12. Finding Components with Metadata in Component Repositories

Lecturer: Dr. Sebastian Götz

Prof. Dr. Uwe Aßmann
Technische Universität Dresden
Institut für Software- und Multimediatechnik
http://st.inf.tu-dresden.de/teaching/cbse
13.04.2017

1. Component Search with Metadata
2. Searching and Browsing with Faceted Classications
3. Faceted Metadata
4. UML Components
5. Searching by Conformance to Protocols

Obligatory Literature

- R. Prieto-Diaz. Implementing Faceted Classification for Software Reuse. CACM May 1991, vol 34(5).
- U. Aßmann. Reuse in Semantic Applications. REWERSE summer school 2005, La Valetta, Malta. Lecture Notes In Computer Science (LNCS) 3564.
  - http://www.springerlink.com/content/blx9yfthkq5xjtjg/

References

- http://flamenco.berkeley.edu
- http://search.express.ebay.com
- FacetMap: Greg Smith, Mary Czerwinski, Brian Meyers, Daniel Robbins, George Robertson, Desney S. Tan. FacetMap: A Scalable Search and Browse Visualization. IEEE Transactions on visualization and computer graphics, vol.12 , No. 5, september/october 2006.
- Thorsten Teschke. Semantische Komponentensuche auf Basis von Geschäftsprozessmodellen. Dissertation. Universität Oldenburg, 2003.
- Facet-based search of computer science literature in DBLP repository
  - http://dblp.l3s.de/
- Luca de Alfaro and Thomas A. Henzinger: Interface automata. ACM SIGSOFT FSE/ESEC, 2001
  - http://doi.acm.org/10.1145/503209.503226

12.1. Component Search in Component Repositories

- It should be as easy to find good quality reusable software assets as it is to find a book on the internet

Component Repositories

- Components must be stored in component repositories with **metadata (markup, attributes)** to find them again
- Description by Metadata:
  - Attributes: Keywords, Author data
  - Contracts (Usage protocols, behavioral specifications)
    - State machines
    - Sequence diagrams
    - Contracts (pre/post/invariants)
- Examples of Component Repositories
  - CORBA
    - implementation registry
    - interface registry
  - COM+ registry
  - Commercial Component Stores www.componentsource.com
  - Debian Linux Component System (apt, dpkg)
  - CTAN TeX Archive

Why Searching Components?

- A public component repository is called a **market**, managed by a **trader (broker)**
  - Distributing or selling components
  - Companies can register components at the the trader
  - Customers can search components in the markets and buy or rent them
- Searching for functionality (interface, contract, protocol)
  - Reuse instead of build
  - Searching for components to replace own ones
  - Semantic substituability (CM-S) should be ensured
- Searching for quality features
  - Performance, energy consumption, reliability

12.2 Searching and Browsing with Faceted Classifications

(thanks to Jan Polowinski)

Faceted Classification for Better Matchmaking

- A **facet** is a dimension of a classification
  - Facets simplify search: Facet classification has been invented in library science to simplify the description and search for books [Ranganathan].
  - A component (or service) is described in several facets, dimensions, which are orthogonal to each other
- Matchmaking engines can look up a service by stating the desired properties for all facets.
- Classifications can be arranged in facets if several partitions of a group of objects exist that are orthogonal
  - In domain modelling, this is often the case
  - Without facets, multiple inheritance hierarchies have to be specified, which are often clumsy and error-prone
- Idea: use facets for better matchmaking

Comparison

Standard Classification

- **B Birds**
  - B1 Breathing of Birds
  - B2 Breading of Birds
- **F Fish**
  - F1 Breathing of Fish
  - F2 Breading of Fish
- **M Mammal**
  - M1 Breathing of Mammals
  - M2 Breading of Mammals
- **I Insects**
  - I1 Breathing of Insects
  - I2 Breading of Insects

- **Gills: F1**

Faceted Classification

- **Processfacet**
  - P Physiology
    - P1 Breathing
    - P2 Breading
- **Animalfacet**
  - 1 Birds
  - 2 Fish
  - 3 Mammals
  - 4 Insects

- **Gills: P1-2**

Facetted Browsing

- Here Facet means: an interesting property of an object orthogonal to other properties
- Incremental refinement of a set of results by restricting values of the data's facets
- Many application domains

TOPHER'S BREAKFAST CEREAL CHARACTER GUIDE

Here is the Exhibit JSON data file.

Search

Brands 1 ☑

- 3 General Mills ☐
- 9 **Kellogg's** ☑
- 2 Nabisco ☐
- 4 Nestle ☐
- 9 Post ☐
- 12 Quaker Oats ☐
- 10 Ralston ☐

Decades

- 1 1930
- 2 1950
- 1 1980
- 1 1990
- 1 2000
- 3 unknown

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ☑ Gruppierung wie Sortierung

Kellogg's (9)

Snap!


Tony the Tiger


Cornelius

Baby Toucan


Mr. Mini-Wheats

C-Rex


Smacks

Broncos Kid

Countries

- 1 Canada
- 1 France
- 1 Great Britain
- 1 Japan
- 5 USA

Forms

- 3 boy
- 1 dinosaur
- 1 frog
- 1 n/a
- 1 rooster
- 1 tiger
- 1 toucan

TOPHER'S BREAKFAST CEREAL CHARACTER GUIDE

Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

Facet

Facet

Facet

Facet

Facet

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetz

brand; sowie nach... • ☑ Gruppierung wie Sortierung

Brands

- 3 General Mills
- 9 Kellogg's ☑
- 2 Nabisco ☐
- 4 Nestle ☐
- 9 Post ☐
- 12 Quaker Oats ☐
- 10 Ralston ☐

Decades

- 1 1930
- 2 1950
- 1 1980
- 1 1990
- 1 2000
- 3 unknown

Kellogg's (9)

Snap!


Tony the Tiger


Cornelius


Mr. Mini-Wheats

C-Rex


Smacks

Broncos Kid

Countries

- 1 Canada
- 1 France
- 1 Great Britain
- 1 Japan
- 5 USA

Forms

- 3 boy
- 1 dinosaur
- 1 frog
- 1 n/a
- 1 rooster
- 1 tiger
- 1 toucan

TOPHER'S BREAKFAST CEREAL CHARACTER GUIDE

Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ☑ Gruppierung wie Sortierung

Brands 1 ☑

- 3 General Mills ☐
- 9 Kellogg's ☑
- 2 Nabisco ☐
- 4 Nestle ☐
- 9 Post ☐
- 12 Quaker Oats ☐
- 10 Ralston ☐

Widget for Restriction of Facet Values

Decades

- 1 1930
- 2 1950
- 1 1980
- 1 1990
- 1 2000
- 3 unknown

Kellogg's (9


Snap!


