



A Flexible Dialogue System for Enhancing Web Usability

Marta Gatus and Meritxell González



Technical University of Catalonia (UPC)

Software Department

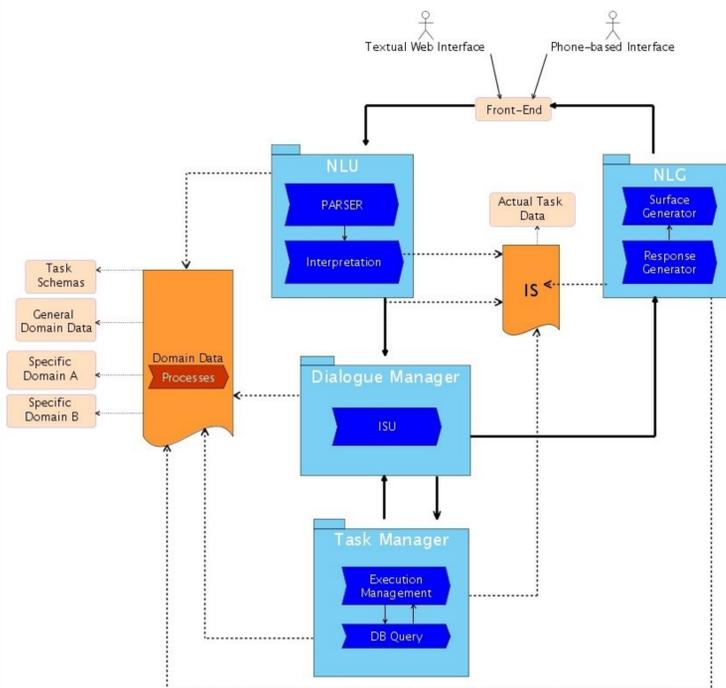
TALP Research Center

The Dialogue System

Functionalities

- ▶ The system guides the user accessing the web
 - ▶ It supports friendly dialogues
 - ▶ The user and the system can take the initiative
- ▶ The system can be adapted
 - ▶ To the web contents
 - ▶ Different domains and presentations
 - ▶ To the users
 - ▶ Different languages, skills, preferences
- ▶ Dialogues strategies can be dynamically adapted

Architecture



The dialogue system consists of separate and interrelated modules

- ▶ The Natural Language Understanding (NLU)
 - ▶ It process the user intervention
 - ▶ It uses grammars and lexicon
- ▶ The Natural Language Generation (NLG)
 - ▶ It produces the system intervention
 - ▶ It uses a syntactic-semantic taxonomy
- ▶ The Dialogue Manager
 - ▶ It controls the communication
 - ▶ It follows the Information State (IS) theory, based on a rich representation of the dialogue state
 - ▶ It includes an adaptive module to automatically adapt the dialogue strategy
- ▶ The Task Manager
 - ▶ It controls the access to the web service
 - ▶ It uses general and reusable knowledge bases to represent the knowledge related to the web services

The Evaluation

Goals

- ▶ Evaluate the performance of the system when accessing different types of web services
 - ▶ Transactional
 - ▶ A service that states a date for large objects collection
 - ▶ Informational
 - ▶ A service that gives information about the cultural events in the city
- ▶ To compare the performance when using different dialogue strategies
 - ▶ Always directed system's messages
 - ▶ Always open system's messages
 - ▶ Adapting dynamically the system's responses
- ▶ To measure the performance improvements when the general linguistic resources were adapted to the web services by human experts

Table 1: Questions related to individual interaction

Average punctuation for the 10 Questions	(0..5)
1. Overall impression is good	3,58
2. The system provided the desired information	3,59
3. You feel understood by the system	3,32
4. You new what the system expected from you	3,74
5. You percived the dialogue as pleasant	3,68
6. You found the system help was appropriate	3,83
7. You were able to understand the system	4,09
8. You found the system is useful	3,64
9. In the future, you would use the system again	3,42
10. You would prefer to access this system by hone	2,91

Metrics and methodology

- ▶ The system was accessible in Spanish through a web site <http://www.lsi.upc.es/~nlp/disi/>
- ▶ Two experiments were carried out
 - ▶ 30 volunteers participated in the first experiment (Test 1) and 31 in the second (Test 2)
 - ▶ In the second experiment the adaptive module was incorporated and the linguistic resources were improved.
- ▶ Following PARADISE four performance metrics were used:
 - ▶ Task success
 - ▶ The user gives all information needed to perform the service task
 - ▶ Dialogue quality
 - ▶ System's errors and appropriateness of system's messages
 - ▶ Dialogue efficiency
 - ▶ Time and number of turns
 - ▶ System usability
 - ▶ A questionnaire about the user satisfaction

Results

- ▶ Users overall impression of the system is good (Question 1 in Table 1)
- ▶ Linguistic resources adapted by experts to the web services result in better performance (as shown in Table 2 the systems errors are fewer and user satisfaction higher)
- ▶ Open system's messages improve task success but are more ambiguous
- ▶ Directed system's messages reduce the system's errors
- ▶ The number of ambiguous messages for the informational service is higher than for the transactional one

Table 2: Test 1 versus Test 2

Performance Features	Test 1	Test 2	P-value
Task success(%)	82%	83,3%	0,783
System Errors	2,37	1,19	0,003
Time (seconds)	124	102	0,222
# user turns	6,42	6,10	0,469
User Satisfaction (1-5)	3,28	3,655	0,004
Ambiguous messages	0,297	0,154	0,085

Table 3: Directed versus automatically adapted messages

Paired-t Test - D vs. A	Directed	Adapt.	P-value
Task success(%)	82%	80%	0,766
System Errors	0,749	1,273	0,038
Time (seconds)	73,5	103,8	0,037
# user turns	6,192	6,198	0,990
User Satisfaction Mean (1..5)	3,720	3,604	0,383
Ambiguous messages	0,1434	0,1320	0,872

Table 4: Open versus automatically adapted messages

Paired-t Test - Open vs. A	Open	Adapt.	P-value
Task success(%)	100%	91,67%	0,339
System Errors	1,827	1,284	0,381
Time (seconds)	81,0	94,7	0,602
# user turns	6,217	5,486	0,409
User Satisfaction Mean (1..5)	3,352	3,500	0,426
Ambiguous messages	0,2675	0,0275	0,021

Conclusions

- ▶ Dialogue systems can guide different type of users when accessing the web
- ▶ Dialogue systems can be adapted to different types of web contents, such as informational and transactional web services
- ▶ Performance is better when general linguistic resources are adapted to the specific domain by experts
- ▶ The use of general knowledge bases reduces the cost of adapting the dialogue system to different web services, although human intervention is still needed to achieve a good performance.

Further information

{gatus,mgonzalez}@lsi.upc.edu