Rapid Multimodal Dialogue Design: Application in a Multimodal Meeting Retrieval and Browsing System

Miroslav Melichar, Martin Rajman
LIA/CGC, EPFL, Lausanne, Switzerland

Agnes Lisowska, Susan Armstrong
ISSCO/TIM/ETI, University of Geneva, Switzerland

The Archivus Interface

The user interface is based on the metaphor of a person interacting in an archive or library.

The input modalities:
- natural language through
- direct manipulation
- mouse
- touchscreen

The output modalities:
- audio
- text
- graphics
- video for showing meetings

The choice of modalities determined primarily by practical motivations – eventual goal is deployment in a commercial or academic institution with standard software and hardware rather than high-end or laboratory prototypes.

Rapid Dialogue Prototyping Methodology (RDPM)

Task Model
Model of domain that:
- lets user select a subset of possible target solutions
- available targets can be described by specific attribute-value pairs useful for the domain

Generic Dialogue Nodes (GDNs)
Perform simple interactions with the user to obtain a valid value for an attribute to which they are associated.

Three types of interactions are possible depending on the type of value sought
1. Simple:
   allow a user to directly specify a value for the associated attribute.
2. List Processing:
   allow the user to browse through a list of values and select by number (position of value in the list).
3. Internal:
   special GDNs invoked by dialogue manager in specific situations – e.g. beginning of the dialogue, over-specification of a request.

Dialogue Strategies
- serve as decision making mechanism of dialogue manager
- provide rules about the progression of dialogue
- RDPM dialogue-management approach handles dialogue strategies at two levels: local and global.

Multimodal Rapid Dialogue Prototyping Methodology (mRDPM)

Multimodal Generic Dialogue Nodes (GDNs)
mGDNs are as at left plus:
- grammars for written and spoken natural language input
- a set of multimodal prompts to guide the user
- information about the graphical representation of the mGDN
- a definition of the role of each GUI element

Multimodal fusion and fission
- is handled at the level of individual mGDNs rather than at the global level.
- mGDNs impose a constrained interaction and define a relatively narrow interpretation context for input and output making it possible to resolve fusion/fission problems.

System Architecture

An example of using Archivus
- new employee of a company is asked by their manager to write a short report about what happened at a particular meeting
- the manager doesn't remember any details about the meeting except that it took place at the end of April, Susan was there and there was a discussion about red sofas

Step 1
- user sees the initial system screen
- user asks 'Show me the meetings which Susan attended'

Step 2
- Susan added to interaction history
- bookcase shows 8 meetings (books) which Susan attended
- system suggests user define a date as a criterion
- user is unsure of exact date and selects the 'topic' button via a mouse click

Step 3
- list of all possible topics contained in the active meetings appears
- user selects through the list and says 'Susan attended meetings'

system identifies it as the relevant topic

Step 4
- topic is added to interaction history and bookcase is updated
- only 2 books match the specifications made by the user
- system suggests specifying the day of month of the meeting (since the other date parameters are the same for both meetings)
- this is the only other piece of information that the user knows, so he does what system suggests
- user clicks on the date at the end of April

Step 5
- system adds the criterion to interaction history and updates bookcase
- only 1 book matches the criteria so it is opened in the interaction pane
- user can browse book by:
  - looking at transcript (flipping the pages of the book using next/previous)
  - watching the video for particular pages by activating the video icon (by clicking on it or by saying "play video")
  - listening to the audio by activating the audio icon (either by clicking or by saying "play audio")
  - since a topic was specified the sections of the meeting related to that topic are marked by tabs on the book and highlighted on the pages to facilitate browsing

Application Specific Interface Design

- immediate feedback to user on effectiveness of search criteria and a permanent window on the selected subset of the database
- available modalities presented as visual cues
- consistency in the way that actions are performed across modalities