Tony the Tiger


Cornelius

Baby Toucan


Mr. Mini-Wheats

C-Rex


Smacks

Broncos Kid

Countries

- 1 Canada
- 1 France
- 1 Great Britain
- 1 Japan
- 5 USA

Forms

- 3 boy
- 1 dinosaur
- 1 frog
- 1 n/a
- 1 rooster
- 1 tiger
- 1 toucan

Sorting and Grouping Mechanisms

Search

VORSCHAUBILDER • ZEITLEISTE

haracters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ✓ Gruppierung wie Sortierung

Countries

1 Canada
1 France
1 Great Britain
1 Japan
5 USA

Forms

3 boy
1 dinosaur
1 frog
1 n/a
1 rooster
1 tiger
1 toucan

Decades

1 1930
2 1950
1 1980
1 1990
1 2000
3 unknown

gg's (9)

2 Nabisco
4 Nestle
9 Post
12 Quaker Oats
10 Ralston

Snap!

Tony the Tiger

Cornelius

Baby Toucan

Mr. Mini-Wheats

C-Rex

Smacks

Broncos Kid

TOPHER'S BREAKFAST CEREAL CHARACTER GUIDE

Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

Brands 1

- 3 General Mills
- 9 Kellogg's
- 2 Nabisco
- 4 Nestle
- 9 Post
- 12 Quaker Oats
- 10 Ralston

Decade

- 1 19
- 2 19
- 1 19
- 1 1990
- 1 2000
- 3 unknown

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • Gruppierung wie Sortierung

Kellogg's (9)

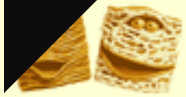
Snap!


Tony the Tiger


Cornelius

Baby Toucan

Result Set


Mr. Mini-Wheats

C-Rex


Smacks


Broncos Kid

Countries

- 1 Canada
- 1 France
- 1 Great Britain
- 1 Japan
- 5 USA

Forms

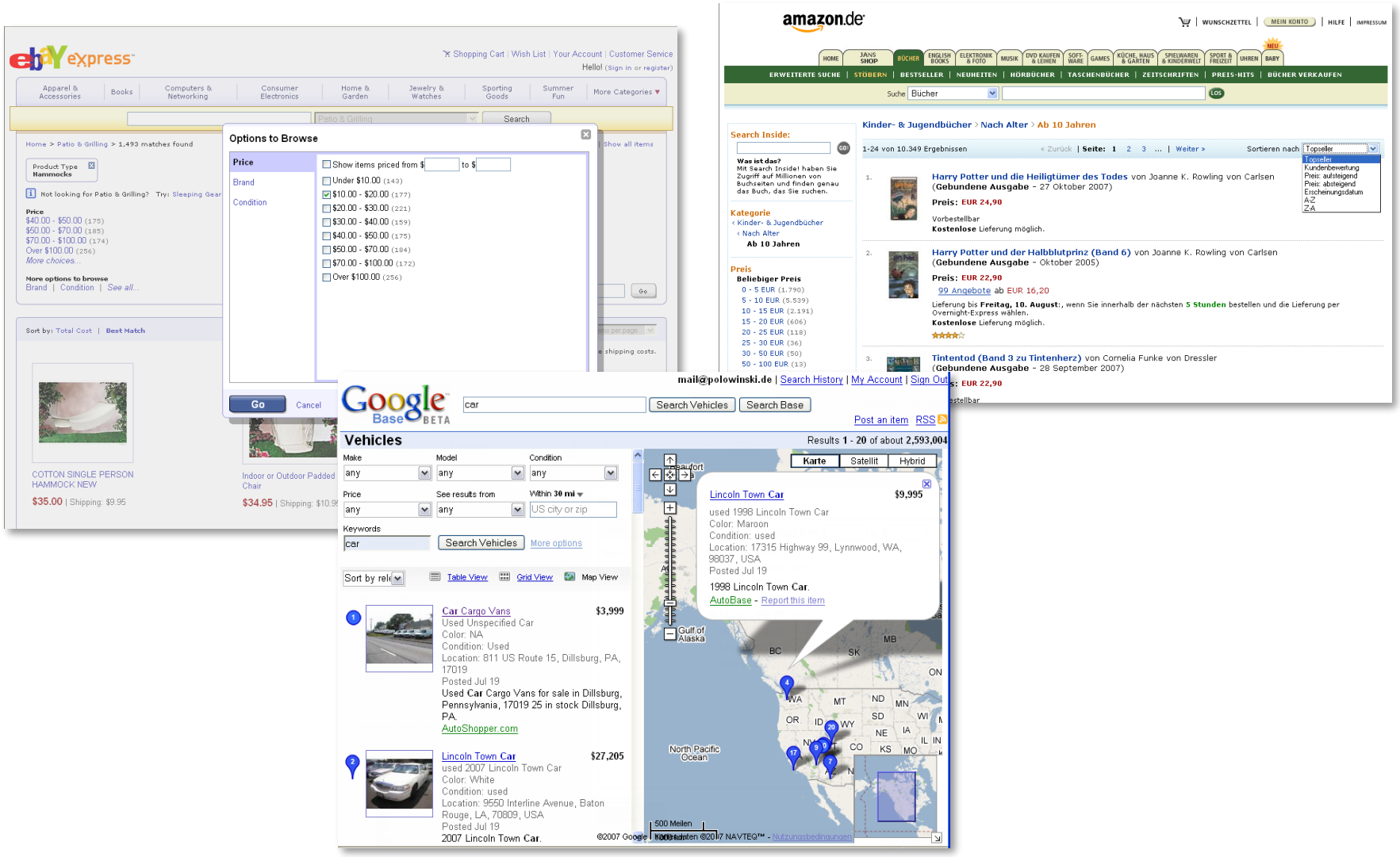
- 3 boy
- 1 dinosaur
- 1 frog
- 1 n/a
- 1 rooster
- 1 tiger
- 1 toucan

More Examples of Facetted Browsers

- Flamenco
  - FLexible information Access using MEtadata in Novel COmbinations
  - University of California, Berkeley

- mSpace
  - http://mspace.fm
  - *University of Southampton*
- FacetMap
  - Microsoft Research

Facetted Browsing in e-Commerce



Google Suche PageRank Rechtschreibprüfung Optionen

mSpace BETA

mSpace Classical Music Explorer

Browser

Search keywords...

