web: suruse.uni-ruse.bg

GRAPHIC EDITOR WITH NATURAL-LANGUAGE INTERFACE

Desislava Baeva, Ivan Ivanov

Angel Kanchev University of Ruse

Abstract: *Graphical editing can be a challenging task because of the small size of the screens portable mobile devices that are now frequently used to capture and edit images have. A system is described herein that allow computing device users to expressions graphically a text messaging session. The multimodal natural-language interface that combines speech and direct manipulation is an interesting way of solving this kind of problems. This paper observes existing image editing practices and derives a graphic editor with natural-language interface. In particular, it is used natural language for expressing desired new parts or changes to an image. We illustrate the need for such a combined system, and also give examples of how a natural language facility beneficially augments the user's ability to navigate a knowledge base graphically.*

Keywords: *multimodal interfaces; natural language processing; graphical image editing.*

INTRODUCTION

We are witnessing an era in the development of technologies which not only interpret letters and sounds, but also phrases and sentences. This way they help the users even more, since they can talk, acknowledge, and understand their language. Pioneers in this area include the Google Translator, which offers free online translations between 57 languages, the question answering computing system IBM Watson, that won against the USA champion of “Jeopardy“ or Apple’s mobile assistant Siri, which reacts to voice commands and answers questions in English, German, French and Japanese.

Studies show the next generation of information technologies are going to be perfected in the use of natural languages, so that users can communicate with them freely using their own language. Natural language processing will offer either automatic translation, or translation assistance. They will provide summaries of the dialogue or of different documents. Computer assisted training is will assist with the ease of integration of minorities and foreigners. This level of processing exceeds the simple collections of symbols and dictionaries, programs for spell checking and rules of pronunciation. Technologies must overcome the simplistic approaches and begin comprehensive modelling of the language by taking into consideration syntax and semantics so they may understand the meaning of the questions and generate extensive and adequate answers [Blagoeva 2009].

There are attempts that provide alternatives to common solutions in which users input information in natural language without using mouse or keyboard and the systems interpret them automatically into formal data (machine language). This method is attractive, but information extracted from free text is unreliable, and in the foreseeable future users must stick to using specific language formats (formal data).

Exposition

Natural language for input suggests a sequence of actions needed to process the text instead of directly comprehending it.

We can summarize four types of natural language processors based on the level of control the user has over the input data.

1. Cursive handwriting recognition software – the user assumes direct control on the graphic level by writing the documents by hand using a digital pen-based input device. In order for this information to be used on other program levels the editor has to run character recognition.

2. Natural language text editor – The user does not have direct control over the graphemes, but can reliably convert his input information in a sequence recognized by the editor.

3. NL-Menu – The user controls the program on syntax level by selecting allowed words or phrases in order to complete his sentence. This way the user loses direct control over the word choice, but reliably encodes the language structure of the sublanguage so that the editor can understand and successfully comprehend the meaning.

4. WYSIWYM Editing – The user has control over the semantic level by choosing conceptual additions towards the current context. The program does not perform analysis or parsing of any kind, since at the highest level knowledge is directly controlled by the user.

A well-known fact is that there is a new generation of technologies that effectively use natural languages, as it is accepted that today computers must be intelligent enough to comprehend text and speech, to translate, to search the web, to answer questions etc. SHRDLU is a program for understanding natural language written by Terry Winograd. It carries on a simple dialog with a user in order to move objects, naming collections and querying the state of a simplified block world. [SHRDLU] PixelTone is software developed by Eytan Adar, Gierad Laput and a team from Adobe. It has a multimodal photo editing interface that combines speech and direct manipulation. It was created to assist users that use small screens on portable devices such as camera phones. Speech interfaces can make complex tasks more accessible because they allow users to simply state goals without first leaning an interface [Laput 2013]. Online natural language editor of the Experimental Media Group at the St. Lucas School of Arts, Antwerp, Belgium, that can be used to assist users when it comes to supporting their design activities. It transforms natural language (English) into entities that can initiate and control the graphical output of NodeBox (software intended to create graphical output) [Gravital].

