AFRL-RH-AZ-TR-2009-0039

Pictures from Words, Pictures from Text: Constructing Pictorial Representations of Meaning from Text

James Cowie
Department of Psychology

Stephen Helmreich
Physical Sciences Laboratory

Hung Huu Dang
Graduate Student
Department of Electrical Engineering

New Mexico State University
P.O. Box 30001
Las Cruces, New Mexico, 88003-8001

November 2, 2009

Final Report for September 2007 – November 2009

Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Warfighter Readiness Research Division
6030 South Kent St. Mesa, AZ 85212

NOTICES

AFRL-RH-AZ-TR-2009-0039 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signed//
JERRY T. BALL
Lab Contract Monitor

//signed//
HERBERT H. BELL
Technical Advisor

//signed//
DANIEL R. WALKER, Colonel, USAF
Chief, Warfighter Readiness Research Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
11-12-2009

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
09-07-2007 – 11-12-2009

4. TITLE AND SUBTITLE
Pictures from Words, Pictures from Text: Constructing Pictorial Representations of Meaning from Text

5a. CONTRACT NUMBER
FA8650-07-1-6487

5b. GRANT NUMBER
FA8650-07-1-6487

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jim Cowie, Stephen Helmreich, Hung Huu Dang

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
1123AT09

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Computing Research Laboratory
New Mexico State University
Anderson Hall
P.O. Box 30002, Mail stop 3CRL

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Human Effectiveness Directorate
Mesa Research Lab, 6030 S. Kent St., Mesa, AZ 85212

10. SPONSOR/MONITOR’S ACRONYM(S)
711 HPW/RHA

11. SPONSOR/MONITOR’S REPORT NUMBER(S)
AFRL-RH-AZ-TR-2009-0039

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A: Approved for public release; distribution unlimited, (approval given by 88 ABW/PA, 88ABW-2010-0077).

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report describes a design study carried out at the Computing Research Laboratory, New Mexico State University. The goal was to investigate the feasibility of creating a visual interpretation of the messages passing between the crew of an unmanned Aerial Vehicle.
The following tasks were carried out by the research team –

- A general analysis of how images could be used to represent dialogs
- A more specific analysis of dialogs and parse trees produced from experimental UAV dialogs
- Developed a set of icons and conventions to represent mission status and the form of crew member dialogs
- Designed a prototype visualization system – Chat2Pix
- Implemented the prototype using Visual Basic to run on Windows systems.
- Redesigned the prototype to display more structured information
- Implemented and delivered a second version of the visualization system – Chat2Pix2

For the limited domain of UAV crew dialogs it seemed possible to produce visual representations of the message. Such messages could be used to communicate between UAV crew members who do not speak the same language, as well as to visually verify the spoken communication.

15. SUBJECT TERMS
pictorial representation; meaning; visualization; picture; image; icon; language comprehension

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Jerry Ball
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED	29	**19b. TELEPHONE NUMBER** *(include area code)* 480-988-6561

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

Overview

This report describes a design study carried out at the Computing Research Laboratory, New Mexico State University. The goal was to investigate the feasibility of creating a visual interpretation of the messages passing between the crew of an unmanned aerial vehicle (UAV). The following tasks were carried out by the research team:

- A general analysis of how images could be used to represent dialogs.
- A more specific analysis of dialogs and parse trees produced from experimental UAV dialogs.
- Participation in discussions and a simulated mission while visiting researchers at AFRL, Mesa.
- Developed a set of icons and conventions to represent mission status and the form of crew member dialogs.
- Designed a prototype visualization system – Chat2Pix.
- Implemented the prototype using Visual Basic to run on Windows systems.
- Redesigned the prototype to display more structured information. This was based on feedback on the first system from the Mesa AFRL research group.
- Implemented and delivered a second version of the visualization system – Chat2Pix2 (see Appendix B).

For the limited domain of UAV crew dialogs it seemed possible to produce visual representations of the message. Such messages could be used to communicate between UAV crew members who do not speak the same language, as well as to visually verify the spoken communication.

Table of Contents

Background

The Air Force Research Laboratory in Mesa, AZ is developing components for a simulated crew member for unmanned aerial vehicle (UAV) training missions under the management of Dr. Jerry Ball,. Normally a crew of three is needed; a mission controller, a pilot, and a payload specialist. While a mission is in progress these three interact both verbally and through a set of shared interfaces. In order to train for a mission a complete crew of three needs to be present. The AFRL team is working on the development of an artificial crew member capable of interacting in a realistic manner with the other, human, team members. If this is successful it will both reduce the cost of training and also make it possible to "refly" missions with consistent behavior by the simulated crew member.

The crew member is being built using the ACT-R behavioral model (Anderson). The input to this model is language and flight status information and the output similarly commands and requests and appropriate actions by the crew member. In order to handle the language input the AFRL team had developed a parser which produced structured representations of the language. Various other components of the ACT-R simulation are also being built, but in the meantime it seemed possible to develop a non-linguistic representation of the crew dialogs as a way to debug the parser and to demonstrate its operation.

Our task was to investigate the possibility of generating pictures from the structured linguistic input,. These pictures need to communicate essentially the same meaning/interpretation as that suggested by the input. We also needed to develop algorithms for mapping the input, into the pictorial representation. Where practicable, we intended to provide a computational implementation of these algorithms. The investigation results are incorporated in this final report focusing on the salient issues of this mapping and generation process.

Language Processing

Information can be assimilated through any number of paths (senses) – vision, hearing, touch, smell, etc. For human beings, one of the common ways is through a specialized system of auditory symbols that we call language. Over centuries, this system has been provided with a visual form – alphabetic writing. These specialized forms of symbolic communication access and activate the much more complex and multimodal internal meaning of the symbols and their relationships. This process is not well-understood, though study of psycholinguistic has been done in this area.

For example, the word “dog” seems to evoke a standard stereotypical image of a dog, with its associated typical characteristics – fur, four-legs, barks, chases cats – and probably also some personal memories of pet dogs and experiences of dogs, such as getting bitten by a dog. It also invokes links to other words, to pronunciation, and to social and pragmatic information.

Individual words may evoke whole social schemas. For example, “voting” might evoke the entire democratic process of electing officials: party primaries, registering to vote, election day, voting booths, etc. Speakers often rely on evoked meanings to help convey their intentions when speaking. Thus meaning is often inferred and implied rather than directly stated. Communication is underdetermined by language.

In addition, the communicative power of language symbols is increased by the concatenating of symbols to convey more complex relationships between the referents of linguist expressions. Hence, “John kissed Mary” conveys some specific information about the relationship between the persons referred to as “John” and “Mary.” Thus, it appears that the language facility in humans consists of both a large memorized vocabulary store of concepts and a fairly simple set of rules as how these concepts are to be combined to produce more complex meanings.