Columns: Album | Arrangement | Recording | Artist | Composer Death Date | Composer Birth Date | Form

INTEREST

Era

- Renaissance
- Baroque
- Classical
- Romantic
- Modern
- Contemporary

Composer

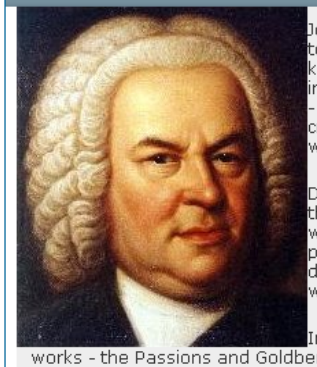
- Bach, Johann Sebastian (1685-1750)
- Biber, Heinrich Ignaz Franz von (1644-1704)
- de Araujo, Pedro (1662 - 1705)
- Gibbons, Orlando (1583 - 1625)
- Maier, Michael (1568 - 1622)
- Pachelbel, Johann (1653 - 1706)
- Purcell, Henry (1659 - 1695)
- Reusner, Esaias Jr. (1636-1679)
- Sweelinck, Jan Pieterszoon. (1562 - 1621)
- Vivaldi, Antonio (1678 - 1741)

Piece

- Canones diversi super thema regium
- Concerto No. 1 in F major, BMV 1046, III. Allegro
- Concerto No. 1 in F major, BMV 1046, I. [-]
- Concerto No. 1 in F major, BMV 1046, II. Adagio
- Concerto No. 1 in F major, BMV 1046, IV. Menuetto -Trio - Polonaise
- Concerto No. 2 in F major, BMV 1047, III. Allegro assai
- Concerto No. 2 in F major, BMV 1047, I [-]
- Concerto No. 2 in F major, BMV 1047, II. Andante
- Concerto No. 3 in G major BMV 1048 I. [-]

page 1 of 11

Information



Johann Sebastian Bach was music's most sublime creative genius. With the notable exception of opera, he composed towering masterpieces in every major Baroque genre: sonatas, concertos, suites and cantatas, as well as innumerable keyboard, organ and choral works. Yet despite the sheer vastness of his output, Bach sustained a rarefied level of musical inspiration that continues to amaze his most gifted successors. Even those who find Bach's music somewhat overwrought - or just plain dull - concede that at a purely technical level he is in a league of his own. Yet during his lifetime his awesome creative talent went largely unrecognised and by the time of his death he was viewed as something of a musical dinosaur who had stubbornly refused to move with the times.

Difficult though it is to believe, in Bach's day musical works were as ephemeral s today's pop songs. Here one day, gone the next. As the tide of taste and fashion rolled inexorably by, so new pieces were required on almost a daily basis for a wide variety of purposes. The notion of rows of music-lovers gathering to listen in hallowed silence to music composed the previous year, let alone centuries before, was practically unheard of. The concert-going experience that Pierre Boulez despairingly referred to as "museum culture" was a concept that lay 100 years ahead. In his lifetime, Bach's reputation was first and foremost as Germany's leading organist - not as a composer.

In his later years Bach faced harsh criticism. During the 1720s and 1730s when he was composing his most important works - the Passions and Goldberg Variations among them - a new Italian style invaded Germany, making his work appear outdated. The composer-critic Adolph Scheibe (1708-1776) attacked Bach's style as "turgid and confused".

It is a common misconception, however, that following his death, Bach's name fell into total obscurity. In fact, his son Carl Philipp Emanuel did much to ensure that the details of his father's life and achievements were recorded for posterity. He published an important biographical obituary in 1754 and, in

Preview

Concerto for Trumpet and Orchestra in D major, I. Adagio

Concerto for Trumpet and Orchestra in D major, II. Allegro

mPlayer ynton Marsalis - Concerto for Trumpet and C

mSpace Classical Music Explorer. Part of the mSpace network.

Fertig

Nobel Prize Winners

1901 to 2004

Powered by Flamenco

Save Search | History and Settings | Return to Search | New Search | Logout

search

all items in current results

Refine your search within these categories:

GENDER (group results)

male (44)

COUNTRY: all > Germany

AFFILIATION (group results)

Berlin University (1)
Germany (38)
Locarno Pact (1)

PRIZE (group results)

chemistry (17)
literature (5)
medicine (8)
peace (3)
physics (11)

YEAR (group results)

1900s (12)
1910s (10)
1920s (11)
1930s (10)
1940s (1)

Recently Viewed Items

Go to Item History

These terms define your current search. Click the × to remove a term.

COUNTRY: Germany ×

Items 1 to 40 of 44 results

Group by: country

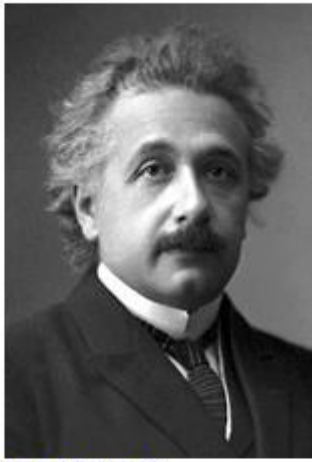
Sort by: usual name, year of birth, year of death, country

1 41

Adolf Butenandt
1903-1995

Adolf von Baeyer
1835-1917

Adolf Windaus
1876-1959



Albert Einstein
1879-1955

12.3 Faceted Metadata for Search in Component Repositories

Example: Service Facets in a UNIX System

- To describe the services of a UNIX system, [Prieto-Diaz] employed a 4-faceted scheme
  - function
  - logical object
  - implementation object
  - tool
- UNIX services can be described with appropriate facet values and looked up in a repository
- Example: "append a line to a file with a text editor"
  - (function = append, logical class = line, implementation class = file, tool = text editor):

	Function	Logical class	Implementation Class	Tool
	edit	Line	File	Text editor

Example: Services in a UNIX System

- [Prieto-Diaz] already suggested to use *controlled vocabulary* (*domain ontologies*) to improve the effectiveness of the search:
  - If every facet is described by an ontology, the service descriptions are standardized for a user group and improve understanding of service semantics.
- Facets simplified the description of the components, improved the understanding of their domain, and facilitated the search in component libraries.

