G.A.T.E.

(Graphics Accessible To Everyone)

http://lsd.fi.muni.cz/gate

© Radek Ošlejšek
Faculty of Informatics MU
oslejsek@fi.muni.cz

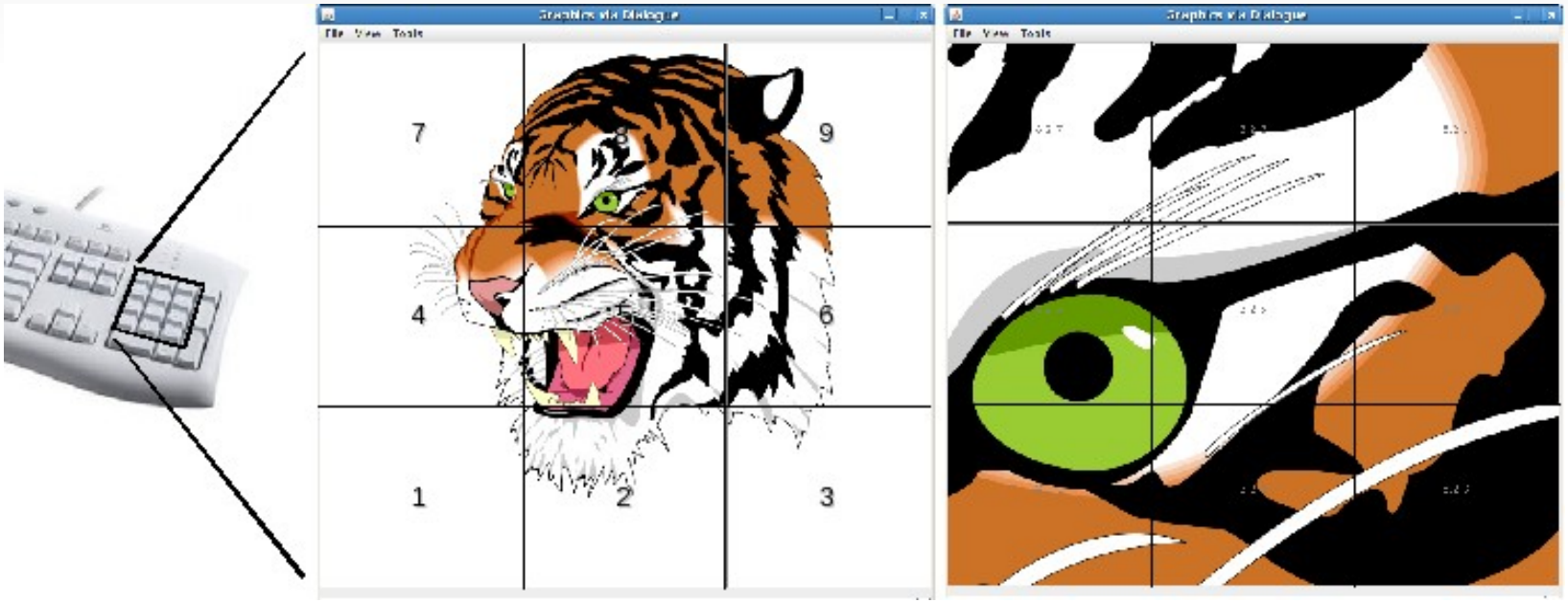
Motivation

- Even a visually impaired people can paint – http://pavla.wu.cz/
- Project goals:
  - graphics accessible *by means of a dialogue* (dialogue is the basic concept for manipulation with images)
  - to develop utilities for picture annotation
  - to provide blind users with support for investigating pictures
  - to develop system for image generation, enabling the blind to easily create some limited form of computer graphics

Navigation

Navigation

- Recursive Navigation Grid
  - Navigation backbone
  - Recursive division of the picture space into nine identical sectors, analogously to the layout of numerical keyboard