The first attempts to communicate visually at a distance most likely involved pictures of the acts or objects to be communicated: petroglyphs, cave paintings, etc. Over time, these pictures became more stylized and abstract. These concept-pictures then became attached to particular words, producing ideographic writing systems such as Chinese. It took some time before there was an attempt to use these pictures to represent the sounds of the word, rather than the concept behind them. Hieroglyphics represent this stage. These were then modified to represent syllables, and finally the alphabet evolved with each symbol corresponding roughly to one phoneme in the language [23]. This elegant and compact solution for visually representing speech has one drawback. Since it represents the sounds of a language, one must be able to speak the language in order to understand the written communication. This has proved a difficulty for logicians, translators, and computational linguists.

Bishop John Wilkins (1614-1672), for example, concocted a Universal Language [20] in which the letter and sound symbols that make up a word in his language are connected to concepts which when combined produce the concept designated by the word. For example, the word for “father” is “Cobara,” consisting of the syllable “co” which signifies an economical relationship, “b” and “a” for the first two subdivisions of this relationship, namely consanguinity and direct ascendant, so that “coba” signifies parent. “Ra” then signifies “male”, thus giving “cobara” as “father” [19]. More contemporary artificial languages include Loglan, based on predicate logic and so universal in the same way that mathematical symbolism is universal [21].

Blissymbolics, invented or developed by Charles K. Bliss (1897-1985), represents an attempt to create a new ideographic, non-linguistic form of communication. Using stylized symbols to represent concepts (for example, a stylized clock face to represent time and tense, a stick figure to represent a man, the same figure with a “skirt” to represent a woman), these symbols can be combined following a rudimentary grammar to produce more complex concepts. For example, a subscript 1, 2, or 3 can be placed beside the male stick figure to represent “I (male)”, “you (male)” and “he” [22].

Computational linguists have attempted over the years to capture the meaning of language symbols and texts for various purposes: automatic translation, natural language understanding, information retrieval, etc. However, all such attempts have at some level simply reformulated language symbols in terms of another (unambiguous) symbol system. Most recently the interlingual symbol system has been enhanced by an ontology, which (supposedly) adds real world information to the interlingual symbols [2] [9]. Actually, this process only adds information about the relationship between the symbols, since they themselves remain ungrounded.

In this project, we explored the possibility of extending the amount of information conveyed by a computational system by capturing the meaning of words and texts through pictorial means. Some initial work has been done in using pictures to communicate across languages [3]. Using pictures for all types of communication raises difficult questions of how to encode pictorially such things as abstract nouns (“sincerity”, “administration”) and complex constructions (relative or subordinate clauses, complex tenses).

Here, we focus on a narrow domain of communication that is highly constrained and closely connected to the observable or measurable world – that of the communication between the crew members responsible for a UAV mission. Much of this communication is in an abbreviated speech, accessing schemas that are shared by the participants, but likely somewhat opaque to others [1]. We expect that focusing on this limited sublanguage will allow us to make significant advances in pictorial representation while avoiding the problems of dealing with the full complexity of standard speech. We hope as well that it will prove useful in providing insights into the development of the language processing components of the ACT-R simulation.

Key Problems

Initial work involved discussions by the NMSU team into what exactly was feasible in terms of producing a useful, dynamic pictorial representation. These discussions were supported by inspection of training session dialogs and the parse structures derived from them by the AFRL parser. The research team also visited the AFRL site in Mesa, AZ and took part in a simulated flight as crew members. This exercise was very useful to understanding the sources of the dialogs provided us to AFRL and also to allow us to see the environment and screens used by crew members. We were also provided with the training software used to train research participants to take part in a mission.

We also struggled throughout the project with deciding what exactly would be the purpose of the final pictorial representations. These discussions continued during our site visit and the delivery of the two prototype systems (see "Purpose of conversation visualization system").

The following issues are important to the project-

- Appropriate representations
- Representation of non-prototypical classes
- Iconicity versus Arbitrariness
- Compositionality
- Schemas

Appropriate representations

Just as language underdetermines communication, pictures both underdetermine and overdetermine communication. Words are often understood as referring to abstract classes of objects and relationships, and having interest only as a symbol for that class. Pictures (except in rare cases) do not have that kind of symbolic relationship to reality. It is not always clear what a picture is intended to be a picture of. That makes the picture ambiguous in a different way from language symbols. In addition, pictures by necessity contain details that are not specified for the class they are intended to point to. For example, a picture of an airplane, even if stylized and iconic, will probably show details that are not found on all airplanes – a specific number of engines, a particular body design, etc. While we expected these problems to be somewhat alleviated by the use of iconic images (less detail) that are reasonably familiar to the communication participants (standard images), this was expected to still be an issue at some level.

Representation of non-prototypical classes

Research [4] [5] has shown that people organize their semantic representations of reality by means of prototypical instances of a class. In addition, these prototypes are most accessible and communicable at a certain level of abstractness. For example, the concept of "chair" is at that medium level of abstractness, which allows common understanding of a prototype – what properties are characteristic of chairs, what exceptions there may be, what a protoypical chair looks like. It is this last property of prototypes that makes them particularly amenable to pictorial representation. Concepts that are more abstract, like "furniture", do not lend themselves to such a representation. There are so many different kinds of furniture that any pictorial representation of a piece of furniture would most likely be taken to represent its prototype rather than the abstraction. Similarly with concepts that are more specific, e.g., "Queen Anne chair". How to represent these overly abstract and overly specific concepts is a challenge. One possible solution was to represent the overly abstract class by a conglomeration of subclass prototypes. Overly specific concepts might be represented by a modifying relationship to a clear prototype. That led us to the next issue.

Iconicity versus Arbitrariness

Though it has been argued since the time of the early Greeks [19, pp.-17ff.], one of the key characteristics of language symbolism is its arbitrariness, as Saussure pointed out [17]. Recently there has been a resurgence of interest in the iconicity of language, particularly with respect to gestures [24], though without essentially challenging Saussure's claim. One goal of a pictorial representation is to communicate less arbitrarily through more iconic representations, those that can be easily recognized even by those untrained in the symbolism. Sign language, for instance, is regarded as perhaps more iconic than spoken language.

However, it is clear that there must be a tradeoff. Certain concepts, as indicated above, will not have such commonly-understood iconic representations and so will be at least partially arbitrary. Iconic representations by their very nature are somewhat abstract representations and also somewhat arbitrary. For example, once a red slash through a picture is understood as negation, it can be used in many icons quite comprehensibly. But the slash itself is only partially iconic.