COMPONENTS FACETED

Name
- 1 ColorChooser
- 1 ColorSelector
- 1 ColorUtils

Information Hiding
- 2 BlackBox
- 1 GreyBox
- 1 WhiteBox

Purpose of the Component
- 1 (Feld fehlt)
- 2 Editing
- 1 Managing

belongs to Layer
- 1 (Feld fehlt)
- 1 CORE
- 2 GUI

Language
- 1 (Feld fehlt)
- 1 C#
- 1 C++

5 Components

sortiert nach: Name und Version; sowie nach... • ☑ Gruppierung wie Sortierung

1. **ColorChooser (release, Versions: 1.1)**
   Last Update on Mo, Jan 1, 2007, 02:00 am (53 days ago). Author: Schmidt
   - Information Hiding: BlackBox
   - Purpose: Editing
   - Layer: GUI
   - License: Free
   - LOC: 2500
   - Language: Java

ColorSelector (, Versions: 1.0 und 1.1)
Last Update on Di, Jan 2, 2007, 02:00 am und Mi, Jan 2, 2008, 02:00 am (days ago). Author: Polowinski
- Information Hiding: BlackBox
- Purpose: Editing
- Layer: GUI
- License:

License
- 2 (Feld fehlt)
- 1 Free
- 1 GNU-GPL

Price
- 2 (Feld fehlt)
- 1 200
- 1 250

Maturity
- 2 (Feld fehlt)
- 1 alpha
- 1 beta

Version
- 1 (Feld fehlt)
- 1 1.0
- 2 1.1

Last Edited
- 1 (Feld fehlt)
- 1 2001-06-03T00:00:00+00:00
- 1 2007-01-01T00:00:00+00:00

COMPONENTS FACETED

Name
1 PersistenceComponent

Information Hiding 1
2 BlackBox
1 GreyBox
1 WhiteBox

Purpose of the Component
1 Persistence

belongs to Layer
1 PersistenceLayer

Language
1 C++

1 Component gefiltert von ursprünglich 5 (Alle Filter zurücksetzen)

sortiert nach: Name und Version; sowie nach... • Gruppierung wie Sortierung

PersistenceComponent (alpha, Versions: 1.6)

Last Update on So, Jun 3, 2001, 02:00 am (12 days ago). Author: Müller

- Information Hiding: GreyBox
- Purpose: Persistence
- Layer: PersistenceLayer
- License: GNU-GPL
- LOC: 155455
- Language: C++

1. ***Buy for 3000 €***

License
1 GNU-GPL

Price
1 3000

Maturity
1 alpha

Version
1 1.6

Last Edited
1 2001-06-03T00:00:00+00:00

Other Advantages

- The facet classification is rather immune to extensions
  - Extending one facet leaves all others invariant
  - Example: If Europe is extended with a new member state, the matchmaking algorithm can deliver new courses from the new member state, without affecting the rest of the semantic specifications at all
- The accuracy can be improved by synonym lists (thesauri)
  - Synonyms increase the chances for a match
  - They permit to search not only for keywords, but also for their *synonyms* (assembled in a *thesaurus*)
  - Beyond synonyms other refinement relations of concepts can be used to improve the search
  - Example: Great Britain is used as a synonym for England, Scotland, and Wales. Synonyms allows for matchmaking on any of the keywords, so that students looking for a course need not bother about geographic and political details.

The Use of Ontologies in Faceted Matchmaking

- Ontologies simplify matchmaking by standardization
  - Since they provide standardized terminology and standardized ontological relations between the terms, queries can specify
    - keywords with a precise, shared, and standardized meaning (semantic search),
    - contextual information for search in context, where the context is defined by the ontological relations of the terms.
- Example:
  - A web course on IT basics can be queried by the standardized word IT-basics (being semantic search)
  - also in context, by relating it to courses such as IT-advanced or IT-preparatory (contextual search)
    - "find me an IT basics course, which has a preceding preparatory IT course and has a follow-up advanced IT course"

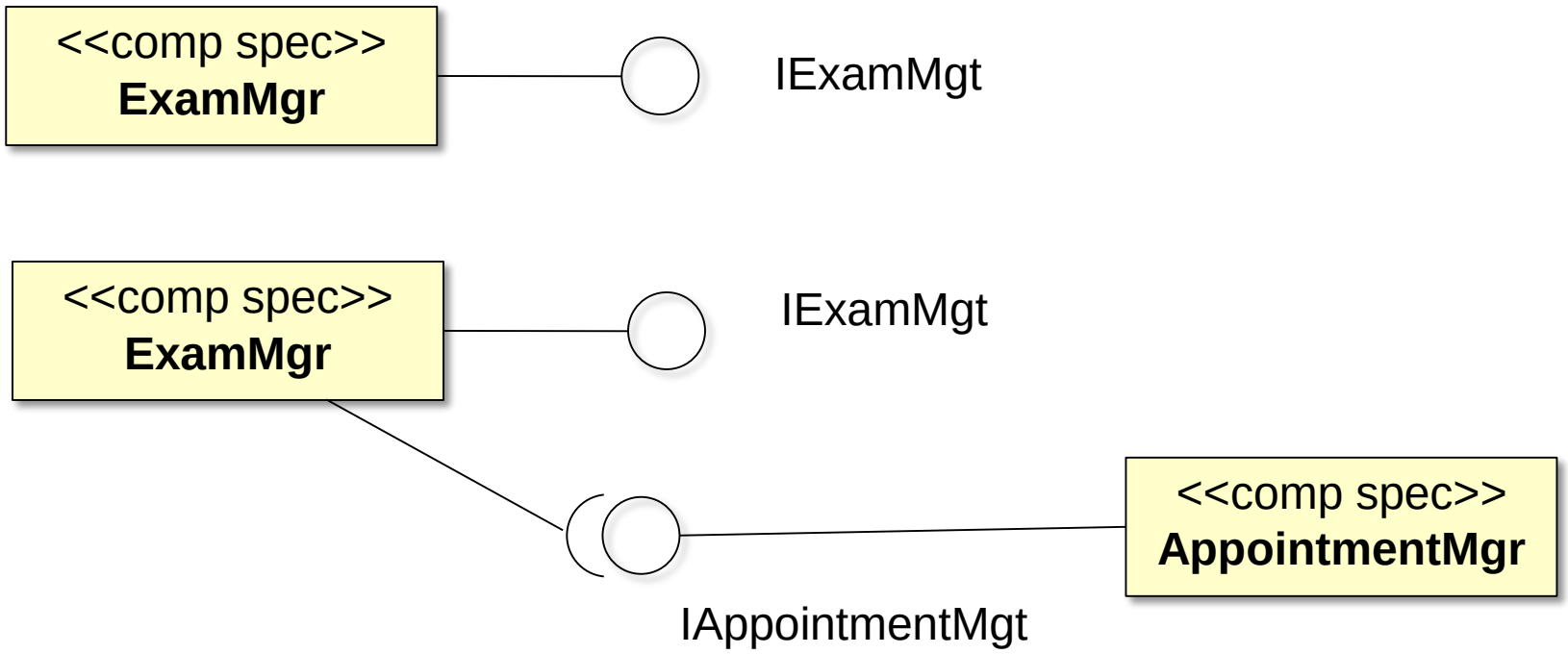
Putting up a Component Repository for Your Company

- Define facets for component metadata
  - If possible, reuse an ontology for a facet
  - Form a thesaurus for synonyms
  - Store the metadata as a tuple in the database
- Realize a search algorithm that uses facets together with thesauri
- Or use a faceted browser with the metadata