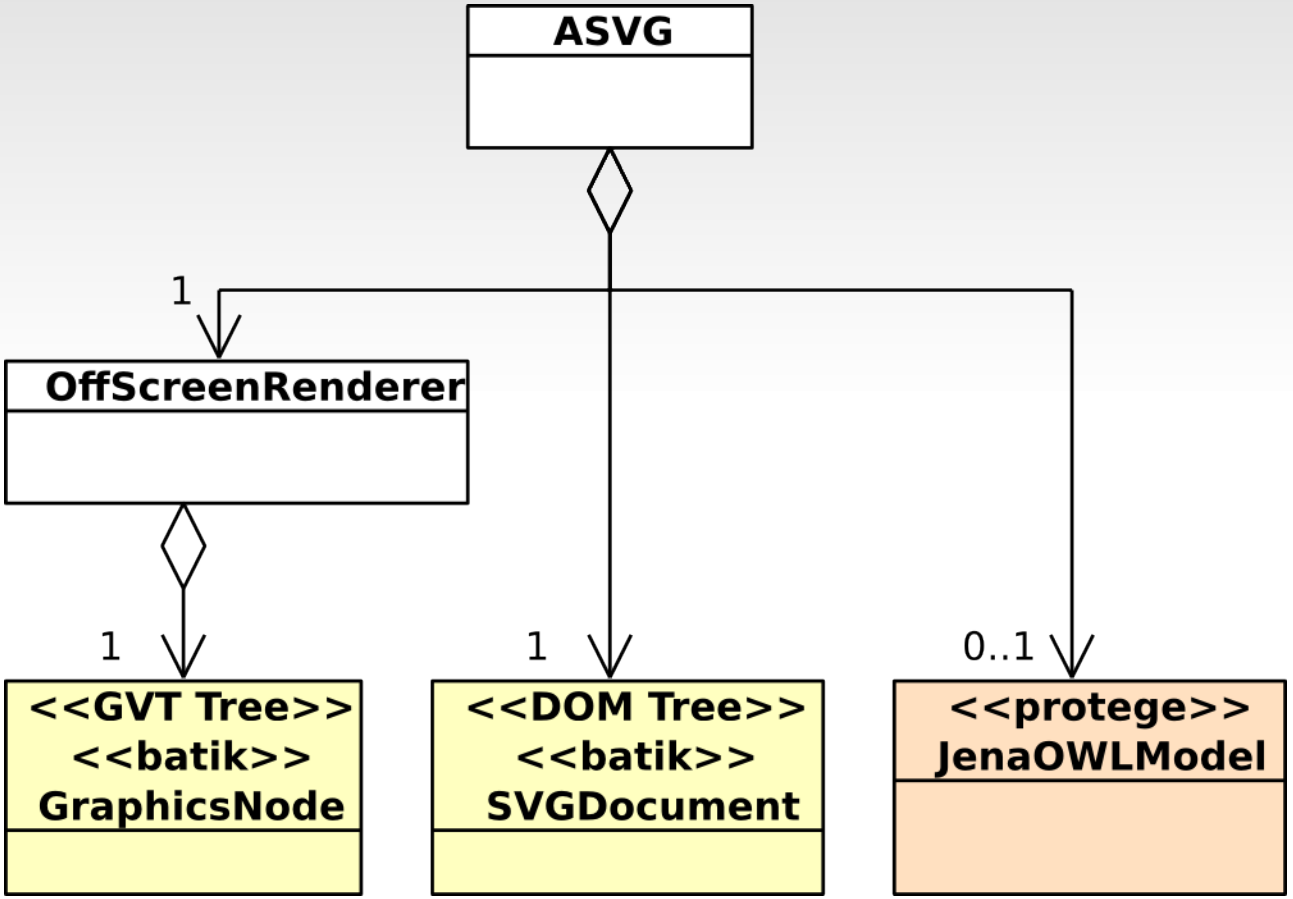


Semantics

Annotation concepts

- SVG (Scalable Vector Graphics) as a basic graphical format
  - defines scene graph, i.e. structural decomposition of graphics
  - based on XML – important feature for the integration of annotation
  - supports vector graphics as well as raster images
  - supports scripting and events handling
- Ontology
  - defines structure and *semantics* of the world
  - OWL (Web Ontology Language), XML-based as well
  - formalization suitable for machine-learning and processing
  - can simplify process of annotation
    - e.g. by offering expected details (sub-objects)
  - can deduce consequences
    - e.g. swan + lake = swimming swan
  - can check correctness of a picture
    - required mainly for the automatic image creation, i.e. painting

Implementation



Ontology-based annotation

- Traditional annotation appends a text description to graphical elements
- Ontology-based annotation formally defines semantics of object by ontology and then classifies graphical elements in the ontology.
  - ✓ Mathematical formalism
  - ✓ Structuring the knowledge
  - ✓ Preventing chaos in terminology
  - ✓ Multilinguality
  - ✗ Generic OWL ontology does not restrict abstraction

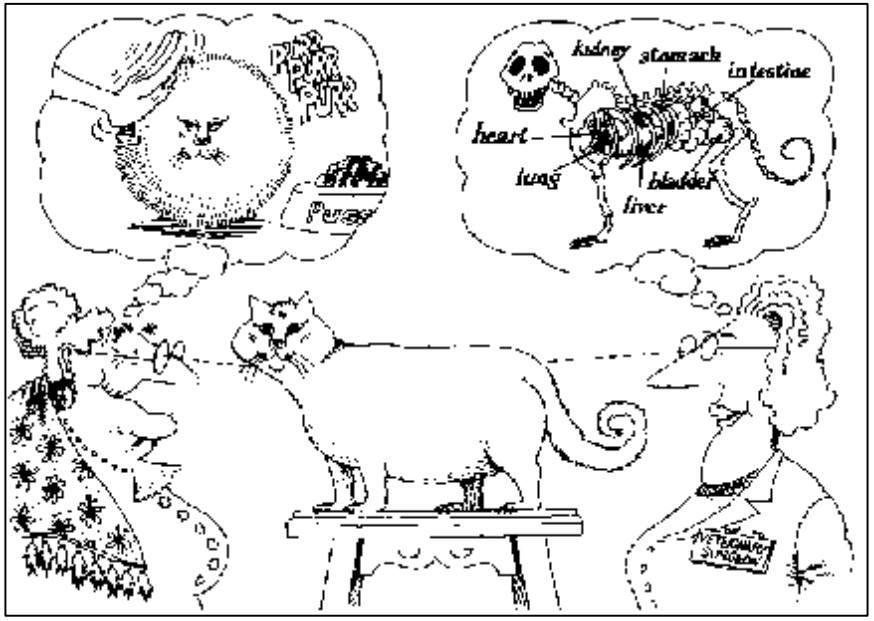
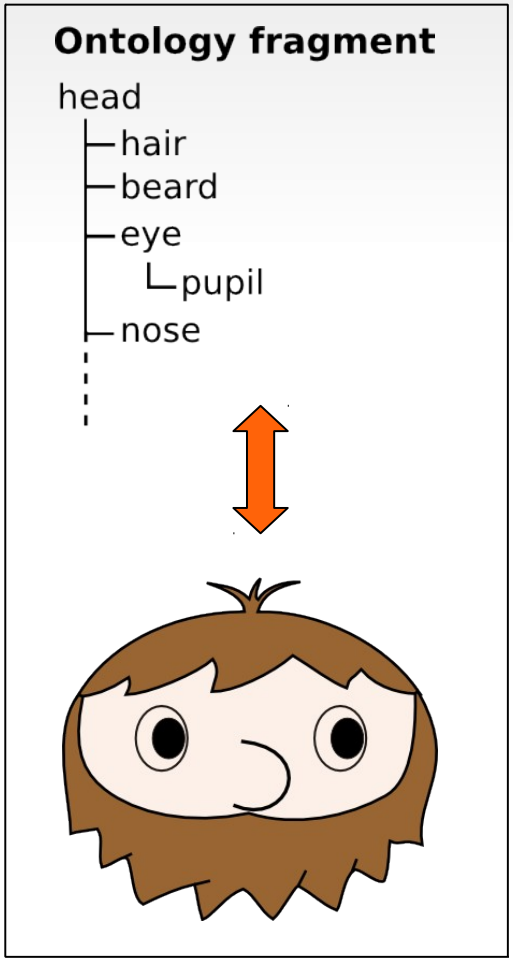