In the field of language technologies there are a number of products, technologies and resources in Bulgarian. There are a number of software products for generating synthesized audio output with appropriate cadence and intonation, simulating speech patterns, spell and grammar checkers. Most programs for automatic translations, on the other hand, do not always offer linguistically correct

translations, especially when the translation is not from Bulgarian. This is caused by the specific language characteristics of the Bulgarian language.

The purpose of this article is to describe and popularize the graphical editor controlled by text in natural language interface in Bulgarian language. The software reads a command sentence input by a user, tokenizes it and compares individual words with its databases. If an incorrect or missing word from the database is encountered, the user will be presented with a predefined feedback message.

Our approach to human-computer interaction provides the basis for an integration of different interaction styles, in our case natural language and graphics, in a multi-modal information system. The choice of an appropriate combination of modes is essential for a successful interface design. As there is not yet a complete theory of multi-modal interaction, this question has to be answered experimentally. Therefore, we propose a modular architecture for multi-modal interfaces composed of interaction tools.

Architectural model

The architectural model of the application consists of the following elements:

- User input commands using the user interface elements.
- Graphical user interface consists of these interface elements (menus, panel for drawing, send button and text field for chat history and command input) and listeners;
- The system core consists of modules for database connection, natural language dialog and processing, feedback messages and drawing and figure manipulation.
- The module for database connection is used to send requests.

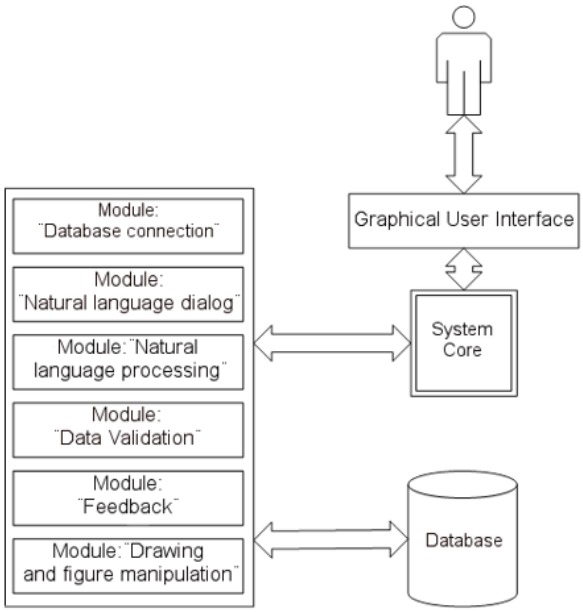

Figure 1. System architecture

- The Natural Language Understanding module receives the user input and analyses it using the keyword matching, speech act detection, and semantic parsing. If the user answers the system's question appropriately, keyword matching should be

able to handle the user utterance. However, the user may choose not to respond to the system's question. In order to handle unexpected user utterances, we are using speech act detection and semantic parsing. The module accepts array of words and uses the functionality of the module for database connection, checks the validity of the command and executes the request. If needed it calls the module for feedback messages.

- The module for drawing and figure manipulation includes drawing, rotating and changing colour of border/background and movement of the current figure [Hettige 2015].

In order to make a thorough automatic analysis of the textual information in a specific field, a conceptual model is needed to assist in the interpretation of objects, identified in the text and the connections between them. The model is a conceptual structure, stored in declarative formalism, which defines the general notion of the field and its relations. An example for these are ontologies and databases. They help finding important pieces of information in texts and their structure is of key importance to developing reliable systems of information excerption with high fidelity.

The model describing our subject area must the data extraction task, as well as future searches in the conceptual templates. It must provide an opportunity for finding more common or specific concepts so that it can assert similarities between user records.

Speech coherence

There are three levels of **speech coherence**: syntactic coherence, semantic coherence and pragmatic coherence [Dilley, Bateman 1992]. Syntactic coherence deals with the immediate connectivity among adjacent segments (in texts this is often called text cohesion). Semantic coherence ensures the well-formed thematic organization of a discourse. Discourse segments are connected by semantic ties. A discourse is pragmatically coherent if it is compatible with the addressees' interpretative ability. In our system syntactic coherence is enhanced by the way we present the natural language output in our graphical environment. Semantic coherence is supported by the use relations which ensure the connectivity between our text segments. Pragmatic coherence is supported by the mere fact that we are using text as a medium for meta-dialogs, as these are difficult to understand on a graphical level.