12.4 UML Components

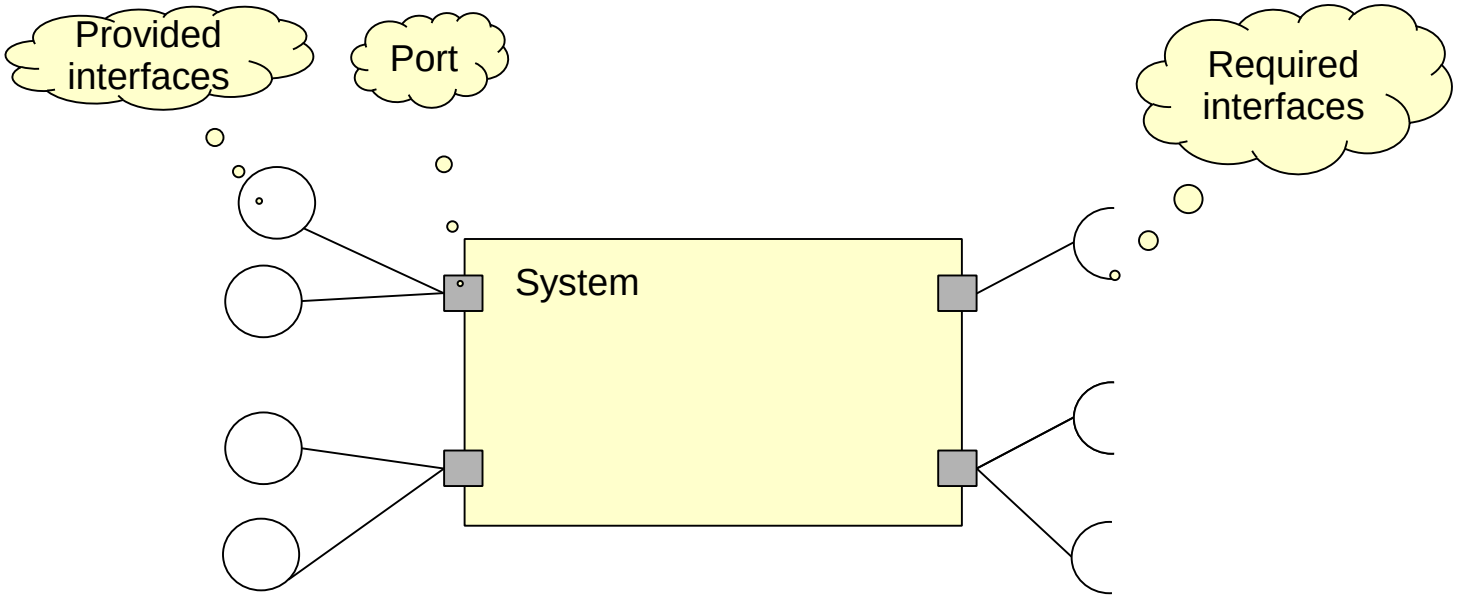
Component Specification with UML Components

- A **UML component** is a hierarchical class for big objects with *provided* and *required* interfaces (roles)
  - Provided interfaces (provided roles) use „lollipop" notation
  - Required interfaces (required roles) use „plug" notation
- Some components are required to use specific other interfaces



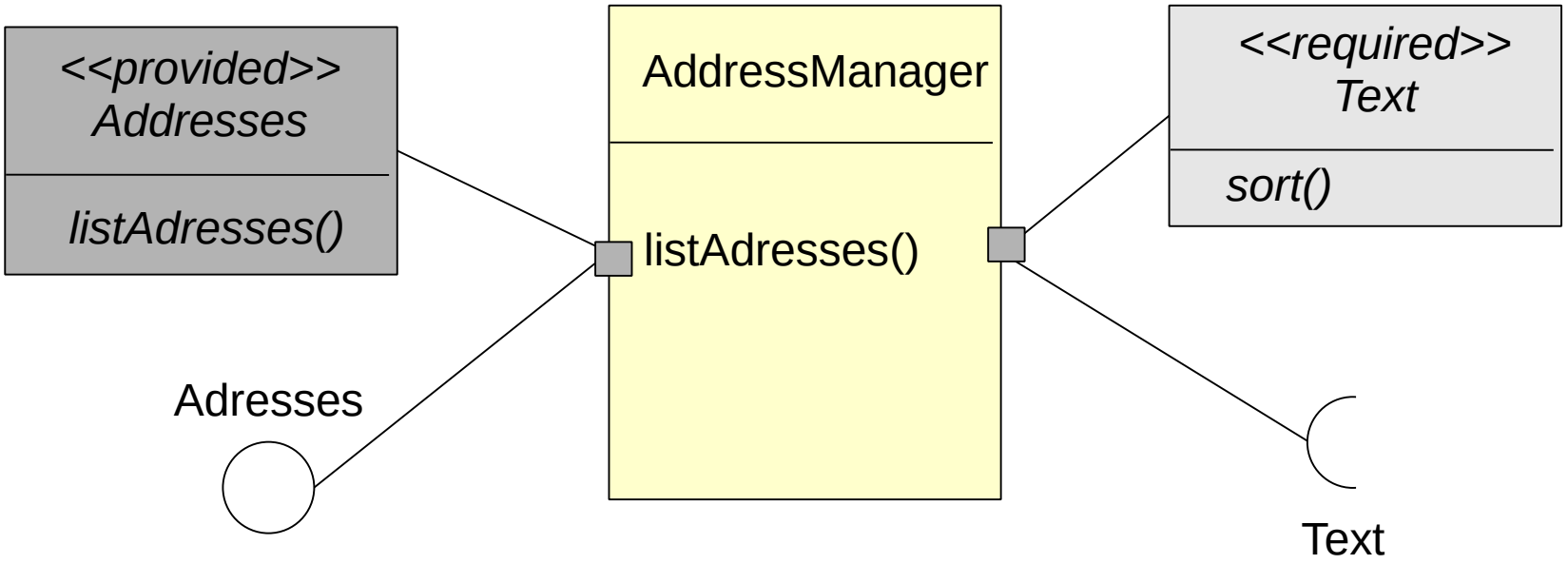
Ports of UML Components

- A **port** is a connection point of a UML component.
  - A port has a set of roles (interfaces)
  - It may be represented by a **port object (gate)**