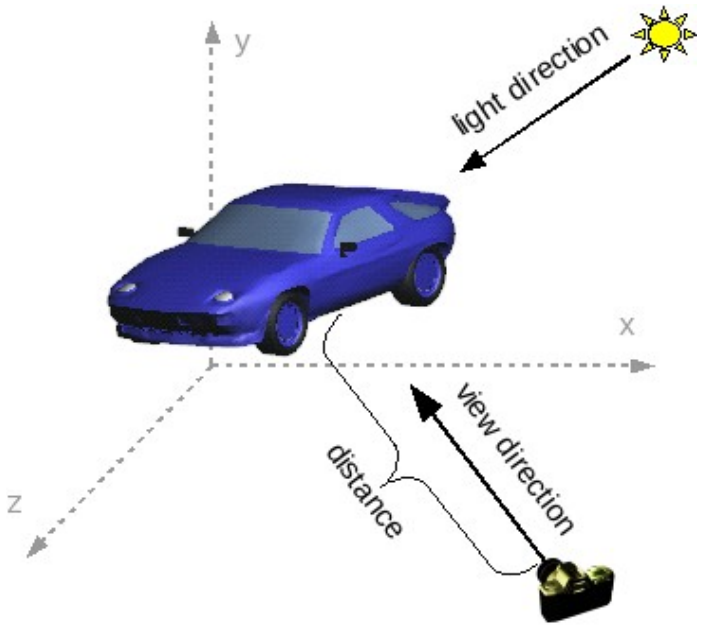
Fig. Ontology-based annotation

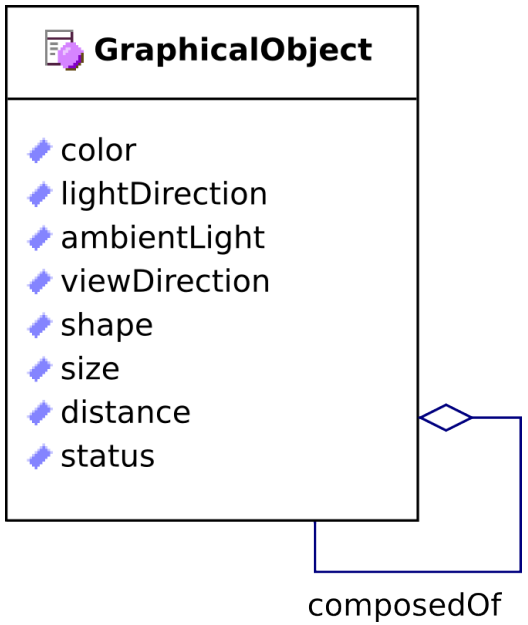
Fig. Abstraction [Booch, G., et al.: Object-Oriented Analysis and Design with Applications. Addison-Wesley Professional (2007)]



Graphical ontology

- Ontology designed from the perspective of graphical objects, their structure and visual properties.
  - E.g. medicinal and poisonous plants are both meaningful ontology categories. But are we able to define their visual characteristics, distinguishing them?
- Predefined visual attributes are similar to those used in energy distribution in 3D scenes, but transformed to approximate verbal form
  - E.g. v*iewDirection = {front_view, rear_view, left_view, ...}*

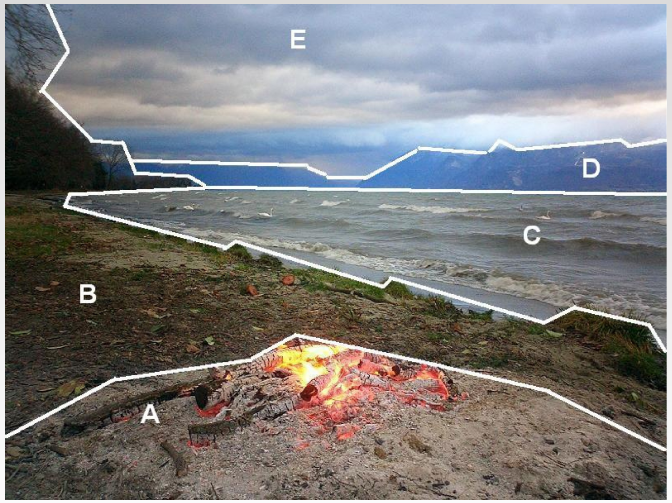


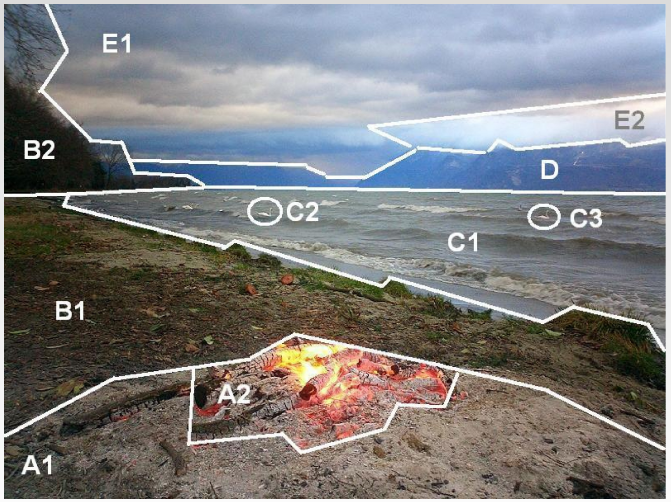


Graphical ontology (cont.)