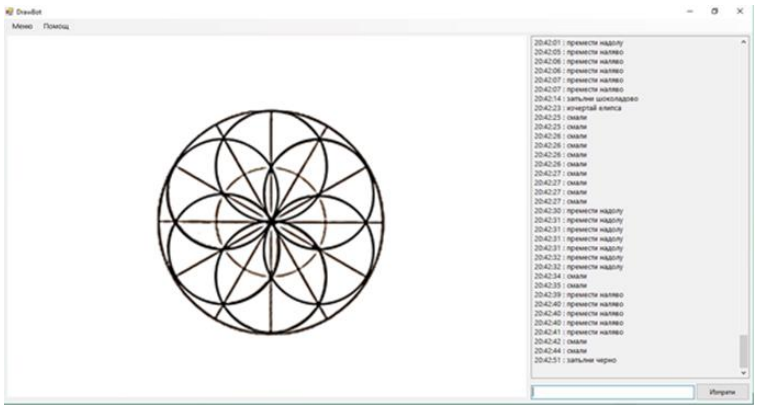

Figure 2. Example of drawn figure

ONTOLOGY

Based on these requirements for graphics editors, the need for ontology that provides access to data with the usage of semantic queries in SPARQL arises.

In this paper we present a prototype ontology-based knowledge database implemented in Protégé 5.0., with which software system can do knowledge elicitation and validation of knowledge facts. The role of ontologies in information systems is refined for knowledge system development. Only classes can't answer many questions so we also need to define link inside or between these classes (such as properties) we use property which show relationship between individual to individual.

Besides concepts and terms, the model describes the connections between the objects in the subject area. The connections that relate to the natural language processing and the understanding of text are of type IS-A, IS-TO, SAVE-AS and others. Connections of type IS-A and IS-TO are linked like parent-child and link specific with more common terms and those of type HAS-ACTIVITY connect manipulations with the object class.

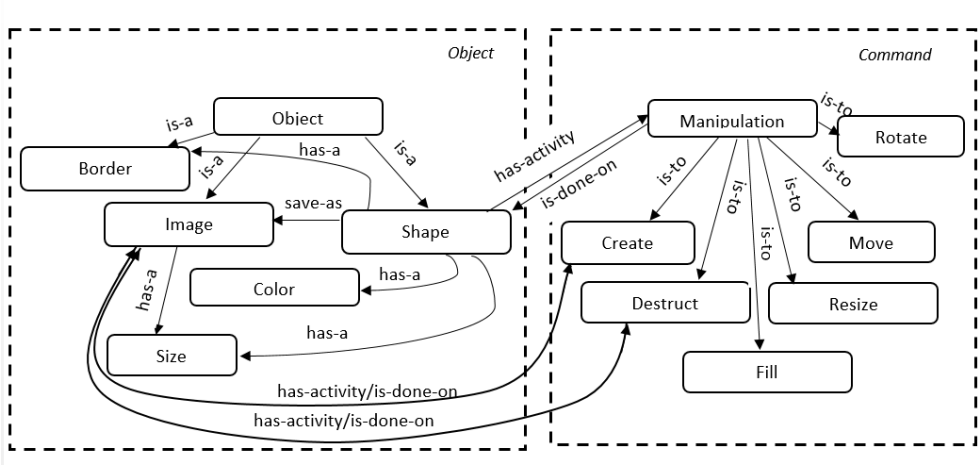

Figure 3. Example of ontology design for domain of 2D computer graphics

The architecture has been introduced in order to realize a user friendly natural language system on the specific domain. We have grouped objects with their properties. We have formalized knowledge about the domain by introducing the main concepts (such as the main kinds of objects and manipulations...). In the following we show a sample of useful predicates: This module forms SPARQL queries to create a set of graphical shapes by the <feature, attribute> pairs specified by the Dialogue Manager. Its output is analysed by the object creator in order to provide clear specifications of each shape that may.