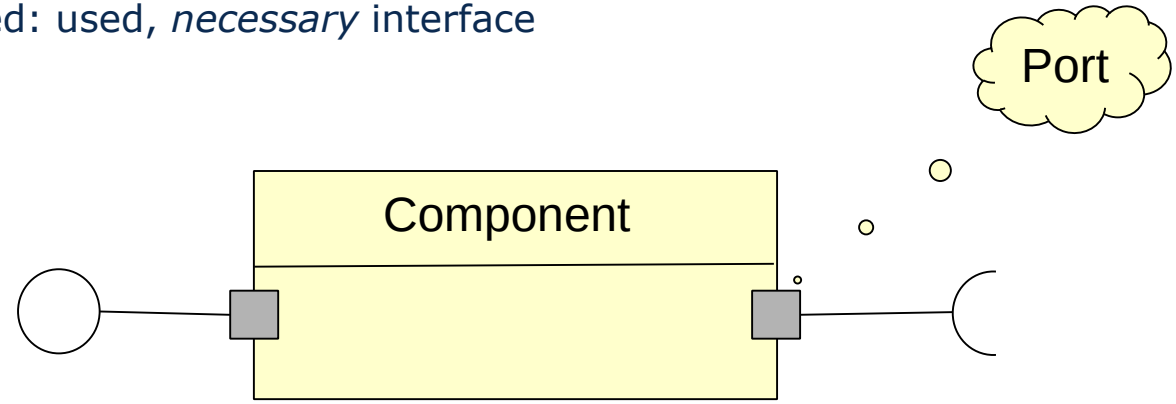
Lollipops und Plugs (Balls and Sockets)

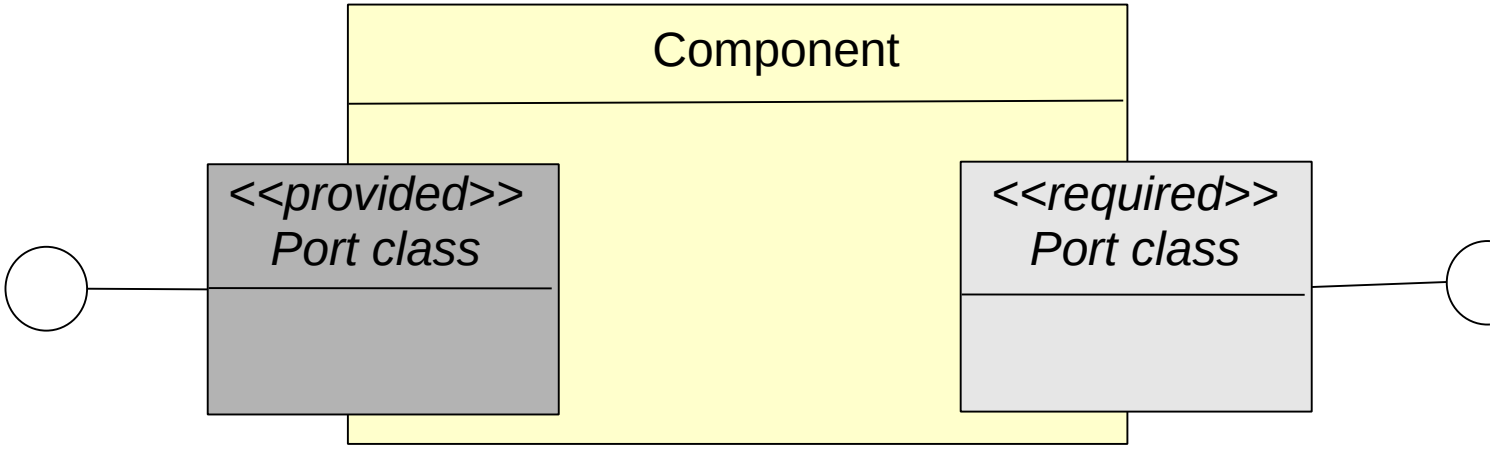
- For a UML component, *provided* and *required interfaces* can be distinguished
  - A required interface specifies what the current class needs to execute.



Ports

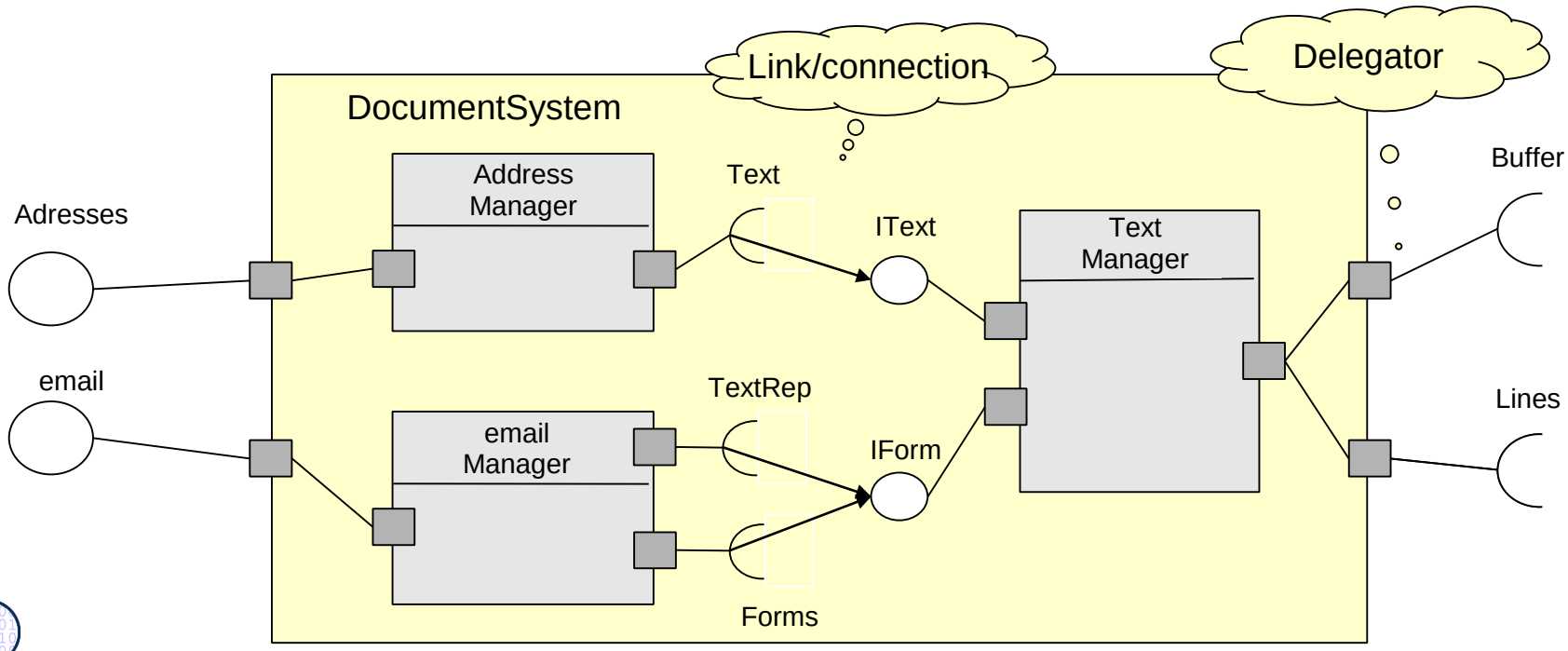
- Ports consist of **port classes** with interfaces and behavior in form of **interface automata**
  - provided: normal, *offered* interface
  - required: used, *necessary* interface