- OWL ontology prescribing important global visual characteristics and constraints valuable for picture investigation by means of dialogues
- Visual aspects are organized in several classification types:
  - Picture classification
    - Properties related to the whole picture
    - E.g. genre (photo, sketch, ...), central focus (dominant objects)
  - Topological classification
    - Compactness and topology of objects
    - Compact objects = significant shape (human being, ...), non-compact objects (sky, smoke, ...), clusters (forest, crowd, ...)
  - Geometric classification
    - E.g. unusual size (big, narrow, ...), significant shape (oval, angular), etc.
  - Color classification
    - Properties related to illumination and colors.
    - E.g. dominant color (ravens are black), light direction (dark silhouettes due to a back light)

Annotation of raster images (photos)





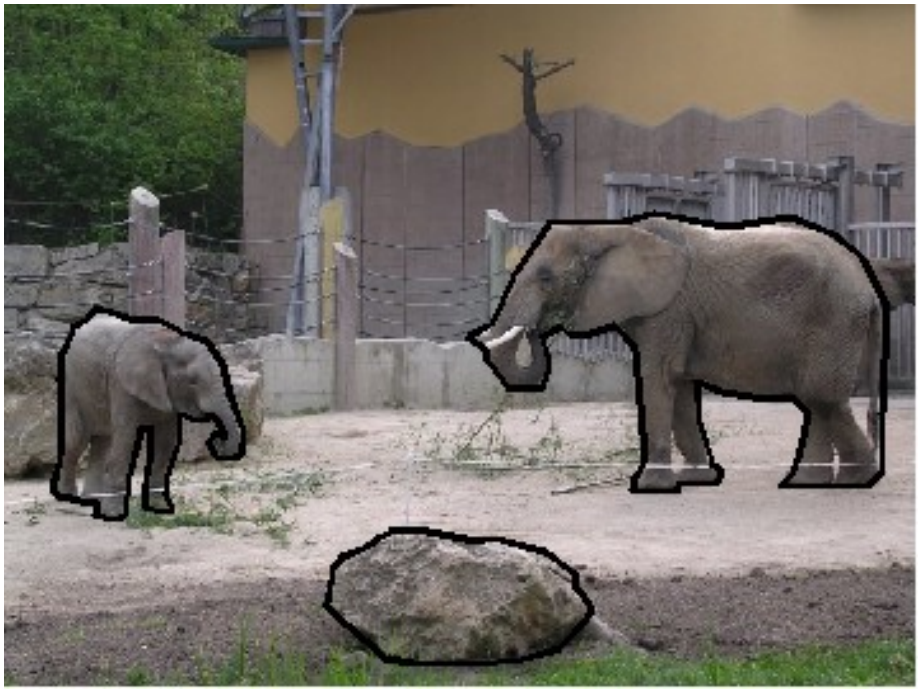
A: an open fireplace with a fire,
B: seashore,
C: Genoa Lake,
D: Alps,
E: cloud sky,
A1: close vicinity of the fire, A2: the fire,
B1: sand shore, B2: seashore vegetation,
C1: the surface of the lake, C2 and C3: a swan on the lake,
E1: dark clouds, E2: light shining from behind the clouds.

Dialogue-based picture investigation

What-where Language

- What-Where Language, WWL
  - Sentences of the form "*WHAT is WHERE*" or "*WHERE is WHAT*"
  - Enables the user to ask simple questions *about the objects* and their *position*
- Objects position
  - Recursive Navigation Grid
- WWL example:

U: What is on the left?
S: On the left there is a small elephant.
U: Where is the elephant?
S: There are two elephants. One is on the left, another on the right.
U: What is in the foreground?
S: In the foreground there is a big gray stone.



Dialogue-based investigation - example

S: This is a color photo of Genoa Lake from its shore. For more information, you can choose Description, Technical information, use keywords or What-Where language.

U: Description

S: In the foreground, there is an open fireplace, shore and a small part of the lake. In the central part, there is the shore near the left-hand part of the picture and the lake is approximately in the right-hand part of the picture. In the background, there is a part of the lake, Alps and a cloud sky.

/Comment: The approximate depth regions and the first semantic level is taken into consideration./

U: What is in the lake?

/Comment: WWL question, the second semantic level is taken into consideration./

S: There are two swans in the lake.

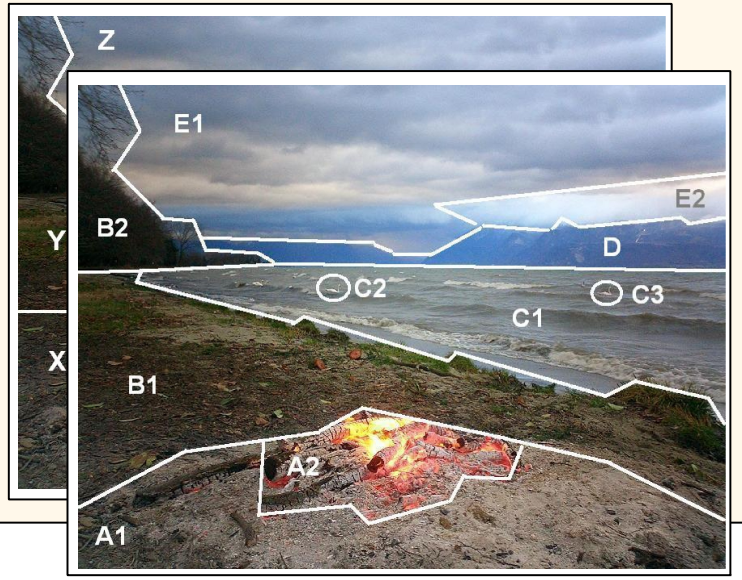
U: Where are the swans?

