TRABHCI 28 March-8 April Valencia (Spain)



Spoken Dialog Systems Human-Computer interaction, audiovisual interfaces and natural interaction

María José Castro-Bleda mcastro@dsic.upv.es

What is a Spoken Dialog System?

- Talking is the most natural way to solve a great number of daily actions between humans: to obtain a certain information, to contract a service, to ask for an order, or to know the state a certain process, etc.
- Inside the research area of Speech Technology a primary target is to obtain systems that make easy the communication between humans and machines in the most natural way.
- In this way, a spoken dialog system can be understood like an automatic system able to emulate a person in a dialogue with another person, with the target of the system performs a certain task (typically to provide a certain information or to carry out a certain task).

What is it needed?

- This type of applications requires a sequence of interactions between the person and the machine to obtain the user target.
 - To keep a history of which it has been said.
 - To take the initiative to lead the dialogue inside the domain defined.
 - To request for information needed to reach the dialog target.
 - To require to the users explanations about the information supply by the user.

State-of-the-art

- To have a conversation person-machine in a natural way is nowadays a challenge.
- Some speech conversational systems has been developed able to interact with certain degree of flexibility (mixed initiative in the development of the dialogue). These systems always are oriented to tasks of very specific information (restricted semantic dominions).
 - Information systems for timetables and prices in public transports, weather information, bank services, applications accessible from a car, facilitate the access to the information to a handicapped person, access by phone or mobile phone to services or control machines, etc..

Evolution of Spoken Dialog Systems

(Pieraccini,2007)



Some Examples of Applications

Call routing

Access to an Information Service:

- flights: Communicator
- trains: SUNDIAL, ARISE, DIHANA

Planning: TRAINS, TRIPS

Restaurants: MIT

Weather: JUPITER

Some Videos

- http://politube.upv.es/play.php?vid=47493
- http://politube.upv.es/play.php?vid=47490
- http://www.youtube.com/watch?v=Y0hl1-06gOo

Spoken Dialog System Scheme



Modules in a Spoken Dialog System



Dialog Manager

- It is the central module of the system, it interacts as much with the modules that compose the input of the system like with the modules that generate the exit of the system.
- The Dialog Manager
 - must keep the dialog history,
 - must have a current state, and
 - finally a strategy to decide what the system must do.

Main actions of the Dialog Manager

- To ask for more information
- To give the answer of the query
- To start a process of error recovery (or hypothesis confirmation)
- To detect inappropriate queries.
- To verify the consistency of the values (Example: the 30th February).
- Relax
- Constraint
- Confirmation strategy: explicit confirmations, implicit or mixed.

Development Methodologies

Rules

- Disadvantages: high cost of model design, specialized personal and low portability from one application to another.
- Stochastic methods based on machine learning (from a corpus of dialogs)
 - Disadvantages: the corpus acquisition and the annotation, the size of samples of the corpus must to be big enough to infer a good model.

Corpus Acquisition: Wizard of Oz



Different behavior of the User when he is talking to a machine or a human.

Corpus Annotation

- □ Morfosyntactic, assigns a grammatical class to each word.
- Syntactic, a hierarchical structure that represents the relationships between words in the same sentence.
- Prosodic, word pronunciation, delays, rhythm
- Pragmatic, user intention, dialog acts.
- Tools: DAT, MMAX, Callisto, AGTK, Alembic, AnnoTag, DiET, Nite, Nb, MUP, EMU

Semantic Representation

- Usually the Semantic representation is the FRAME:
 - A concept or several concepts and the constraints associated to it/them.
- Example:

I would like the flights from Denver to Pittsburgh for tomorrow please.

Concept :	flight
Constraints :	Departure-town : Denver
	Arrival-town : Pittsburgh
	Departure-day : tomorrow

Typical Discourse Phenomena in Conversational Systems

- Anaphoric (e.g., pronominal) reference:
- 1. Show me the restaurants in Cambridge.
- 2. What is the phone number of the third one?
- 3. How do I get there from the nearest subway stop?

• Ellipsis:

- 1. When does flight twenty two arrive in Dallas?
- 2. What is the departure time ()?

• Fragments:

- 1. What is the weather today in Denver?
- 2. How about Salt Lake City?
- Semantic ambiguity:
- 1. Not to Denver, to Chicago
- 2. Not, to Denver

Dialog Flow

- □ The dialog flow can have the following strategies:
 - System-initiative dialog: The system is in control to guide the dialog at each step.
 - Mixed-initiative dialog: Users can barge in and change the dialog direction. The system follows the user request, but tries to direct the user back the original course. This is the most commonly used dialog strategy in today's dialog systems.
 - User-initiative dialog: The user takes lead, and the system respond to whatever the user directs.

Finite State Model



Explicit Confirmation



Implicit Confirmation



System initiative Dialog example

S: Would you like movie showtime or theater playlist information?

U: Movie showtime.

S: What movie do you want showtime information about?

U: Saving Private Ryan.

S: At what theater do you want to see Saving Private Ryan?

U: Paramount theater.

S: Saving Private Ryan is not playing at the Paramount theater.

Mixed initiative

- Conversational initiative can shift between system and user.
- Example:

S: How may I help you?
U: When is Saving Private Ryan playing?
S: For what theater?
U: The Paramount theater.
S: Saving Private Ryan is not playing at the Paramount theater, but it's playing at the Madison theater at 3:00, 5:30, 8:00, and 10:30.

- S: How may I help you?
- U: When is Saving Private Ryan playing?
- S: For what theater?
- U: The Paramount theater.
- S: Saving Private Ryan is playing at the Fairmont theater at 6:00 and 8:30.
- U: I wanted to know about the Paramount theater, not the Fairmont theater.

S: Saving Private Ryan is not playing at the Paramount theater, but it's playing at the Madison theater at 3:00, 5:30, 8:00, and 10:30.

Comparison

Dialogue 1:

- System-initiative
- Implicit confirmation
- Merely informs user of failed query

Dialogue 2 and 3:

- Mixed-initiative
- No confirmation
- Suggests alternative when query fails
- More natural
- Most efficient