Compositionality

Even if one has iconic representations of individual lexical items, we needed to decide how to identify their relationships in context and to display those relationships meaningfully. The goal was to represent modified concepts by means of a composition of pictures. Ideally, we wanted to have a basic store of images or pictures, representing basic classes. These would then be combined to represent intersection of classes (adjectival modification) or other relationships between the classes. For example, the modified concepts "asphalt runway" and "concrete runway" could be represented by an iconic runway colored black, in the first case, and white, in the second. These separate images would be fine if they were the only kinds of runway that were relevant. However, there are grass runways, dirt runways, gravel runways, etc. To represent all of these possibilities, it would be better to have one picture for a generic runway, and other pictures for various types of material and then combine these two images to produce one that signifies a runway composed of that material. But that requires a representation of a material or substance, which is difficult to represent iconically, as well as a representation of the relationship "composed of".

Schemas

Much information conveyed by language is not conveyed directly. Rather it is communicated through pragmatic devices that are in the immediate context of communication (e.g., deictics, pronouns) or through shared context (shared knowledge) evoked by the communication. Schemas are in the latter category and represent shared understandings of relatively complex, but frequently recurring scenarios – eating out at a restaurant, taking an academic class, flying a mission, etc. [6]. When the scenario is evoked or activated, the entire scenario is available for reference. It appears that these scenarios are often evoked in the course of a mission by means of well-defined language schemas that reference parts of these scenarios in short-hand form. Thus the communication participants share not only semantic scenarios or schemas, but linguistic schemas that reference these semantic scenarios in a standard form. As part of this project, it was necessary to identify and understand these schemas, both in terms of the linguistic formalism used and of the underlying semantic frame.

Purpose of conversation visualization system

The research group struggled for some time with the question - “Why should we want to convert text to images?”. The initial goal was to validate the output of the parser. Even here, however, there were issues that needed to be understood. For example “Who would be using the images to understand the parser output?” Initially, the answer to this question was the developers of the picture system themselves, who regularly produced odd looking sets of images as they had not fully understood the structures produced by the parser. In the long term the users of the system would be the parser developers, who would be given an alternative, independently developed, view to the annotated parse trees they normally produced from their analysis (See Appendix B. Sample Corpus and Parse Trees).

As the project developed, other potential applications of a more general type emerged that might be supported by the dialog visualization system. Researchers investigating the system might be able to use a dynamic series of images to understand the dialog better. One aspect of this would be the possibility of showing disparities between the individual mission views of each of the participants, the “total” team view, and the actual state of the system. The final version of the system only shows a “total” status, but individual differences might be added. In terms of training and exercise reviews, the visualization system has the possibility of helping team members to understand what happened at critical points in an exercise, rather than just replaying the actual mission.

The problem with supporting any visualization application is that mapping to any image, or series of images requires the use, and understanding of conventions. The conventions considered by the team were: the meaning of icons, how space is represented, how time is represented, how causality is represented, how intent is represented, and how beliefs are represented.

The possibilities are in fact very large, however the limited domain of UAV dialog reduces the number of possibilities we need to visualize. In addition there are images used on the control displays, maps, gauges etc., used by the team, which already have a meaning for team members. This is due to the crew having a shared context, through their mission goals and training. Other information that needs displayed such as reference to places, events, and things were less easy to define. In the end the developers decided on canonical images for various objects, and then added labels, if needed, to indicate which particular object was being considered. We also decided to use coloring conventions on boxes placed around images to indicate the pragmatics of a dialog – thus the images related to a question are surrounded by a green box, arrows are used to indicate the producer and recipient of each statement. Given the limited screen real-estate available to the developers we decided not to handle shared viewpoints or multiple viewpoints. The system shows one single pictorial representation of the situation (see Appendix A). The prototype system shows a single snapshot of the mission, time is represented by a progression through the series of messages, which are displayed at the top part of the prototype interface to support its parser debugging function. One possibility here would be to have a separate display using the *sparklines* suggested by Edward Tufte [18] to show successive snapshots of system and/or belief status. This was not possible within the time frame of this project.

Schedule of Work

In our initial proposal, a significant part of the work would be spent on the development of a corpus of conversations and then analyzing this corpus to identify the common concepts and relations present in the corpus either explicitly or implicitly. We then proposed to understand the schemas involved in the dialogs and their interpretation in context. When the project started we were given both a corpus and a set of parsed structures from the corpus. At this point we could focus on the schemas and their pictorial representations.

Technical Discussion

We discuss each of the tasks in more detail.

Developing a concept list

We developed a list of the most common concepts (and their linguistic referents). This was done using a word count, program and by examination of the dialogs provided to the research team. The result was a list of lexical items (and the concepts they represent) that need pictorial representation. This is a kind of lexical-semantic analysis.

We also developed a list of common relations represented in the corpus. These relations can be expressed lexically by means of verbs or nominalizations. They are also be represented by prepositions which relate lexical items. In the context of UAV dialogs these appear to be fairly unambiguous.

The result of this stage of the research was a set of concepts and relations that require pictorial representation.

Pictorial Representation

This stage of the research involved developing iconic representations for the set of concepts and relations identified in the previous stage. There have been a number of such semiotic systems proposed and implemented. Traffic signs, rebuses, hieroglyphic writing, universal technical pictographs (mathematical equations, musical representation, orchesography, chemical symbolism, etc.) are only some of these systems. The control systems used for UAV missions also provide a set of potential images. Our interaction with the AFRL research team helped us locate standard iconic representation for concepts that are already in current use or well understood. (3) We then searched the web for potential representations of the key concepts and relations.

The result of this stage of the research is a set of pictorial representations connected to the sets of concepts and relations identified in the previous stage.

Computational Implementation

The research focused on implementing a program that takes as input UAV dialogs and outputs a pictorial representation of that communication. This implementation is a proof-of-concept rather than a full-fledged software product. Although it is not a broad coverage system, it has been tested with all the dialogs provided by AFRL and it appears to work accurately.

The system is implemented in Visual Basic and uses as input the original dialogs and the parses developed by the AFRL team to analyze text to concepts, relations, and schemas; the program connects these items to pictorial representations; and combines these individual pictorial representations into a representation of the text.

Research Team

The research team consisted of the Principal Investigator, Dr. Stephen Helmreich; a psychology faculty member also trained in computer science, Dr. James Cowie; and a graduate student from electrical engineering, (now) Dr. Hung Huu Dang.

The project group met on weekly to develop ideas and produce mockups for the prototype system. Helmreich also served as chief administrator for the project and primary supervisor of the graduate student. He provided the linguistic expertise for the project. Cowie was responsible for designing and overall implementation of the computational implementations for the project. Dang was the main implementer of the system.

Regular consultations were held with the funders to exchange information, present results, and guide the direction of the research project. One site visit in each direction was held during the course of the project.

