Integrating Dialogue Systems with Images

Jaromír Plhák, Ivan Kopeček and Radek Ošlejšek

{xplhak, kopecek, oslejsek@fi.muni.cz

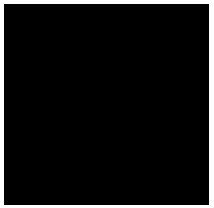
Faculty of Informatics
Masaryk University, Brno
Czech Republic

Motivation

- Looking at this photo from a holiday approximately three years ago: It is somewhere in the Czech Republic, but where exactly? And what is the name of that castle? And who is the man behind your girlfriend? ...

Figure 1: Photo from a holiday

> ***Communicative image***
>
> *An "intelligent" image that is able to discuss its content with the user*

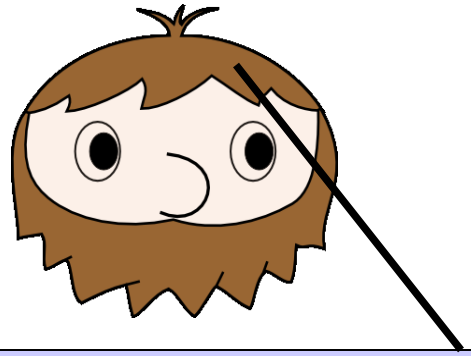


Key Concepts

- Dialogue-based communication in natural language
- Ontology-based knowledge base
  - Ontology Web Language
- Scalable Vector Graphics
  - Raster image and objects annotation
  - owl:import
- Information sources and algorithms
  - EXIF data; Face detection; Similarity search algorithms
  - Information from users
  - Crowdsourcing

Ontologies

- Graphical ontology
  - General visual characteristics
    - Unusual size
    - Typical shape
    - Dominant color
  - Navigation
    - General and mutual position
- Domain-specific ontologies
  - Family and friends
  - Sights
  - Nature



```
<Hair rdf:ID="head">
   <hasColor rdf:resource="#brown" />
</Hair>
```

Ontology fragment

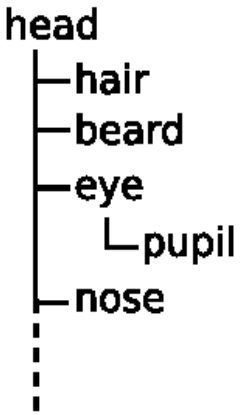
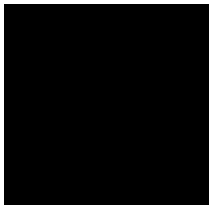


Figure 2: Part of graphical ontology



Dialogues with Images

- Communication modes
  - Information retrieval mode
  - Image information supplementing mode
  - Free communication mode
- Communication Analysis
  - Small fragment of natural language
    - Relatively simple grammars
    - Frames technology
    - Standard techniques for misunderstanding solving

How far is it from this hotel to the nearest beach?

How far is it from <SLOT1> to <SLOT2>?

Dialogues with Images

- Grammar generator
  - General ontology structure (concepts and relations)
  - Set of objects in the image
  - The restrictions on the given slots (cardinality or value type)
  - Actual dialogue strategy
  - Dialogue context
  - History of the communication