In the following we show a sample of useful predicates:

Predicates

<Manipulation> is-to <Create>
<Create> hasIndividual <Н а р и с у в а й>
<Create> is-done-on <Shape>
<Shape> has-a <Color>
<Color> hasIindividual <Ч е р в е н>
<Shape> hasIindividual <К в а д р а т>

OWL/XML

```
<ObjectPropertyRange>
        <ObjectProperty IRI="#is-to"/>
        <Class IRI="#Create"/>
</ObjectPropertyRange>
<Declaration>
        <NamedIndividual IRI="#н а р и с у в а й"/>
</Declaration>
<ObjectPropertyRange>
        <ObjectProperty IRI="#is-done-on"/>
        <Class IRI="#Shape"/>
</ObjectPropertyRange>
<ObjectPropertyRange>
        <ObjectProperty IRI="#has-a"/>
        <Class IRI="#Color"/>
</ObjectPropertyRange>
<Declaration>
        <NamedIndividual IRI="#ч е р в е н"/>
</Declaration>
<ClassAssertion>
        <Class IRI="#Square"/>
        <NamedIndividual IRI="#к в а д р а т"/>
</ClassAssertion>
```

A domain lexicon, containing the definitions of lexical items of all the words that may appear in the application domain.

CONCLUSION AND FUTURE DEVELOPMENT

Creation of graphic objects with the help of different application require preliminary preparations and sufficient knowledge in software products on the user side. Specific terminology will also make it difficult for people that are not accustomed with it. Attempts to create graphic images on mobile devices make the interaction between human-computer even more difficult, due to the small screen. Natural language interfaces allow users to request their own goal without initial training.

The next generation of information technologies must improve the use of natural languages to such extent that users must communicate freely using their own language.

After the implementation of voice recognition and improvement of the knowledge base the presented product will be helpful for:

- People with disabilities, to visualise specific objects
- To help create raster images on small displays

- To create play scenarios for educational activities for kids.

Natural language possesses properties that make it preferable over the graphical mode of expression for such meta-interaction and hence natural language generation needs to be supported even in graphics-oriented interfaces. Addition of voice recognition will significantly make it easier for users to work with the application.

REFERENCES

[Blagoeva 2009] Blagoeva D., S. Koeva, V. Murdarov, White Paper Series the Bulgarian language in digital age.

[Dilley, Bateman 1992] Dilley, S., Bateman, J., Thiel, U., & Tissen, A. Integrating natural language components into graphical discourse. In Proceedings of the third conference on Applied natural language processing (pp. 72-79). Association for Computational Linguistics, 1992.

[Laput 2013] Laput G. P., M. Dontcheva, G. Wilensky, W. Chang, A. Agarwala, J. Linder, and E. Adar, "PixelTone: a multimodal interface for image editing," presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 2013.

[Hettige 2015]. Hettige, B.,A. Karunananda, "Octopus: A multi-agent chat-bot", Proceedings of 8th International Research Conference KDU 2015

[SHRDLU] http://hci.stanford.edu/winograd/shrdlu/.

[Gravital] www.clips.uantwerpen.be/~vincent/gravital/GLPDaemon/.

[Yokobot] http://yokobot.com/index.php?p=about.

[NLprocessing] http://www.slideshare.net/DivyaSugumar1/natural-language-processing-in-alternative-and-augmentative-communication.

[Protégé] http://protegewiki.stanford.edu/wiki/Main_Page.

[Нишева 2009] Нишева М., Д., Изкуствен интелект, Издателство Интеграл, Добрич 1995.

CONTACT ADDRESSES

Pr. Assist. Desislava Baeva, PhD
Department of Informatics and Information Technologies
Faculty of Natural Sciences and Education
Angel Kanchev University of Ruse,

8 Studentska Str., 7017 Ruse, Bulgaria
e-mail: dbaeva@ami.uni-ruse.bg

Ivan Ivanov, BSc student
Department of Informatics and Information Technologies
Faculty of Natural Sciences and Education
Angel Kanchev University of Ruse,
8 Studentska Str., 7017 Ruse
e-mail: i.zenitch@gmail.com

ГРАФИЧЕН РЕДАКТОР С ЕСТЕСТВЕНОЕЗИКОВ ИНТЕРФЕЙС

Десислава Баева, Иван Иванов

Русенски университет „Ангел Кънчев”