/Comment: Another WWL question./

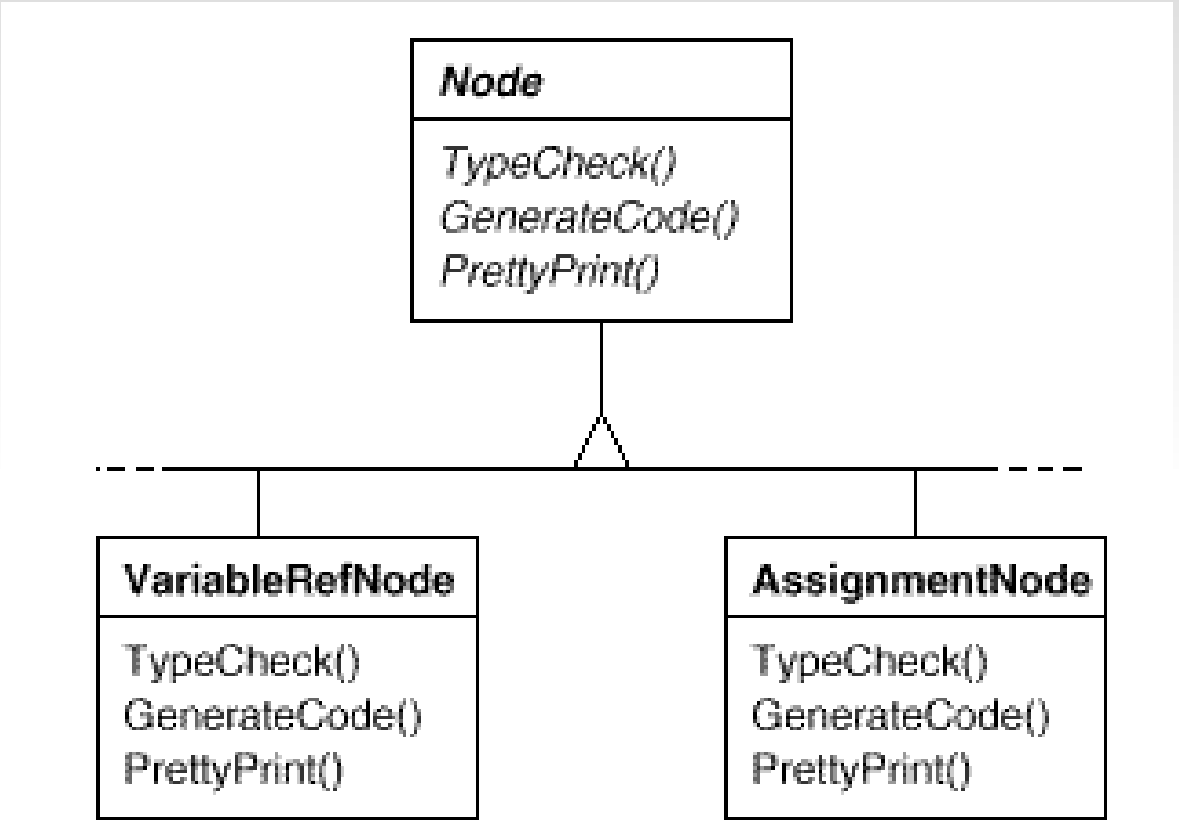
S: Do you wish the approximate localization or use Recursive Navigation Grid?

U: Approximate.