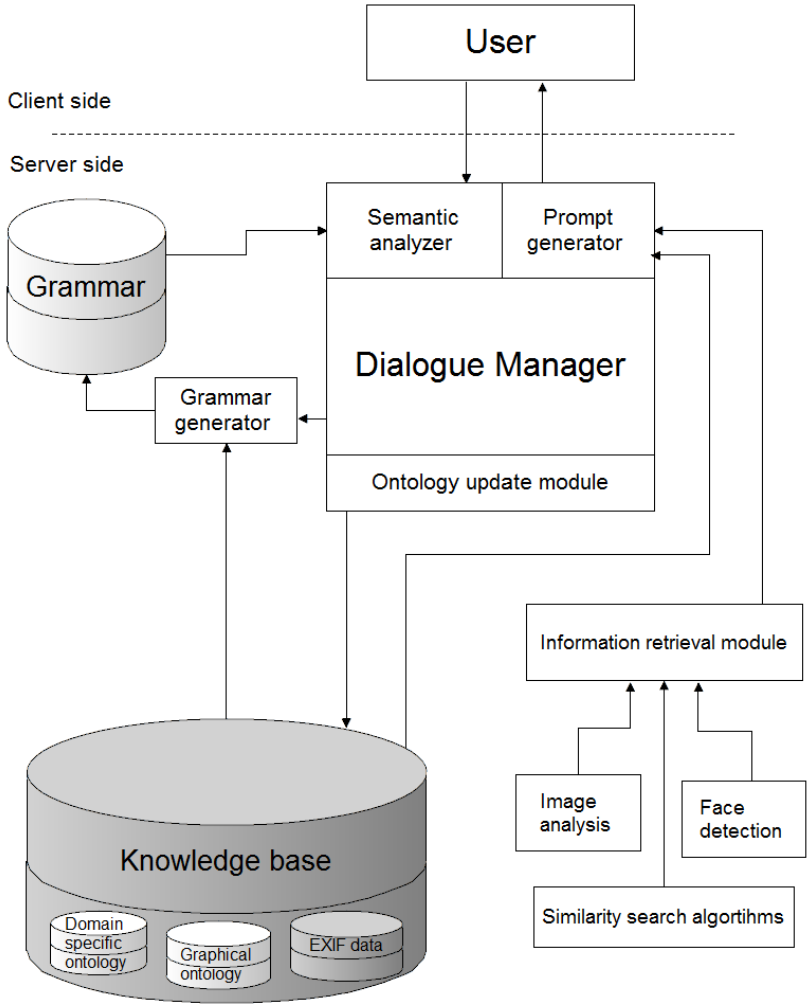


Figure 3: Example of application structure

Dialogues with Images

S: *This is a photo of a group of your friends standing before the Pernstejn castle. Now, you can open and view the photo or start the communication with image.*
U: Open the image, please.
S: */opens the photo/ File is opened. Now, you can start the communication with the image.*
U: What is the resolution of this picture?
S: *The resolution of this picture is 8 MPix. Enter your next query, please.*
U: Who is in the picture?
S: *There is a group of your friends. George, John and Elisabeth. There are also four other groups of tourists.*
U: Give me the detailed information about this caste.
S: *Pernstejn Castle is a castle located on a rock above the village of Nedvedice and the rivers Svratka …*
U: Thank you, bye.
S: *Good bye!*

Figure 4: Example of a dialogue communication in Information retrieval mode

Implemented Modules

- GateDialogueBean
  - Dialogue manager
- GateOwlBean
  - Processing of ontologies
  - OWL reasoner Pellet
- GateSvgBean
  - SVG processing
  - Batik library
- Client scripts
  - Web browser plug-in

Applications Domains

- Personal photos and albums
  - Organize
  - Search
- E-learning study materials linked to domain-specific knowledge base

S: In the picture there is the chemical structure of antioxidant resveratrol.
U: What is antioxidant resveratrol?
S: Antioxidant resveratrol consists of two benzene nucleus and three hydroxyl groups.
U: What is benzene nucleus?
S: Benzene nucleus …
/The system uses definition from the chemical ontology/

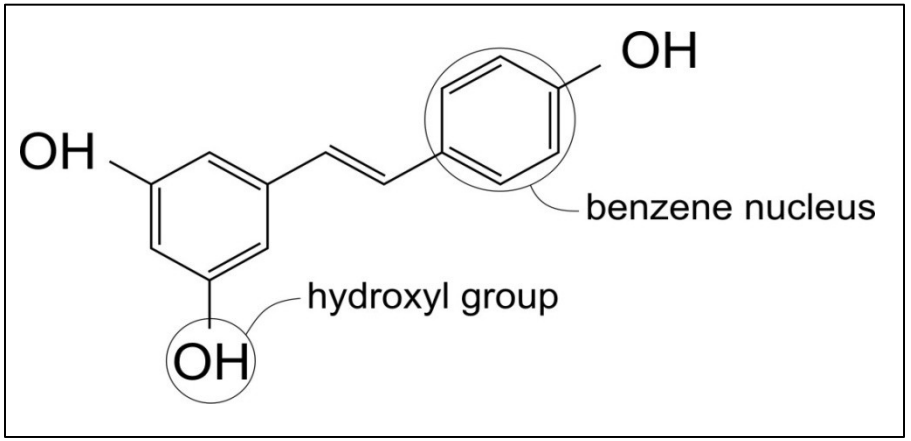
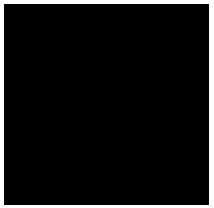
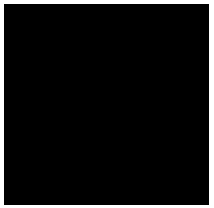


Fig. 2: Antioxidant resveratrol



Application Domains

- Internet graphics exploration
  - Social networks
  - Electronic news
  - Web page editing
- Applications for people with special needs
  - Visually impaired people
  - Older adults
  - People with any constraint in communication



Conclusion and Future Work

- Communication with image by means of natural language
- Efficient images investigation using frame based dialogue management
  - Supported by the ontologies
- Automated generation of grammars from knowledge base

- Testing of the prototype
- Abstraction of ontologies
  - Semantic terms, attributes and relations valuable for the system
- Modules that search for additional knowledge
  - Internet, picture analysis

Thank you for your attention!

xplhak@fi.muni.cz