Appendix A: Chat2Pix USER'S GUIDE

I. Software overview:

- *Chat2Pix* provides a pictorial interpretation of a text chat by displaying a series of pictures representing sentences in the chat. Each picture consists of images that are structurally organized.
- *Input*: A text file that contains a pre-processed conversation (chat) between the AVO, PLO, and DEMPC during a flight. The text chat has been pre-processed by an external, separated interpreter to be free of filler words, and put in a format described by pre-defined keywords.

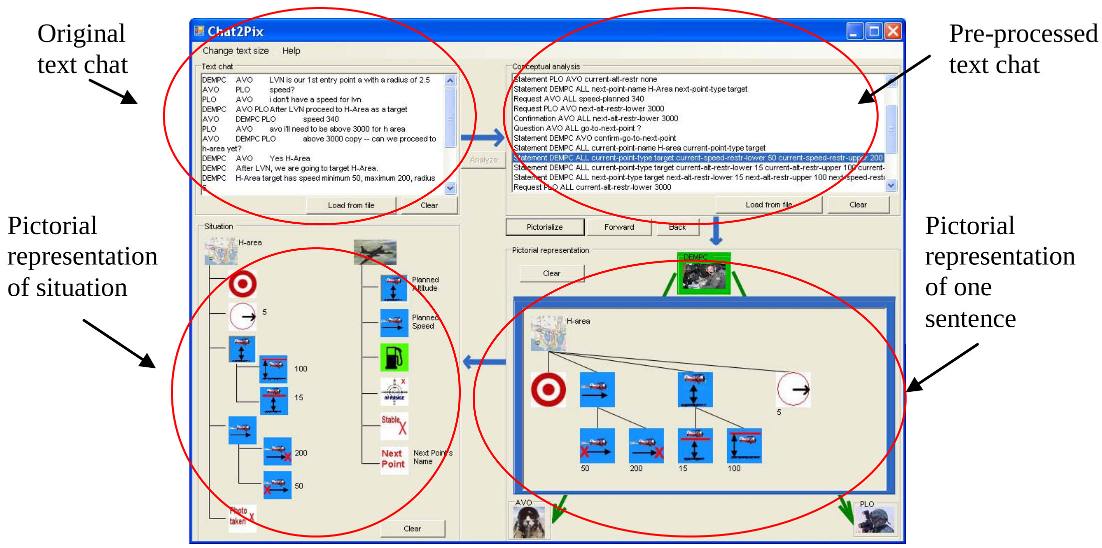


Figure 1. *Chat2Pix* main screen

- *Output*: A series of pictures, each represents one complete sentence spoken by one party in the conversation. Each picture consists of images that are organized structurally. Each picture is drawn with the following goals:
    - Imply the sentence's type by using different colors for the picture's frame: blue for a statement, red for a question, yellow for a request, green for a confirmation.
    - Indicate the speaker and his/her targeted audience by showing the speaker's representative image in a green frame, and green arrows pointing from the speaker to the targeted audience.
    - Interpret the meaning of the sentence pictorially by displaying images that represent the objects involved in the sentence.
    - Illustrate the relations between the objects.
- As seen in Figure 1, *Chat2Pix*'s main screen consists of four frames:
    - The *Text chat* frame contains the original text chat.
    - The *Conceptual analysis* frame contains the pre-processed text chat.

- The *Pictorial representation* frame contains the pictorial interpretation of the highlighted sentence in the *Conceptual analysis* frame.
- The *Situation* frame contains the pictorial representation of the overall situation, which displays information regarding the current waypoint, the airplane, and the mission.

II. How to run

- *Step 1*: Double click on file *pictorialization.exe* to launch the software. The main screen will appear:

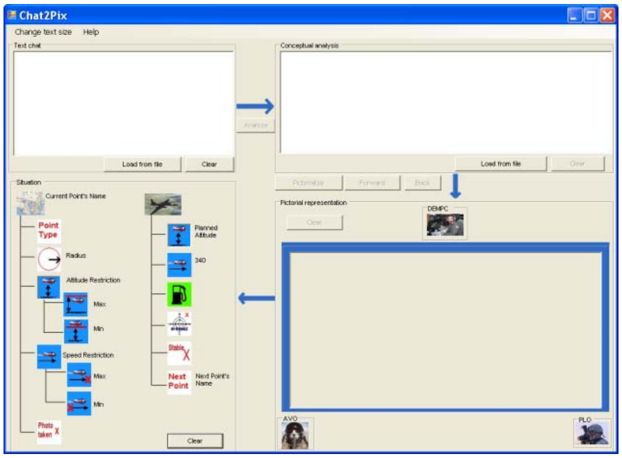


Figure 2. *Chat2Pix* main screen at the beginning of a session

- *Step 2*: Click on *Load from file* in the *Text chat* frame to load an original text chat from a text file:

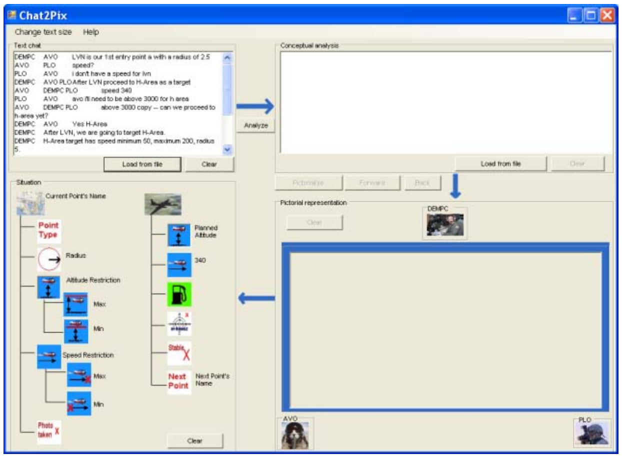


Figure 3. Main screen after the loading of an original text chat

- *Step 3*: Click the *Analyze* button to load the pre-processed text chat from a text file. This text file must be in the same folder with the file loaded in the previous step. Its name must be the name of the file in the previous step concatenated by *"-analyzed"*. For example, if the original text chat file loaded in the previous step is *textchat.txt* then the pre-processed text chat file must be *textchat-analyzed.txt*. *Chat2pix* does not handle the creation of *textchat-analyzed.txt* from *textchat.txt*, which has to be done by a separate text processor.