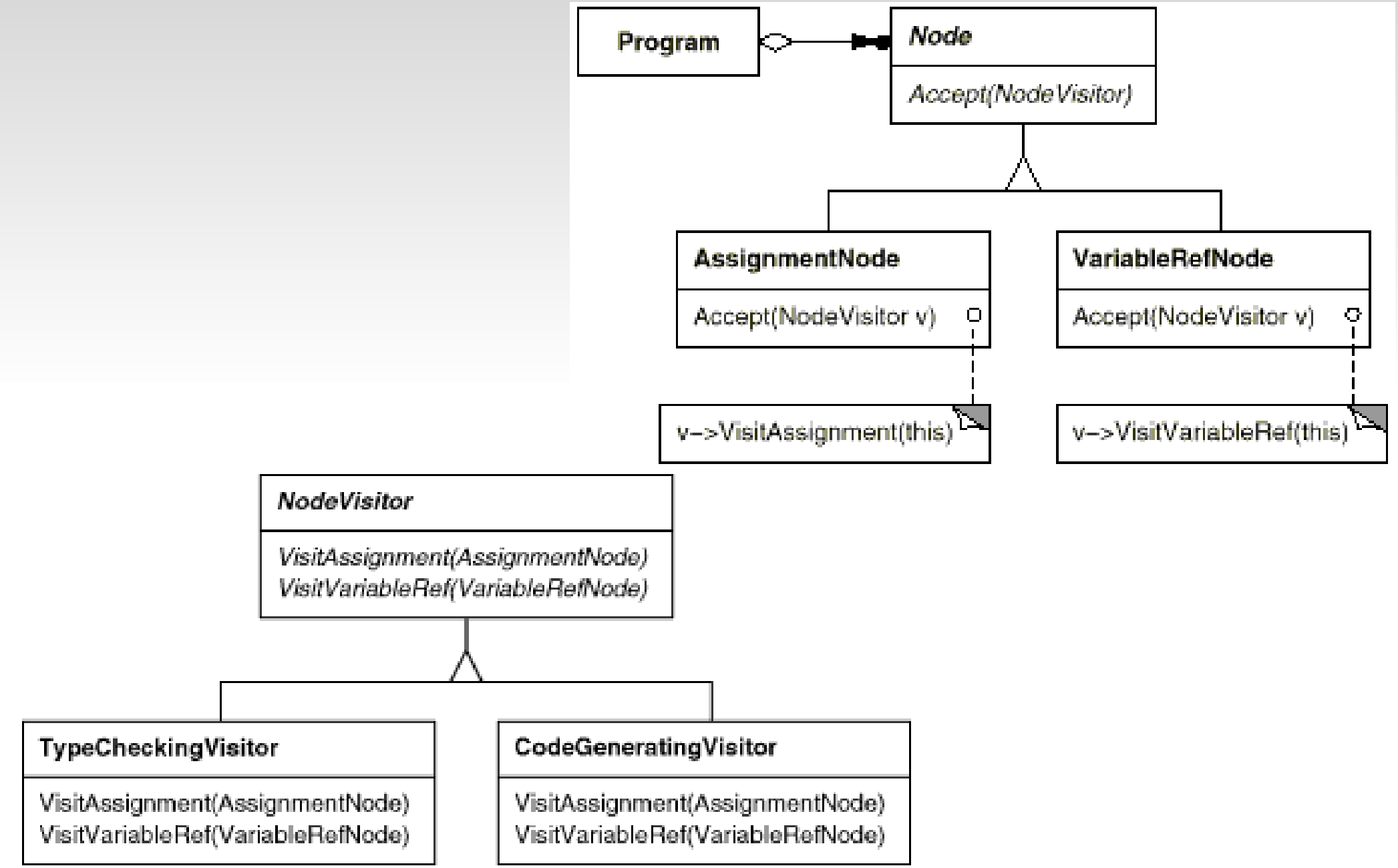
S: One is near the center of the picture, the second one is horizontally near the right-hand side, vertically near the middle.



Visitor GoF Pattern

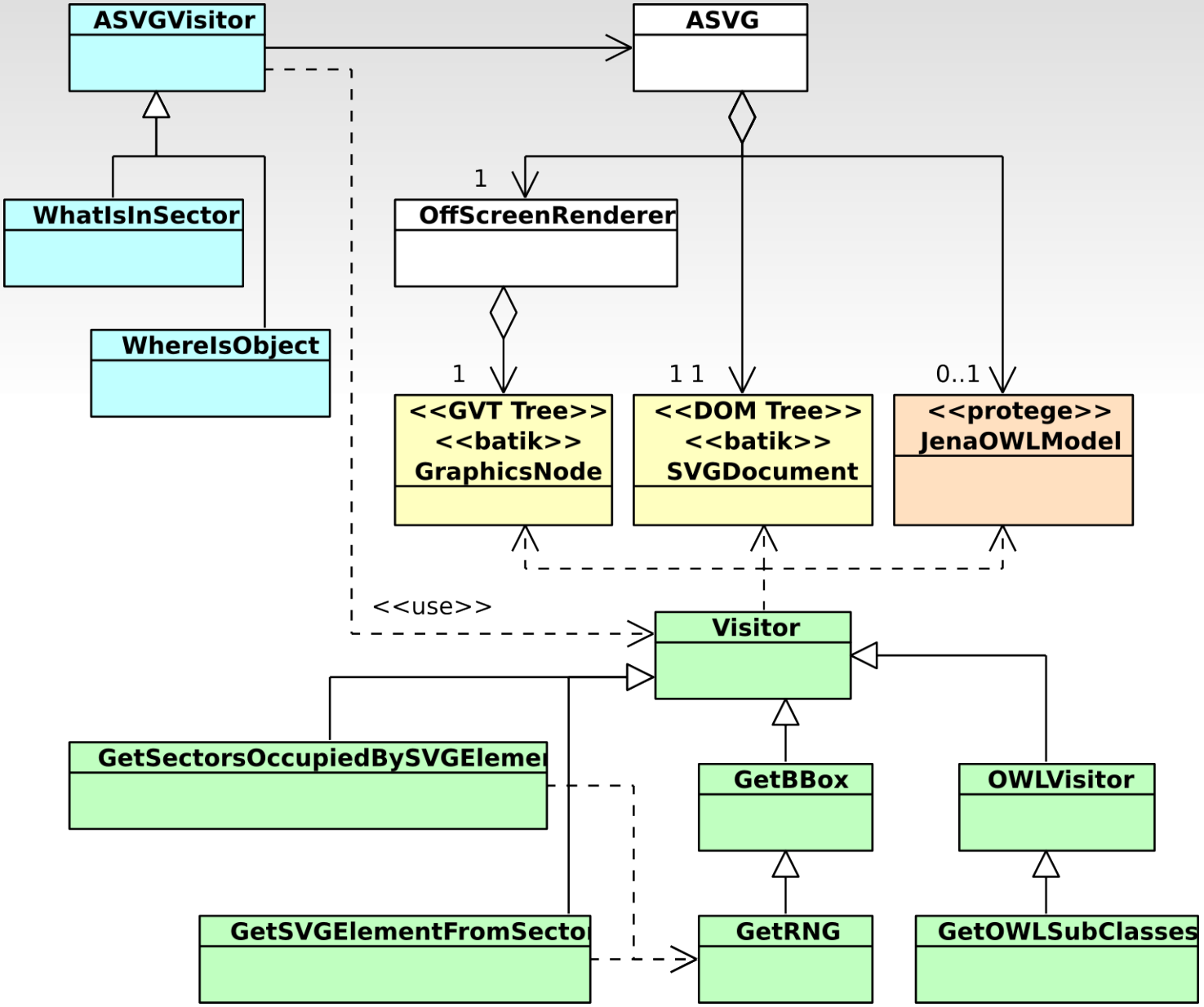


Visitor GoF Pattern (cont.)



