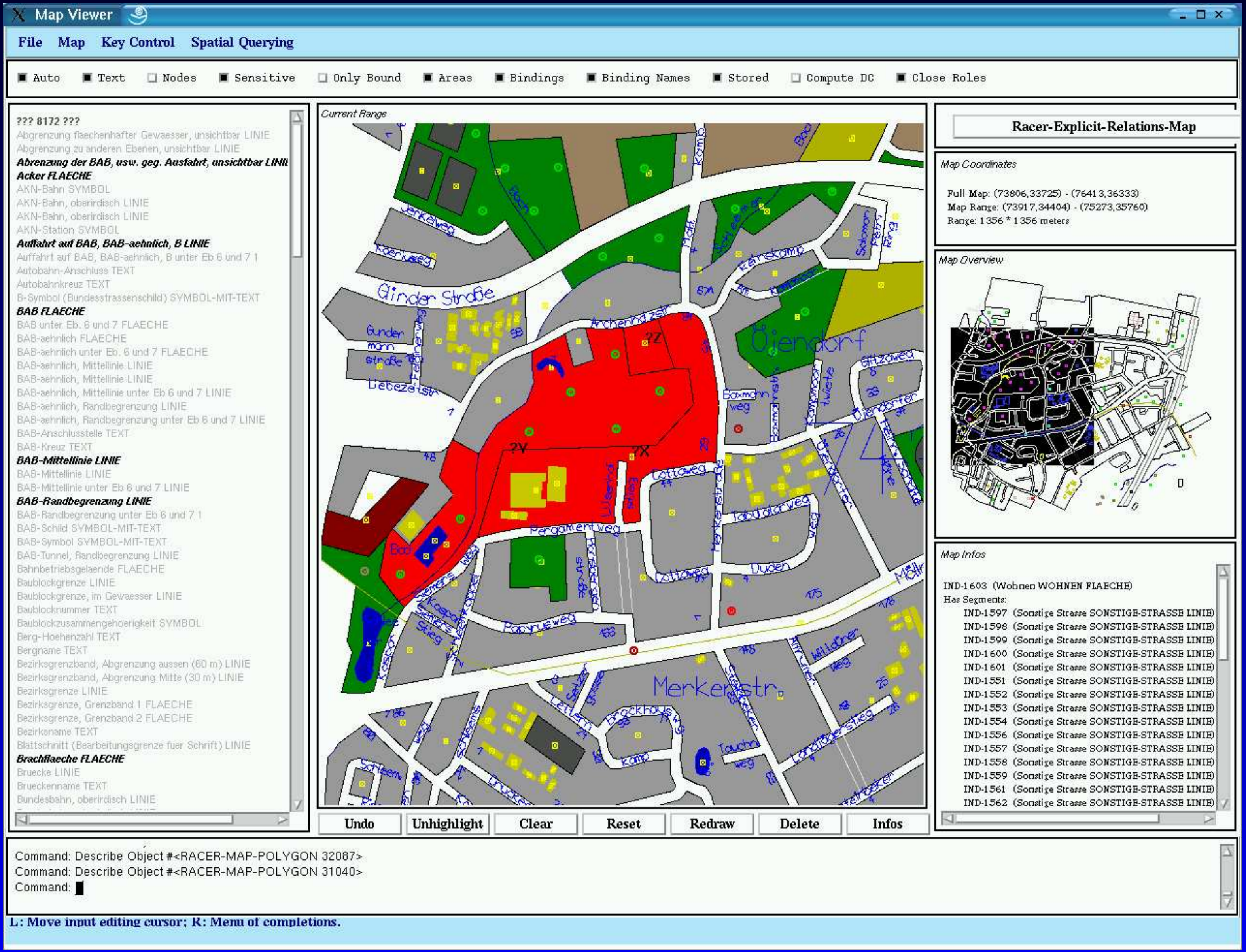
- Simple non-recursive conjunctive queries
  - "Find a living area, a green area and a parking lot which are pairwise adjacent"
  - $query(?x, ?y, ?z) \leftarrow$