- Note: The pre-processed text chat can be loaded without the original text chat by simply bypassing *Step 2* and clicking on the *Load from file* button in the *Conceptual analysis* frame.
-

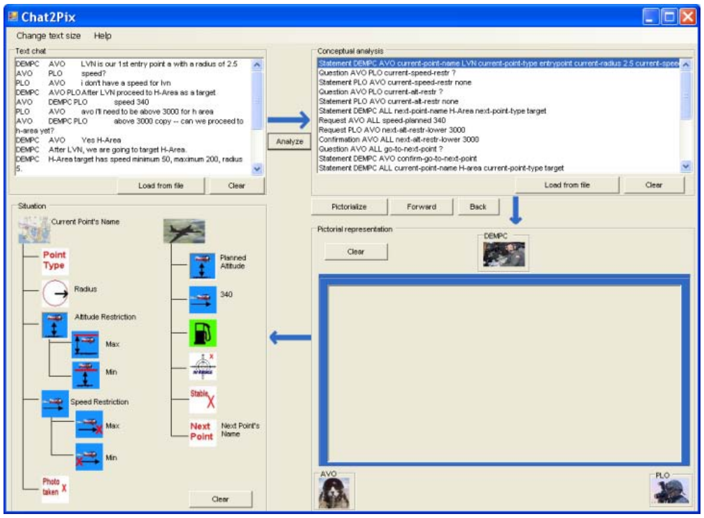


Figure 4. Main screen after the loading of a pre-processed text chat

- *Step 4*: Click on the *Pictorialize* button to make *Chat2Pix* create a pictorial representation of the highlighted sentence from the *Conceptual analysis* frame. The pictorial representation is displayed in the *Pictorial representation* frame.
-

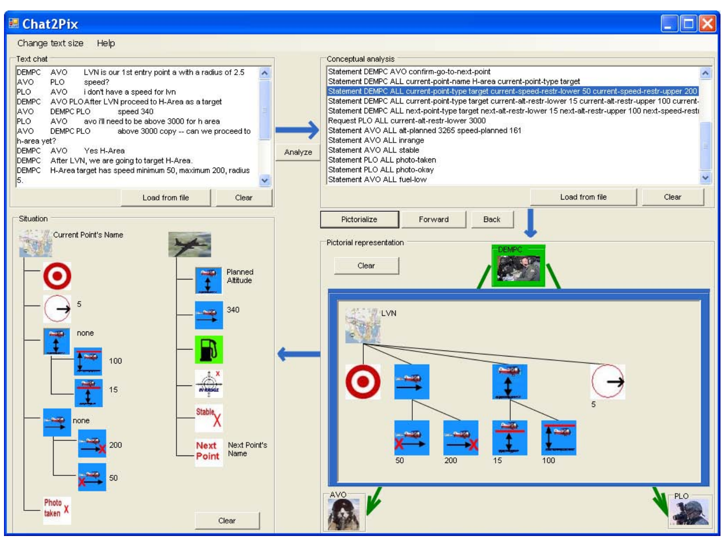


Figure 5. Pictorial representation of the highlighted sentence from the *Conceptual analysis* frame

Use the *Forward* or *Back* buttons to browse the sentences in the *Conceptual analysis* frame. The *Pictorial representation* frame will automatically display the pictorial version of the highlighted sentence.

- The picture in the *Situation* frame will constantly update itself based upon the information given by the sentences that have been browsed.

III. Interpret the pictures

The picture in the *Pictorial representation* frame illustrates the highlighted sentence in the *Conceptual analysis* frame:

- The color of the picture frame indicates the type of the sentence: *blue* for a statement, *red* for a question, *yellow* for a request, *green* for a confirmation.
- Actions, such as *"go-to-next-point"*, are illustrated by animated GIF images.
- Objects, such as *"speed-restriction"*, are illustrated by static JPEG images.
- Images are connected to show the relations between them

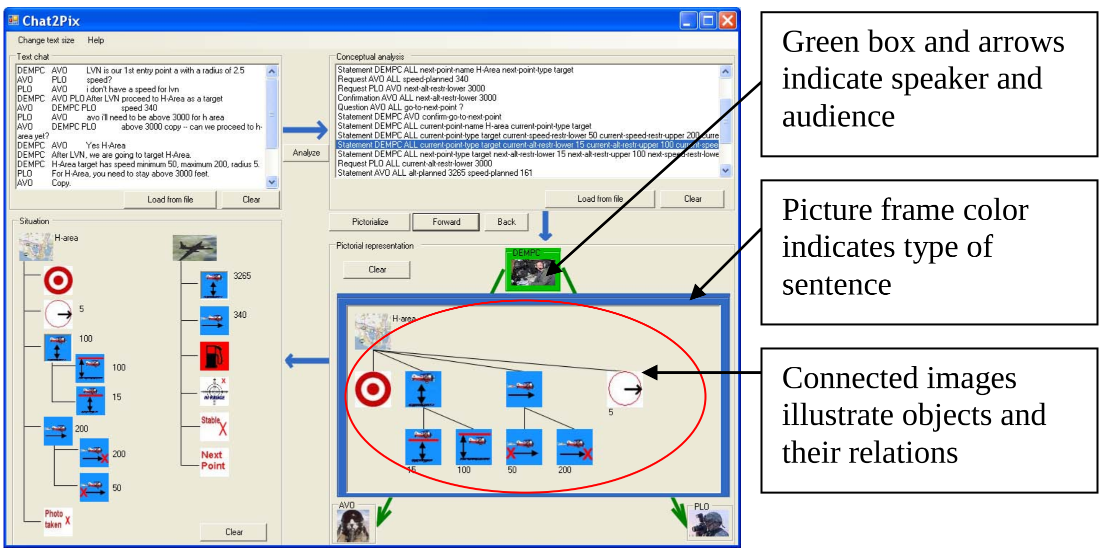


Figure 5. Pictorial representation

The picture in the *Situation* frame illustrates the overall situation:

- The left side of the picture shows the current waypoint with its properties: name, speed restrictions, altitude restrictions….
- The right side of the picture shows the airplane and its mission-related information: fuel level, next waypoint's name, whether the plane is within range of the target for taking photo, whether the plane is stable enough for taking photo…