Implementation

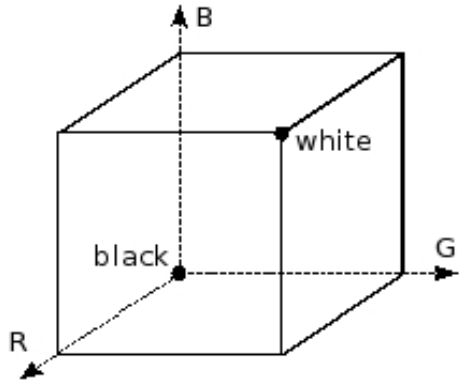
- Tip na BP: Vytvořit sadu visitorů. Vhodné pro 1-2 studenty.



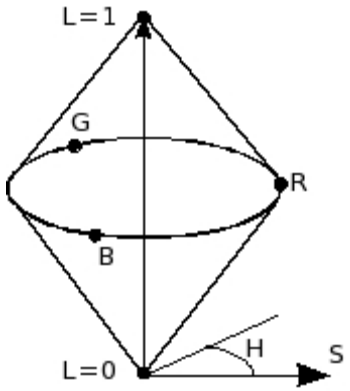
Sonification

Sonification

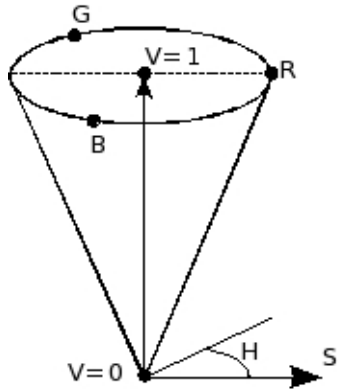
- Color is represented as a triple of primary colors
  - e.g. red, green and blue (RGB)
- Sonification transforms color into the combination of sounds
  - three primary colors have assigned sounds
  - intensity of color component in sonified color drives loudness of produced sound
  - information about the intensity of three components usually does not help the user to imagine the color



RGB color model



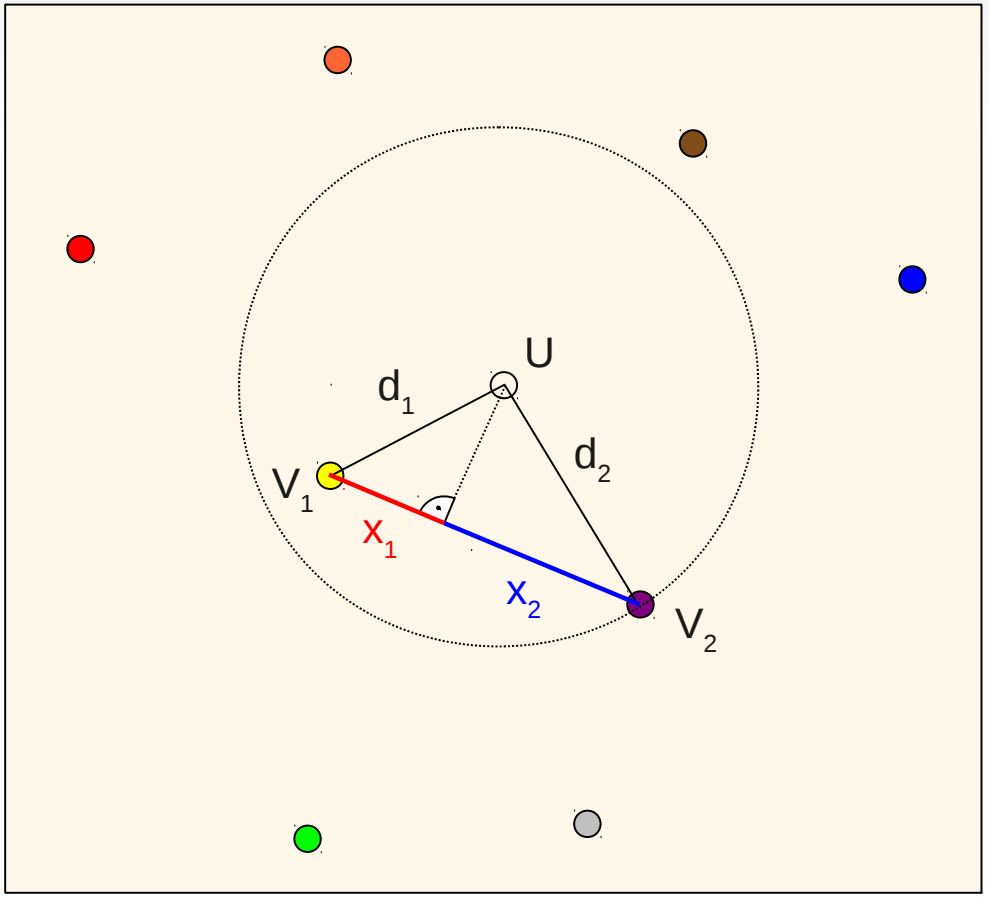
HSL color model



HSV color model

Semantic color model

- Semantic Color Model (SCM) is approximative model which attempts to express a color as a combination of two well-known colors.
  - it is easier to imagine colors described in term like: yellow-green, gay with orange shade, etc.
  - it is easier to distinguish two sounds coding two well-known colors



Implementation

- PHP + Flash inside client
- Tip na DP: Přepsat v Javě, přesunout na stranu serveru (ustanovit interface), testování použitelnosti (latence), optimalizace a úpravy SCM, ...