    $living\_area(?x), green\_area(?y), parking\_lot(?z),$

    $adjacent(?x, ?y), adjacent(?x, ?z), adjacent(?y, ?z)$

$\Rightarrow$ not expressible with standard DL concepts

More Expressive Queries

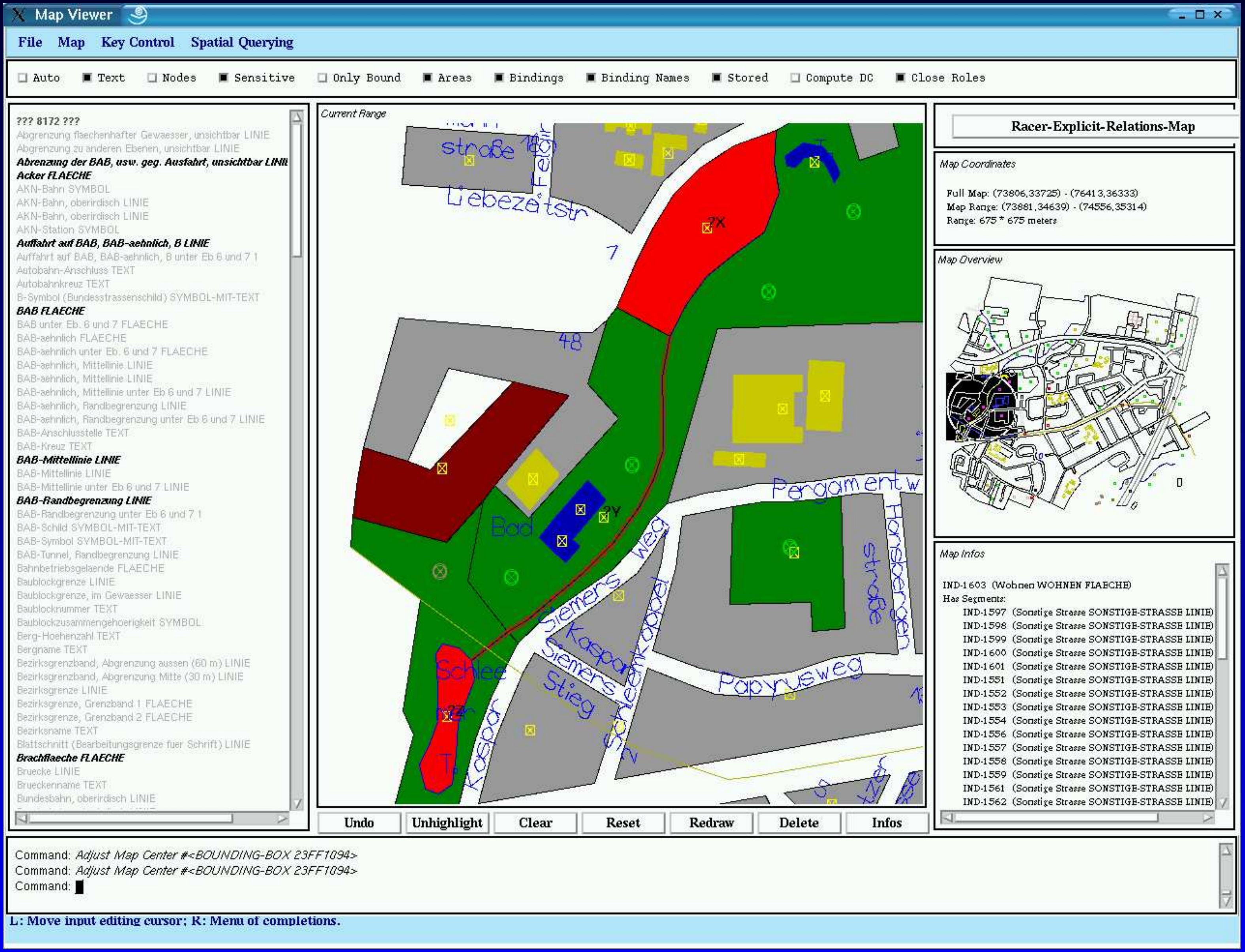


More Expressive Queries

- Simple non-recursive conjunctive queries
  - "Find a contaminated lake in a park in which a creek flows which borders an industrial area containing a chemical plant"
  - $query(?x, ?y, ?z, ?f) \leftarrow$

$$industrial\_area(?x), creek(?y), lake\_or\_pond(?z), contaminated(?z), chemical\_plant(?f), park(?u), borders(?y, ?x), flows\_in(?y, ?z), contains(?u, ?z), contains(?x, ?f)$$

More Expressive Queries



More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies
  $\longrightarrow$ "Plug in" definitions of terms