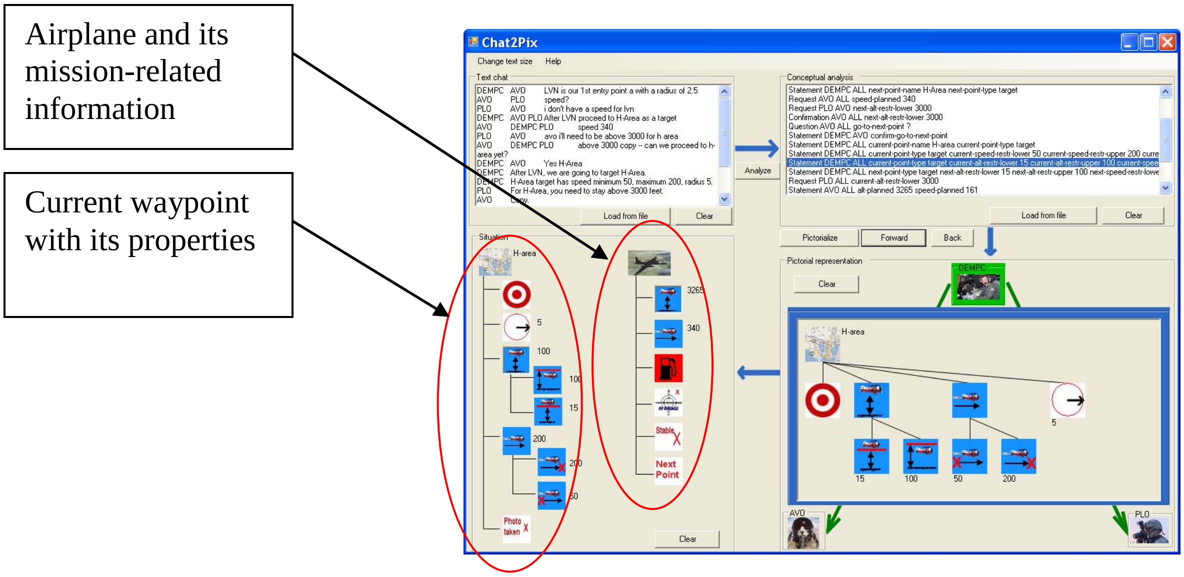


Figure 6. Situation frame

IV. Format of the input text files

The original text chat can be in an arbitrary format, for example:

DEMPC AVO LVN is our 1st entry point a with a radius of 2.5
AVO PLO speed?
PLO AVO i don't have a speed for lvn
DEMPC AVO PLO After LVN proceed to H-Area as a target
AVO DEMPC PLO speed 340
PLO AVO avo i'll need to be above 3000 for h area
AVO DEMPC PLO above 3000 copy -- can we proceed to h-area yet?
DEMPC AVO Yes H-Area
DEMPC After LVN, we are going to target H-Area.
DEMPC H-Area target has speed minimum 50, maximum 200, radius 5.
PLO For H-Area, you need to stay above 3000 feet.
AVO Copy.
AVO Flight level for H-Area will be 3265, airspeed will be 161.
PLO Copy.
AVO Within range for H-Area.
AVO Stabilized.
PLO We have a good photo for H-Area.

The pre-processed text chat must contain sentences in the following format:

- 1st word indicates sentence type: can be one of *Statement*, *Question*, *Request*, or *Confirmation.*
- 2nd word indicates the speaker: can be one of *AVO*, *PLO*, or *DEMPC*
- 3rd word indicates the targeted-audience: can be one of *AVO*, *PLO*, *DEMPC* or *ALL*
- The rest of the sentence must be a series of pairs of *Object-Value* such as *speed-planned 340*, or *Actions* such as *confirm-go-to-next-point*.

For example, the processed text chat of the original chat above is as follows:

Statement DEMPC AVO current-point-name LVN current-point-type entrypoint current-radius 2.5
Question AVO PLO current-speed-restr ?
Statement PLO AVO current-speed-restr none
Question AVO PLO current-alt-restr ?
Statement PLO AVO current-alt-restr none
Statement DEMPC ALL next-point-name H-Area next-point-type target
Request AVO ALL speed-planned 340

Request PLO AVO next-alt-restr-lower 3000
Confirmation AVO ALL next-alt-restr-lower 3000
Question AVO ALL go-to-next-point ?
Statement DEMPC AVO confirm-go-to-next-point
Statement DEMPC ALL current-point-name H-area current-point-type target
Statement DEMPC ALL current-point-type target current-speed-restr-lower 50 current-speed-restr-upper 200 current-radius 5
Request PLO ALL current-alt-restr-lower 3000
Statement AVO ALL alt-planned 3265 speed-planned 161
Statement AVO ALL inrange
Statement AVO ALL stable
Statement PLO ALL photo-taken
Statement PLO ALL photo-okay

V. Extra features

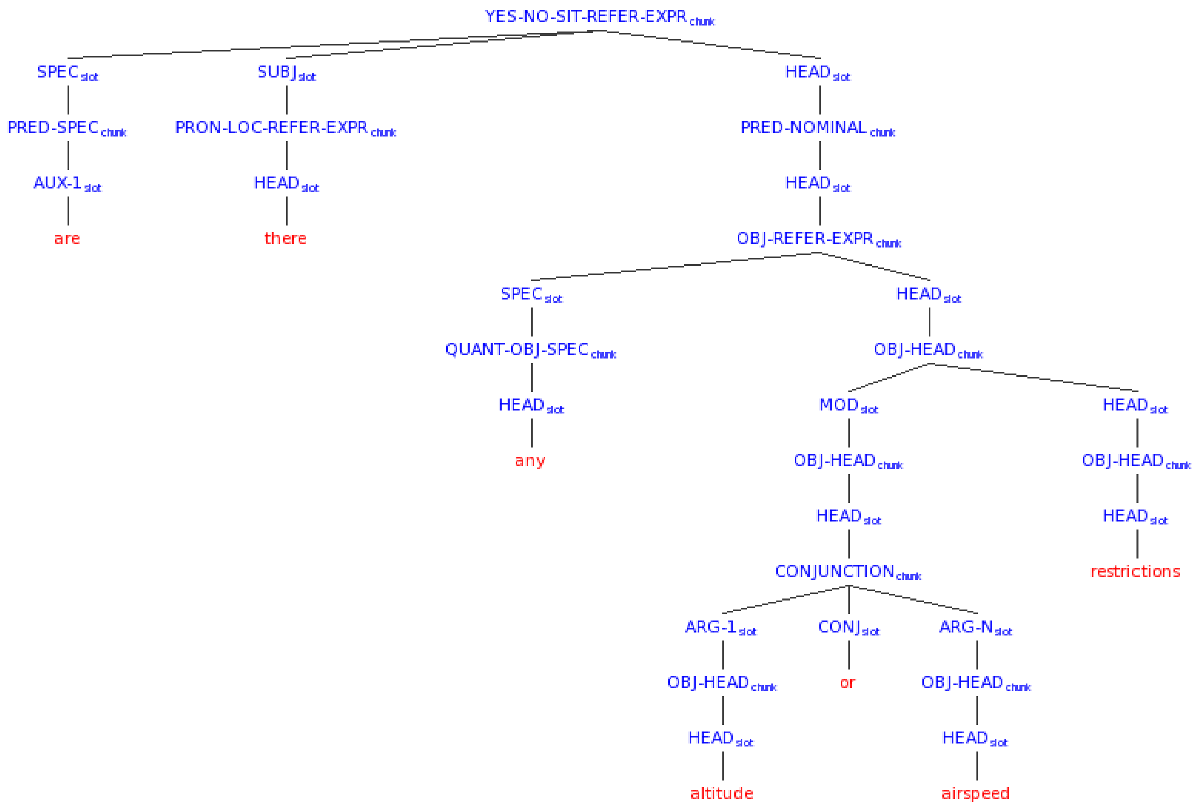
- Change font size: click on menu *Change text size* to change the size of the text in *Chat2Pix*.
- Image tour: click on menu *Help* and then *Image Tour* to see a list of the images used in *Chat2Pix.*