Резюме: *Графичното редактиране на изображения на малки по размер екранни на преносими мобилни устройства, може да бъде предизвикателство. Мултимодалният интерфейс на естествен език, който съчетава речта и пряката манипулация, е интересен начин за решаване на този проблем. Тази статия разглежда съществуващите практики за редактиране на изображения и описва пример за графичен редактор с естественоезиков интерфейс. По-специално, използван е естествения език за изразяване на желанието за създаване на нови елементи или правене на промени на определено изображение. Илюстрирана е комбинирана система и е представен пример за това, как естественият език увеличава способността на потребителя да се движи графично в базата от знания.*

Ключови думи: *мултимодални интерфейси; естественоезиков интерфейс; графично редактиране на изображения.*

Requirements and guidelines for the authors - "Proceedings of the Union of Scientists - Ruse"

The Editorial Board of "Proceedings of the Union of Scientists - Ruse" accepts for publication annually both scientific, applied research and methodology papers, as well as announcements, reviews, information materials, adds. No honoraria are paid.

The paper scripts submitted to the Board should answer the following requirements:

1. Papers submitted in Bulgarian, Russian and English are accepted. Their volume should not exceed 8 pages, formatted following the requirements, including reference, tables, figures and abstract.

2. The text should be computer generated (MS Word 97 for Windows or higher versions up to Word 2003) and printed in one copy, possibly on laser printer and on one side of the page. Together with the printed copy the author should submit a disk (or send an e-mail copy to: desi@ami.uni-ruse.bg).

3. Compulsory requirements on formatting:

font - Ariel 12;
paper Size - A4;
page Setup - Top: 20 mm, Bottom: 15 mm, Left: 20 mm, Right: 20mm;
Format/Paragraph/Line spacing - Single;
Format/Paragraph/Special: First Line, By: 1 cm;
Leave a blank line under Header - Font Size 14;
Title should be short, no abbreviations, no formulas or special symbols - Font Size 14, centered, Bold, All Caps;
One blank line - Font Size 14;
Name and surname of author(s) - Font Size: 12, centered, Bold;
One blank line - Font Size 12;
Name of place of work - Font Size: 12, centered;
One blank line;
abstract – no formulas - Font Size 10, Italic, 5-6 lines ;
keywords - Font Size 10, Italic, 1-2 lines;
one blank line;
text - Font Size 12, Justify;
references;
contact address - three names of the author(s) scientific title and degree, place of work, telephone number, Email - in the language of the paper.

4. At the end of the paper the authors should write:

The title of the paper;
Name and surname of the author(s);
abstract; keywords.

Note: If the main text is in Bulgarian or Russian, parts in item 4 should be in English. If the main text is in English, they should be in Bulgarian and have to be formatted as in the beginning of the paper.

5. All mathematical signs and other special symbols should be written clearly and legibly so as to avoid ambiguity when read. All formulas, cited in the text, should be numbered on the right.

6. Figures (black and white), made with some of the widespread software, should be integrated in the text.

7. Tables should have numbers and titles above them, centered right.

8. Reference sources cited in the text should be marked by a number in square brackets.

9. Only titles cited in the text should be included in the references, their numbers put in square brackets. The reference items should be arranged in alphabetical order, using the surname of the first author, and written following the standard. If the main text is in Bulgarian or Russian, the titles in Cyrillic come before those in Latin. If the main text is in English, the titles in Latin come before those in Cyrillic. The paper cited should have: for the first author – surname and first name initial; for the second and other authors – first name initial and surname; title of the paper; name of the publishing source; number of volume (in Arabic figures); year; first and last page number of the paper. For a book cited the following must be marked: author(s) – surname and initials, title, city, publishing house, year of publication.

10. **The author(s) and the reviewer, chosen by the Editorial Board, are responsible for the contents of the materials submitted**.

Important for readers, companies and organizations

1. Authors, who are not members of the Union of Scientists - Ruse, should pay for publishing of materials.

2. Advertising and information materials of group members of the Union of Scientists – Ruse are published free of charge.

3. Advertising and information materials of companies and organizations are charged on negotiable (current) prices.

Editorial Board

ISSN 1314-3077