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies

$\longrightarrow$ "Plug in" definitions of terms

- $lake\_or\_pond(?z) \longrightarrow$
  $?z^* : (lake \sqcup pond)$

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies

  $\longrightarrow$ "Plug in" definitions of terms

  - $borders(?y, ?x), flows\_in(?y, ?x) \rightarrow EC(?y, ?x)$

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies

$\longrightarrow$ "Plug in" definitions of terms

- $contaminated(?z) \rightarrow$
  $?z^* : \exists water\_quality.poisoned$

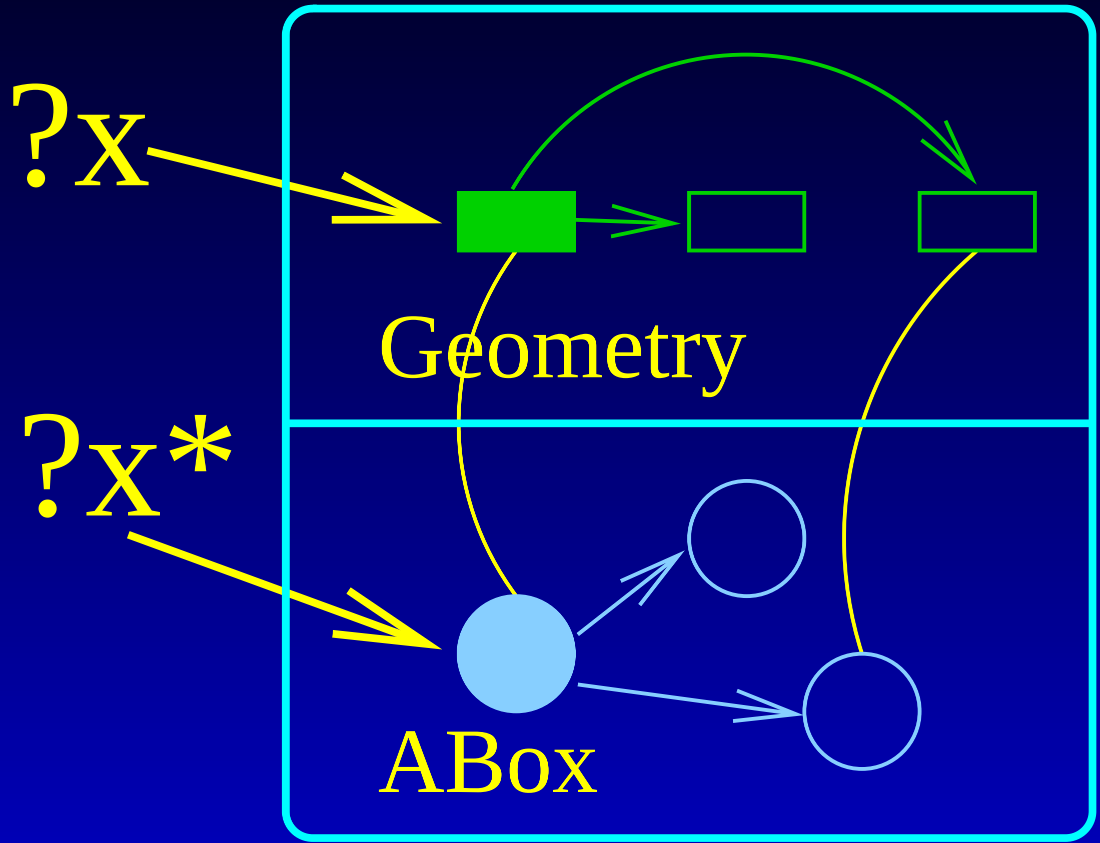
More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies
- Vision: according to where and how the data/information is represented ("sources"), queries will be "rewritten"

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies
- Vision: according to where and how the data/information is represented ("sources"), queries will be "rewritten"
- Result of the reformulation process: a "hybrid" spatio-thematic query which we can already process

More Expressive Queries



$?x^* / ?x$ are bound in parallel

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of “hybrid” spatio/thematic vocabulary from the ontologies
- Vision: according to where and how the data/information is represented (“sources”), queries will be “rewritten”
- Result of the reformulation process: a “hybrid” spatio-thematic query which we can already process
- “Active domain” semantics for variables (note: conjuncts like $?x^* : \exists R.C$ can be used)

More Expressive Queries

- Simple non-recursive conjunctive queries
- Queries make use of "hybrid" spatio/thematic vocabulary from the ontologies
- Vision: according to where and how the data/information is represented ("sources"), queries will be "rewritten"
- Result of the reformulation process: a "hybrid" spatio-thematic query which we can already process
- "Active domain" semantics for variables (note: conjuncts like $?x^* : \exists R.C$ can be used)
- Query processing: parsing $\longrightarrow$ plan generation $\longrightarrow$ plan optimization $\longrightarrow$ compilation $\longrightarrow$ execution

Reasoning about Queries

- Example: query consistency

Reasoning about Queries

- Example: query consistency
  - Two kinds of conjuncts: “RCC” and “ABox assertion” conjuncts

Reasoning about Queries

- Example: query consistency
  - Two kinds of conjuncts: “RCC” and “ABox assertion” conjuncts
  - Check satisfiabilty seperately

Reasoning about Queries

- Example: query consistency
  - Two kinds of conjuncts: "RCC" and "ABox assertion" conjuncts
  - Check satisfiabilty seperately
  - ABox assertions: construct an ABox from the conjuncts, replacing variables with individuals (RACER offers unique name assumption), check for ABox satisfiability

Reasoning about Queries

- Example: query consistency
  - Two kinds of conjuncts: “RCC” and “ABox assertion” conjuncts
  - Check satisfiabilty seperately
  - ABox assertions: construct an ABox from the conjuncts, replacing variables with individuals (RACER offers unique name assumption), check for ABox satisfiability
  - RCC conjuncts: construct an RCC network and check for its consistency

Reasoning about Queries

- Example: query consistency
  - Two kinds of conjuncts: “RCC” and “ABox assertion” conjuncts
  - Check satisfiabilty seperately
  - ABox assertions: construct an ABox from the conjuncts, replacing variables with individuals (RACER offers unique name assumption), check for ABox satisfiability
  - RCC conjuncts: construct an RCC network and check for its consistency
  - Conjecture: somehow “weak” since no interaction, but quite useful in this scenario, and complete (unlike using RACER concepts)