Appendix B: Sample Corpus and Parse Trees

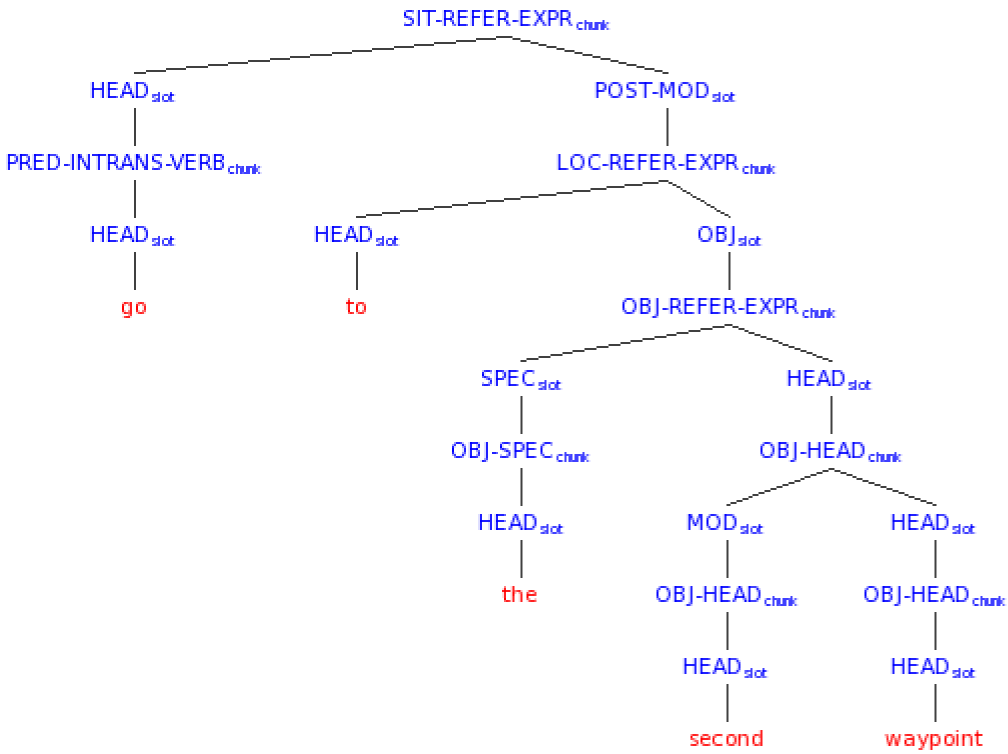
This appendix contains an example of a dialog and the types of annotated parse trees produced by the AFRL parser from inputs of this type.

0:09	PLO	0	AVO, can I please be about 3000 feet or higher, please?
0:28	PLO	0	Cancel.
0:28	PLO	1	Cancel.
0:36	**AVO**	**0**	**Do I need to change my airspeed?**
0:43	**AVO**	**0**	**I mean my altitude.**
0:47	DEMPC	0	Once I get the first, uh, sequence figured out, I'll let you know.
2:20	DEMPC	0	First waypoint LVN is an, uh, ROZ access point.
2:20	DEMPC	1	There is no flight restrictions, but the, uh, radius is, uh, 2.5 miles.
3:43	DEMPC	0	I'm pretty sure you can take a bearing towards HAREA now.
3:43	DEMPC	1	It looks like you're in within the 2.5 required for this entry point.
4:02	PLO	0	AVO, can I please, uh, keep, uh, altitude over 3000 feet for this picture, please?
4:14	PLO	0	Can you give me a range?
4:19	DEMPC	0	The next target HAREA has a range of 5 miles.
4:26	PLO	0	Copy.
4:31	**AVO**	**0**	**Was that above 3000?**
4:41	PLO	0	Yes, please.
4:50	PLO	0	Can you also keep this current airspeed?
4:57	**AVO**	**0**	**OK.**
5:10	DEMPC	0	Next waypoint is HAREA.
5:10	DEMPC	1	There is no altitude restriction, but the speed restriction is between 50 and 200.
6:32	PLO	0	I have a picture, you can go.

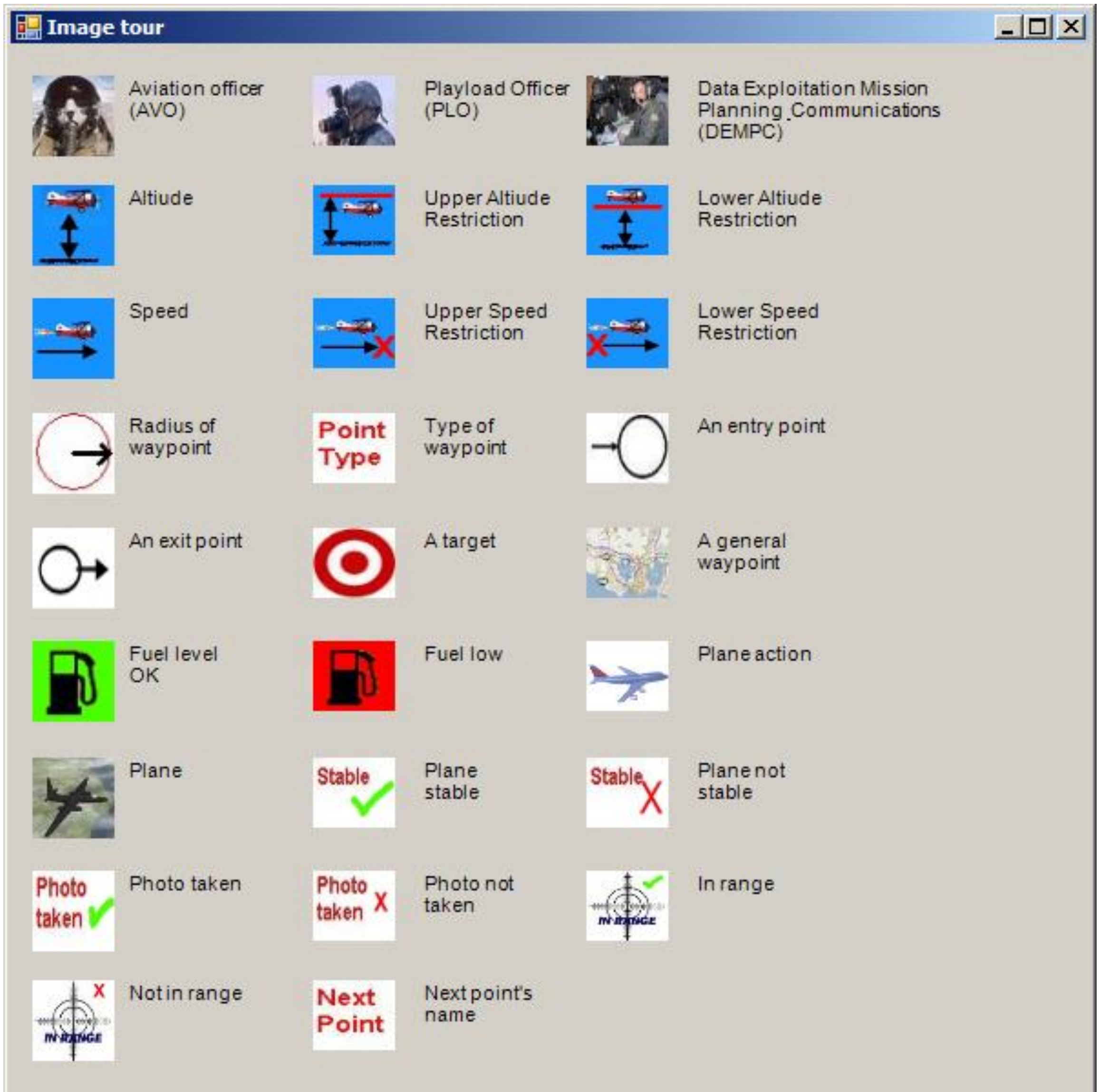
Parse tree for: “Are there any altitude or airspeed restrictions?”