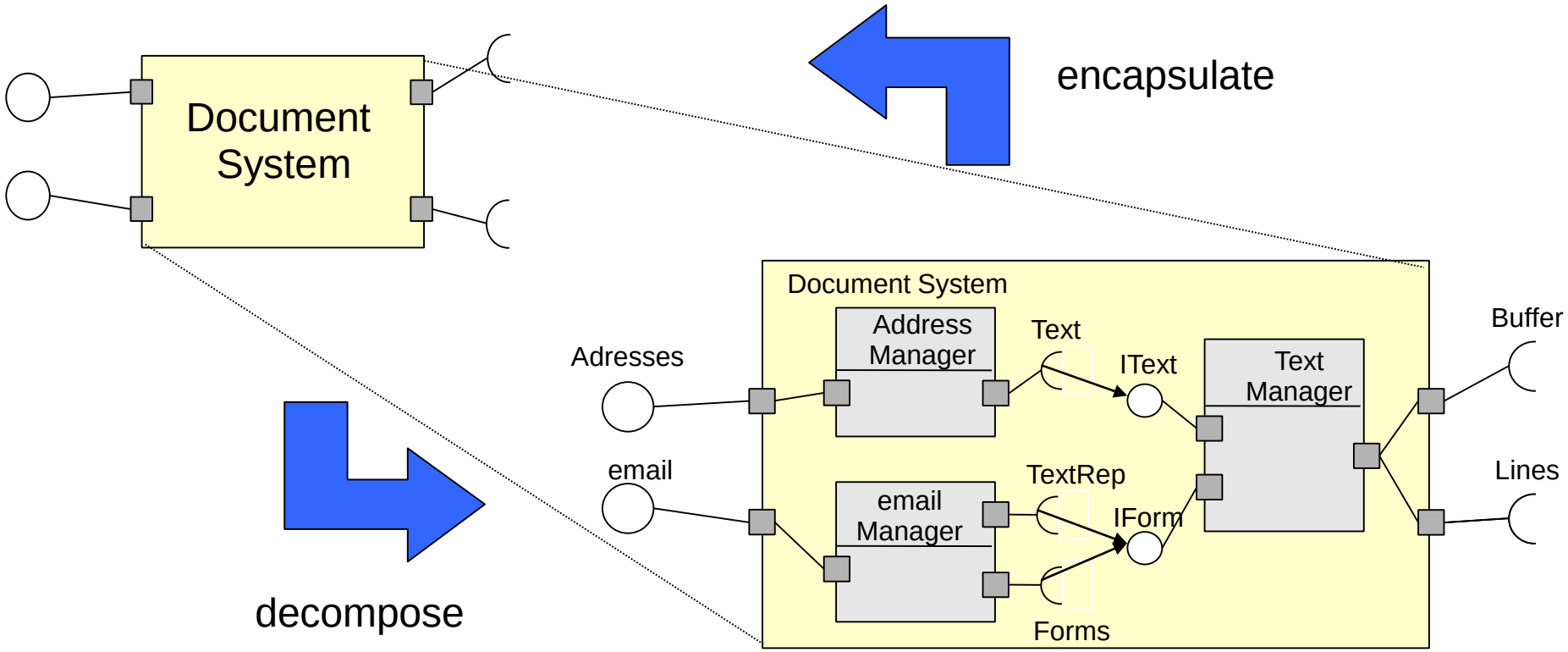
Nesting of UML Components

- UML components
  - Ports are connected by *links (connections)*
  - *Delegation link*: links outer and inner port



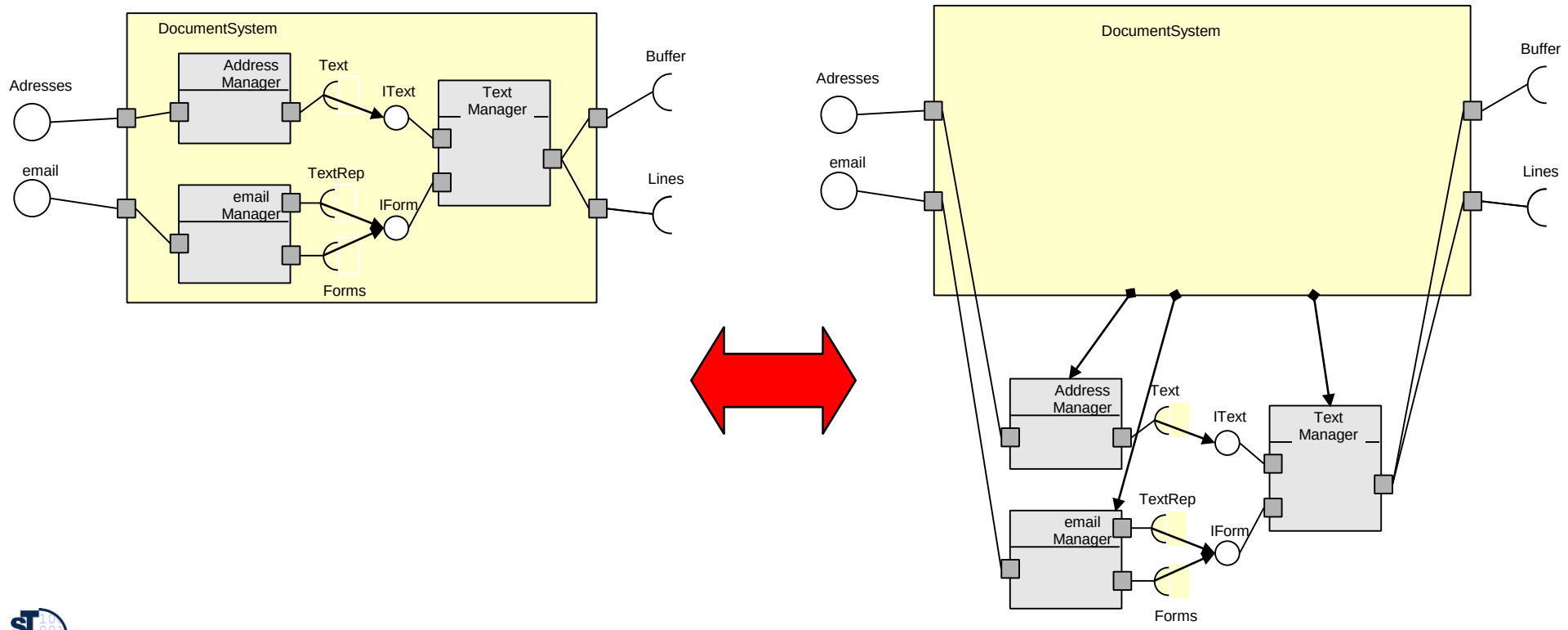
Refinement of UML Components

- UML components can be nested.
- Nesting is indicated by *aggregation* and *part-of relationship*.
- Nesting is introduced by an encapsulation operator encapsulate.