Reasoning about Queries

- Example: query consistency

  $\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks

Reasoning about Queries

- Example: query consistency

  $\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks

- Example: hybrid query containment

Reasoning about Queries

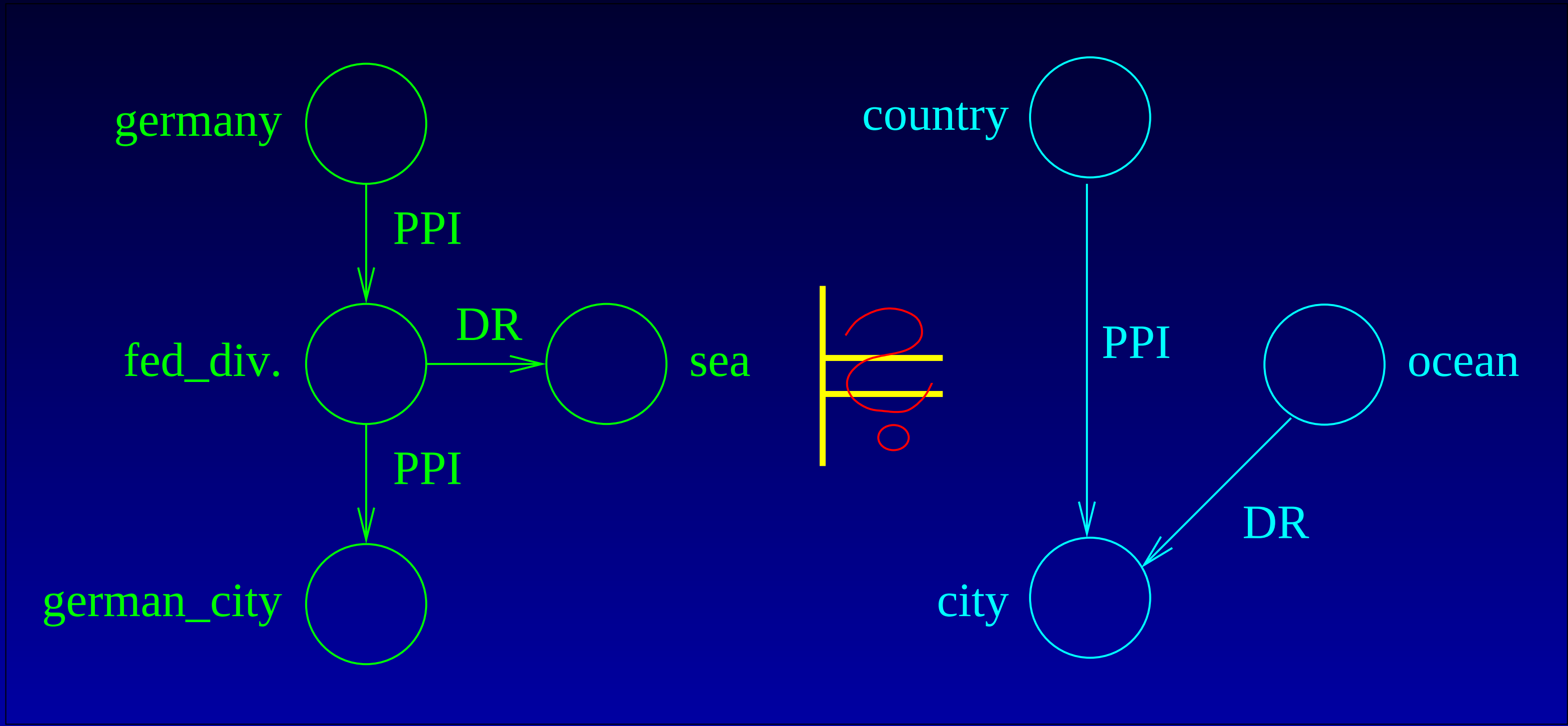
- Example: query consistency

$\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks

- Example: hybrid query containment

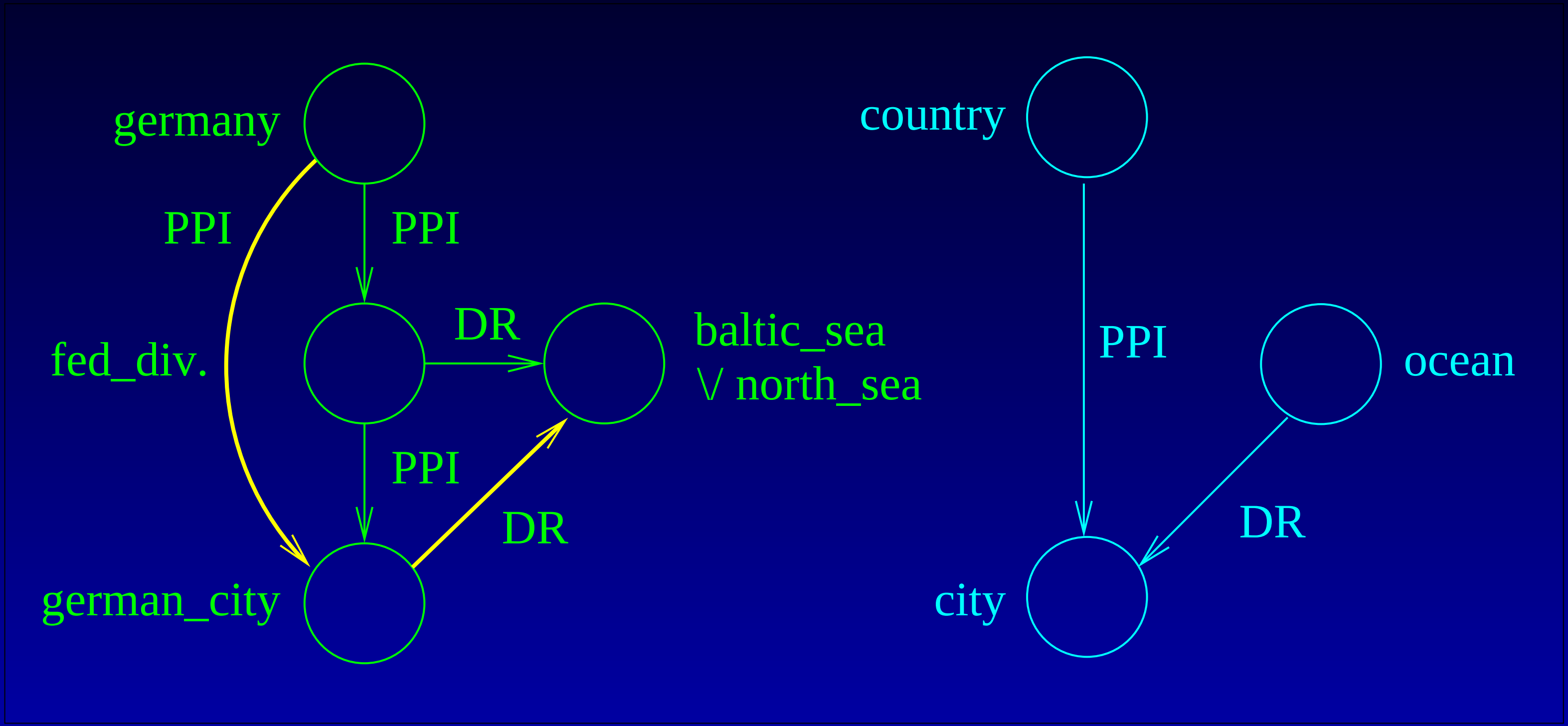
$$query(?germany, ?city, ?sea) \leftarrow$$
$$germany(?germany^*), federal\_division(?division^*),$$
$$german\_city(?city^*), (baltic\_see \sqcup north\_sea)(?sea^*),$$
$$\boxed{\text{PPI(?germany,?division), PPI(?division,?city),}}$$
$$\boxed{\text{DR(?division,?sea)}}$$
$$\models$$
$$query(?country, ?city, ?ocean) \leftarrow$$
$$country(?country^*), city(?city^*), ocean(?ocean^*),$$
$$\boxed{\text{DR(?ocean,?city), PPI(?country,?city)}}$$