Painting by means of dialogue

Result depends on experience

?

Result depends on experience

Paint Mona Lisa

Mona Lisa is a smiling woman

What is mona lisa?

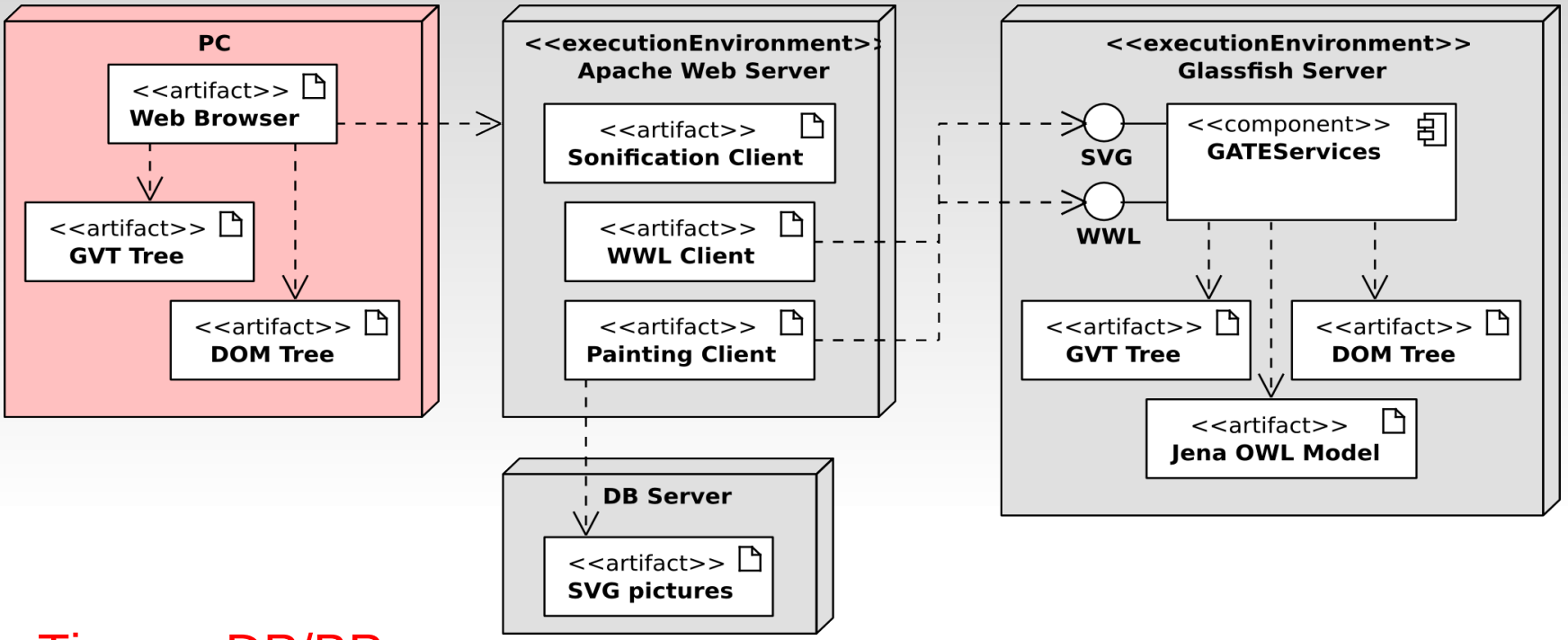
Knowledge base

Painting example

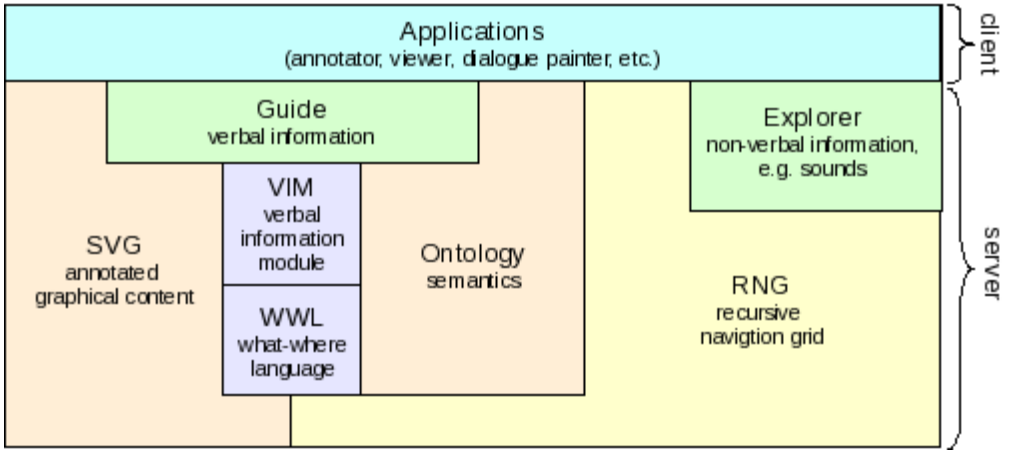
U: Put a comet in the sector 9.
U: Put a snowman into the bottom left corner.
U: Write the text „Merry Christmas and Happy New Year“ into the horizontal center, color yellow.
U: Write the text „PF 2010“ into the bottom right corner, color blue.
U: Set background to snowflakes.
U: Generate.

Fig. The Chrismas card generated by a blind user

Implementation



- Tipy na DP/BP:
  - Webové služby pro znalostní bázi (přesun DB obrázků na stranu serveru, strukturované záznamy, …)
  - Propojení generátoru s WWL
  - VIM modul
  - ...



Thank you for your attention!