Parse tree for: “Go to the second waypoint”



Appendix C: Image Tour



References

[1] Ball, Jerry. 2004. A Cognitively Plausible Model of Language Comprehension. In *Proceedings of the Behavior Representation in Modeling and Simulation (BRIMS) Conference*. Scottsdale, AZ.

[2] Mahesh, Kavi and Sergei Nirenburg. 1995. A Situated Ontology for Practical NLP. *In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95): Workshop on Basic Ontological Issues in Knowledge Sharing,* 19-20 August 1995. Montreal, Quebec, Canada: IJCAII.

[3] Mihalcea, Rada and Chee Wee (Ben) Leong. 2006. Toward Communicating Simply Sentences Using Pictorial Representations. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas,* August 8-12, 2006 (pp. 119-127). Cambridge, MA.

[4] Lakoff, George. 1987. *Women, Fire and Dangerous Things: What categories reveal about the mind*. Chicago, IL: University of Chicago Press.

[5] Rosch, Eleanor. 1978. Principles of Categorization. In E. Rosch and B. Lloyd, Eds., *Cognition and Categorization.* Hillsdale, NJ: LEA.

[6] Schank, Roger C., and Robert P. Abelson. 1977. *Scripts, Plans, Goals and Understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates.

[7] Foltz, Peter W., Melanie J. Martin, Ahmed Abdelali, Mark Rosenstein, and Rob Oberbreckling. 2006. Automated Team Discourse Modeling: Test of Performance and Generalization. *Poster presented at the 28th Annual Conference of the Cognitive Science Society* (pp. 1317-1322), 26-29 July 2006. Vancouver, BC, Canada: CSJ.

[8] Ball, Jerry. 2003. *Double R. Grammar*. http://www.DoubleRTheory.com/DoubleRGrammar.pdf

[9] Mahesh, Kavi and Sergei Nirenburg. 1995. Semantic Classification for Practical Natural Language Processing. In *Proceedings of the Sixth American Society for Information Science Special Interest Group / Classification Research (ASIS SIG/CR) Classification Research Workshop* (pp. 79-94) 8 October 1995. Chicago, IL: ASIS.

[10] Borman, A., Rada Mihalcea, and P. Tarau. 2005. Picnet: Augmenting semantic resources with pictorial representations. In *Proceedings of the American Association for Artificial Intelligence spring symposium: Knowledge collection from volunteer contributors* (pp.1-7). Menlo Park, CA: AAAI Press.

[11] Miller, George. 1995. Wordnet: A lexical database. *Communication of the ACM, 38*(11):39-41.

[12] Nirenburg, Sergei and Victor Raskin. 2004. *Ontological Semantics*. Cambridge, MA: MIT Press.

[13] Farwell, David, Stephen Helmreich, Florence Reeder, Keith Miller, Bonnie Dorr, Nizar Habash, Eduard Hovy, Lori Levin, Teruko Mitamura, Owen Rambow, and Advaith Siddharthan. 2004. Interlingual Annotation of Multilingual Text Corpus. In *Proceedings of the Human Language Technologies/Annual Conference of the North American Chapter of the Association for Computational Linguistics Workshop: New Frontiers in Corpus Annotation, 6 May 2004* (pp. 55-62). Boston, MA: NAACL.

[14] Helmreich, Stephen and David Farwell. 1998. Translation Differences and Pragmatics-Based MT. J*ournal of Machine Translation 13*(1):17-39.

[15] Cowie, Jim, Sergei Nirenburg, Tanya Korelsky, Stephen Helmreich, Benoit Lavoie, Valeriy Sibirtsev, Stephen Beale, and Ted Caldwell. 2004. MOQA: Meaning-Oriented Question-Answering. In *Proceedings of the Computer-Assisted Information Retrieval (Recherche d'Information et ses Applications) - RIAO 2004, 7th International Conference, University of Avignon, France, April 26-28, 2004.* Vaucluse, France: Centre de Hautes Etudes Internationales d'Informatique Documentaire (CID), Institut de Recherche en Informatique de Toulouse (IRIT).

[16] Cowie, Jim, Sergei Nirenburg, and Hugo Molina-Salgado. 2000. Generating Personal Profiles. In *Proceedings of The International Conference On Machine Translation and Multilingual Applications in the New Millennium (MT 2000).* Exeter, United Kingdom, 20-22 November 2000. The British Computer Society.

[17] Saussure, Ferdinand de. 1959. *Course in General Linguistics*. New York, NY: McGraw-Hill.

[18] Tufte, Edward R. 2006. *Beautiful Evidence*. Cheshire, CT: Graphics Press.

[19] Robins, R. H. 1979. *A Short History of Linguistics.* London: Longman.

[20] Wilkins, John. 1688. *An Essay towards a Real Character and a Philosophical Language. [With An alphabetical dictionary, wherein all English words ... are either referred to their places in the Philosophical Tables, or explained by such words as are in those Tables]*. London.

[21] Brown, James Cooke. 1960. Loglan. *Scientific American 202*(6), 53-63.

[22] Blissymbolics Communication International. 2004. *The fundamental rules of Blissymbolics : creating new Blissymbolics characters and vocabulary.* Retrieved from http://www.blissymbolics.org.

[23] Fromkin, Victoria, Robert Rodman, and Nina Hyams. 2007. *An Introduction to Language*, 8th Ed. Boston, MA: Thomson-Wadsworth.

[24] Simone, Raffaele. 1995. *Iconicity in Language*. Amsterdam, Netherlands: John Benjamins.