Reasoning about Queries



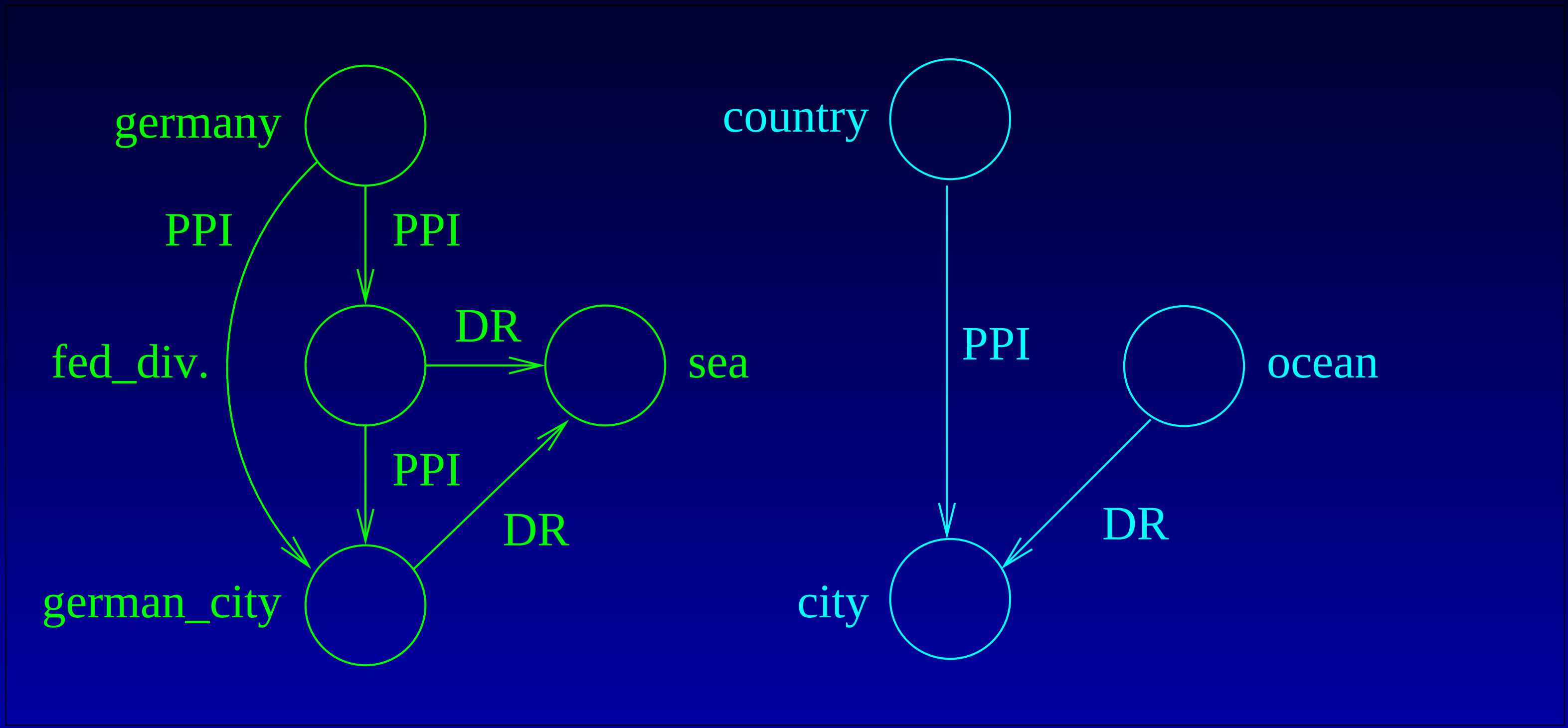
Two queries - does Green entail Blue?