Encapsulation means Aggregation

- *Nesting* means *Aggregation*
  - A UML component is a package and a facade for all subcomponents

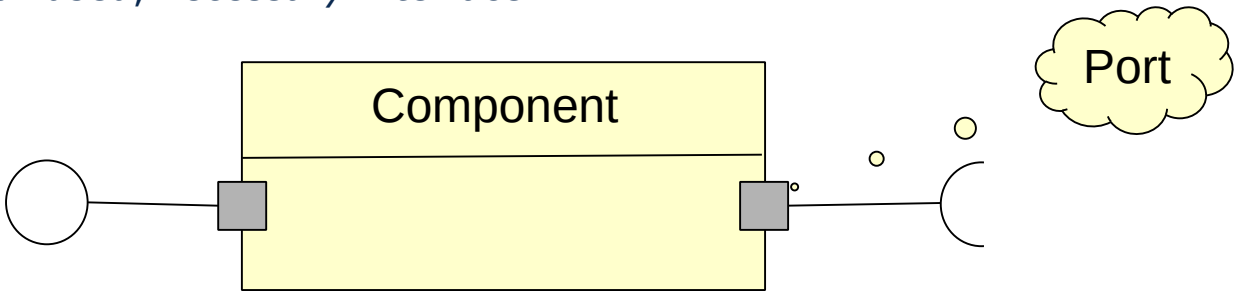


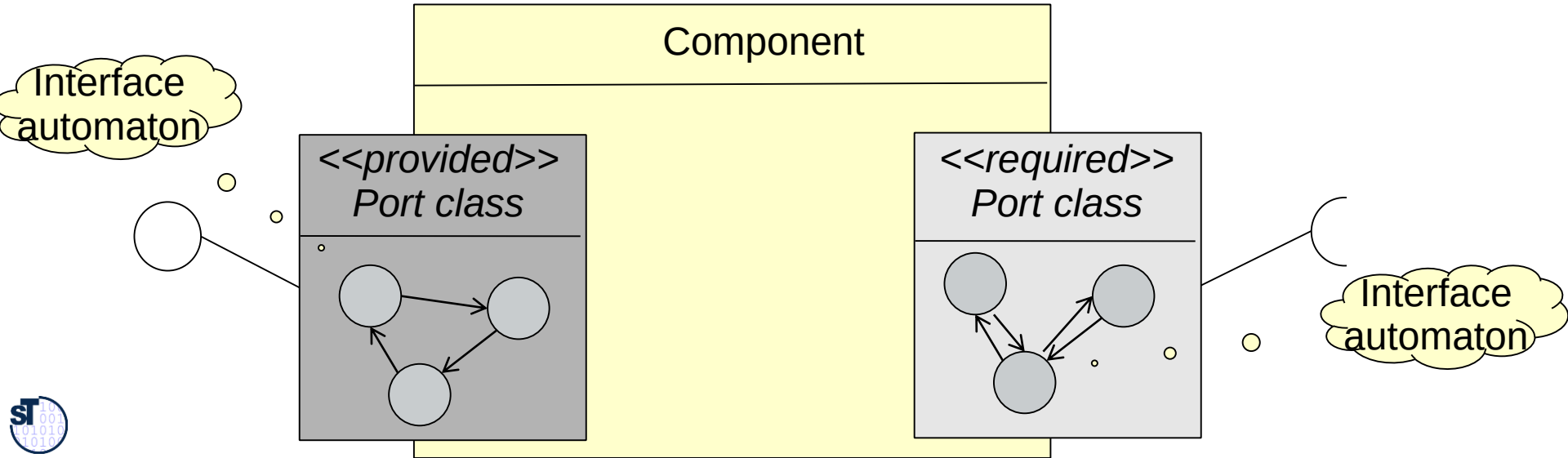
12.5 Searching in Component Repositories by Contract Conformance

- Contract Conformance means semantic substituability

Ports can be Equipped with Interface Automata Contracts

- Ports consist of **port classes** with interfaces and behavior in form of **interface automata (port automata, protocol automata)**
  - provided: normal, *offered* interface
  - required: used, *necessary* interface





Component Protocols with Operational Contracts

- The port protocol automata can be composed to a **component protocol automaton**
- Components may have a **protocol automaton** in which their ports, services, procedures should be called, invoked, or signalled
  - The **provided protocol** specifies in which order the services can be invoked (given by a provided interface automaton)
  - The **required protocol** specifies in which order the services can be invoked (given by a requried interface automaton)
- **The order of component invocation can be specified by a language over the alphabet of the ports,** services, procedures (**state-based protocol contract, operational contract**)
  - Language contains **sets of paths over the alphabet**
  - Finite state automaton (regular language) specify regular sets of paths
    - UML state chart (Hierarchical finite state machine, *prococol machines)*
    - Data flow diagram
  - Stack machine (context-free language)
  - Petri net (regular dialects, context-free and context-sensitive dialects)
- The contract provides an *abstraction* of the implementation of the component
  - Implementations must be proven to be **conformant** to the procotol
- The **conformance checking** is decidable if the protocol language is decidable
- Sets of paths over states (words over state and edge alphabet)

The Golden Rules of Substitutability

- Component A can replace component B if it **offers more** and **requires less**

- Two conditions:
  - A‘s provided protocol must be **stronger (richer, larger)** than B‘s – it must guarantee **more**
  - A‘s required protocol must be **weaker (smaller)** than B‘s – it must assume **less**

- If those conditions hold for all component instances of two component types AT and BT, we say that AT can substitute BT in a program.

Searching by Protocol

- A component C can be **found** in a repository, if a query protocol Q is given with Q <= P(C)
- Search consists of subsumption checking with all component protocols in the repository

- Query protocols can be:
  - Metadata about the component
  - Provided protocols
  - Required protocols
  - Provided **and** required protocols

Declarative Protocols

- A protocol can also be specified as predicates over the states of a component **(declarative contract)**
  - Preconditions (assumptions)
  - Postconditions (guarantees)
  - Invariants
- Then, the protocol consists of logic expressions. The logic should be decidable
  - OCL
  - Description logic
  - Datalog
  - Temporal logic (propositional logic with temporal quantifiers, such as LTL and CTL)
- Subsumption checking of protocols and conformance can be done by reasoning
  - E.g., by subsumption checking of an OWL class hierarchy

The End - Acknowledgements

- Faceted browsing slides are courtesy to Jan Polowinski.