Reasoning about Queries

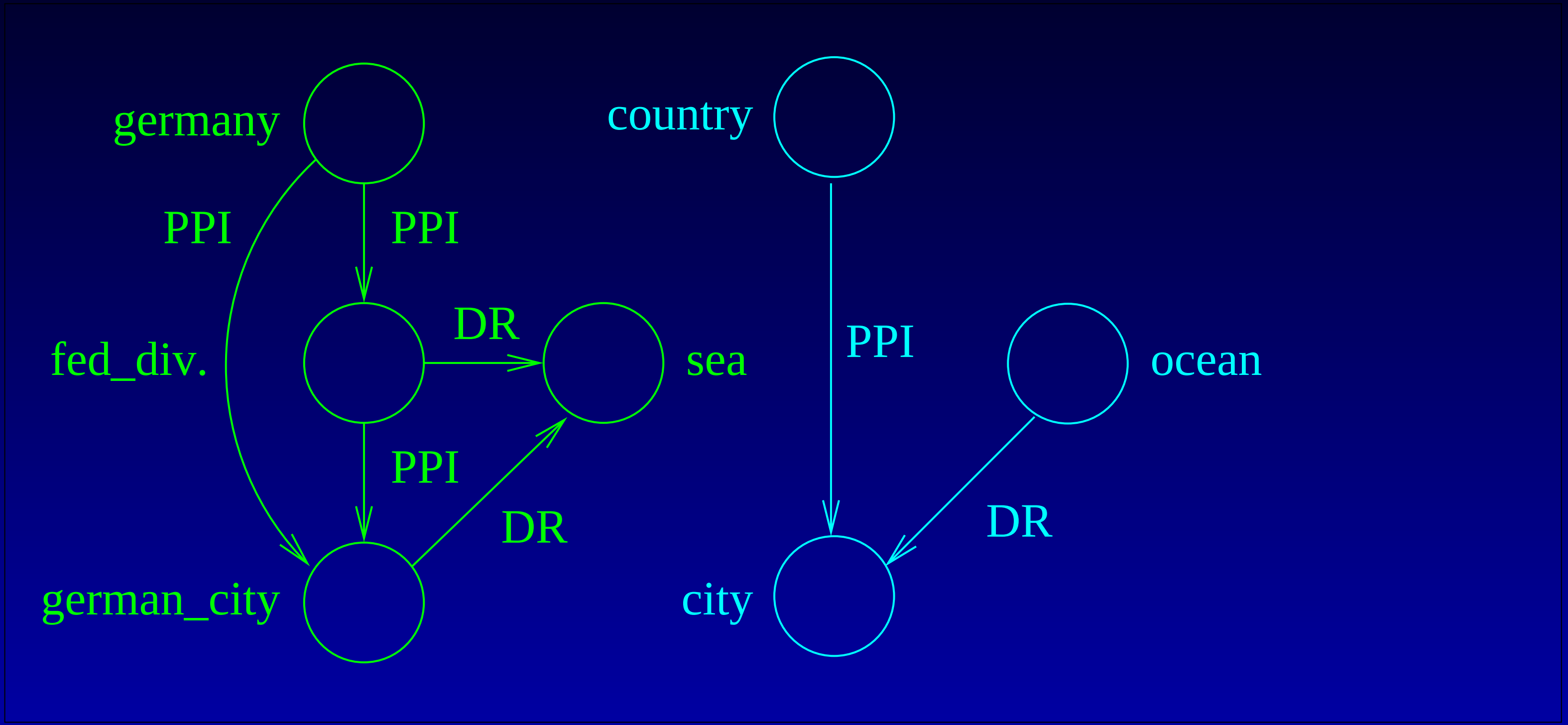


Adding entailed constraints for Green

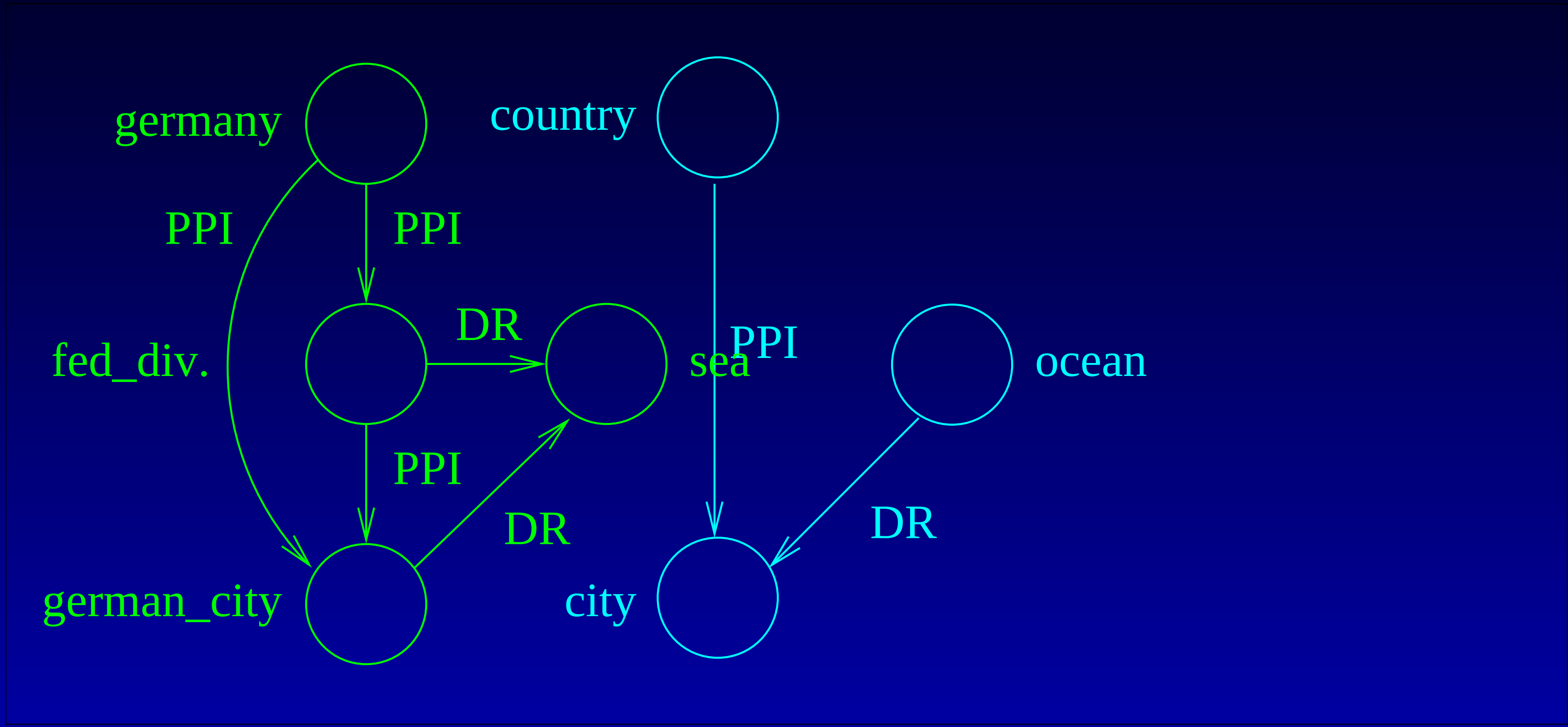
Reasoning about Queries



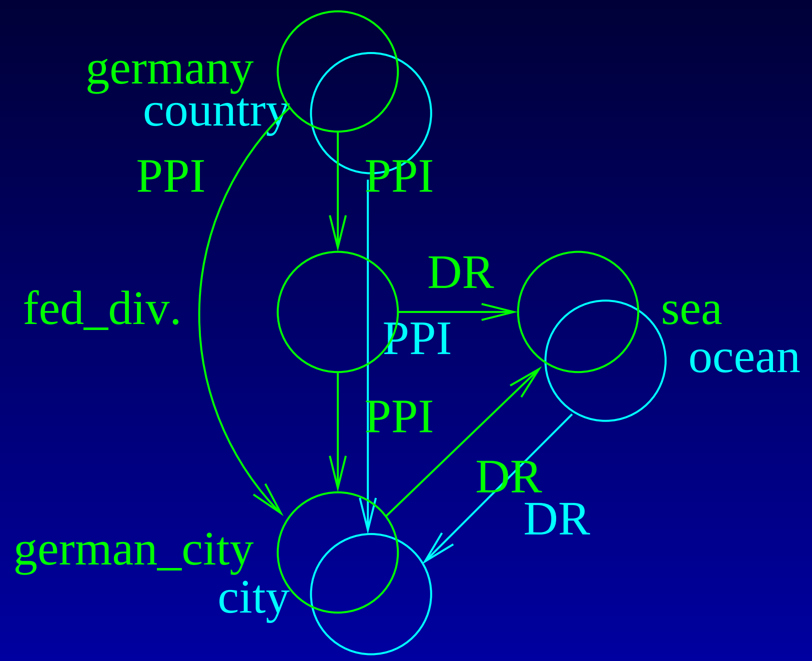
Reasoning about Queries



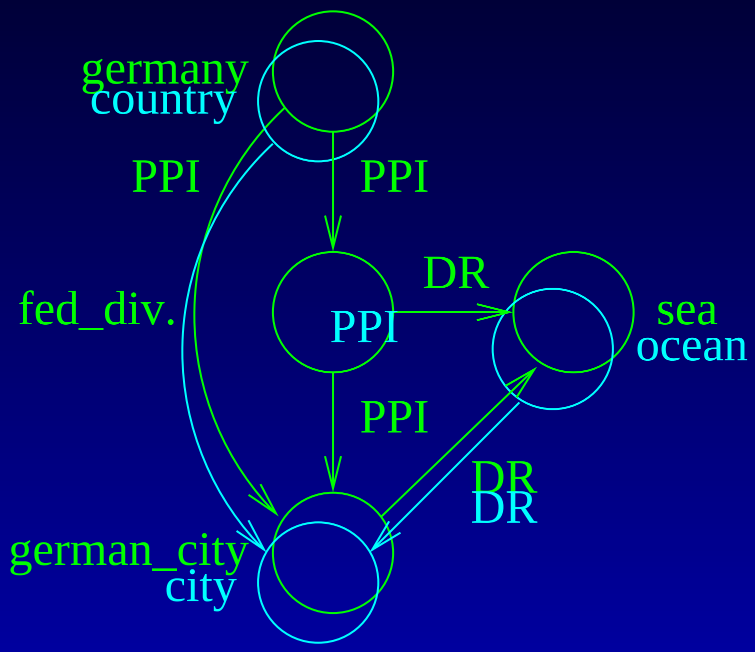
Reasoning about Queries



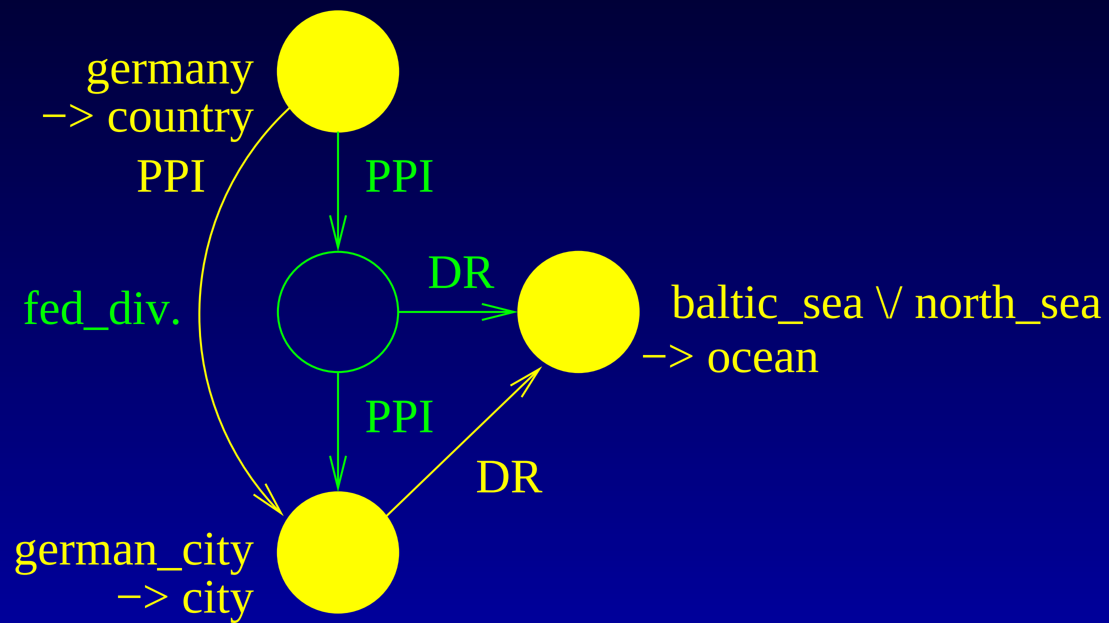
Reasoning about Queries



Reasoning about Queries



Reasoning about Queries



Match - Green is more specific than Blue

Reasoning about Queries

- Example: query consistency
  - $\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks
- Example: hybrid query containment
  - $\Rightarrow$ By reduction to query consistency

Reasoning about Queries

- Example: query consistency

  $\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks

- Example: hybrid query containment

  $\Rightarrow$ By reduction to query consistency

- Customizable: notion of consistency has to be provided by the framework user (implementation of specialized methods)

Reasoning about Queries

- Example: query consistency

  $\Rightarrow$ Reduction to appropriate ABox / RCC consistency checks

- Example: hybrid query containment

  $\Rightarrow$ By reduction to query consistency

- Customizable: notion of consistency has to be provided by the framework user (implementation of specialized methods)
- Vision: since queries can also be seen as "concept definitions" it might be reasonable to base the ontology $\mathcal{I}$ on them (instead of a truly spatio-thematic description logic)

Summary

- Usage of (one of) the fastest DL systems does not guarantee good overall performance in arbitrary application contexts

Summary

- Usage of (one of) the fastest DL systems does not guarantee good overall performance in arbitrary application contexts
- Application-specific reasoners and/or reasoning services are still needed

Summary

- Usage of (one of) the fastest DL systems does not guarantee good overall performance in arbitrary application contexts
- Application-specific reasoners and/or reasoning services are still needed
- Application-specific index structures and optimizations are needed

Summary

- Usage of (one of) the fastest DL systems does not guarantee good overall performance in arbitrary application contexts
- Application-specific reasoners and/or reasoning services are still needed
- Application-specific index structures and optimizations are needed

⇒ It would be nice if DL systems were more open and "customizable" using inheritance (where is the DL system with arbitrary user-definable concrete domains?)

Summary

- Usage of (one of) the fastest DL systems does not guarantee good overall performance in arbitrary application contexts
- Application-specific reasoners and/or reasoning services are still needed
- Application-specific index structures and optimizations are needed

$\Rightarrow$ It would be nice if DL systems were more open and "customizable" using inheritance (where is the DL system with arbitrary user-definable concrete domains?)

$\Rightarrow$ An object-oriented DL-system architecture can have advantages

Summary

Thanks for your attention!