What is a dialogue system?

- McTear (2002 p.92) (Spoken Dialogue System)
  - "Computer systems that use spoken language to interact with users to accomplish a task."
  - "Computer systems with which humans interact on a turn-by-turn basis and in which spoken natural language plays an important part in the communication."

- Flycht-Eriksson (2001) (Dialogue System)
  - "A computer system that interacts with users utilizing connected natural language dialogue, where the use of natural language need not consist of predefined commands."

Modalities
- text or speech, and possibly
- any number of accompanying modalities

Complexity of interaction
- adequate response to user initiative,
- extended conversation,
- topical coherence,
- goal-directedness,
- meta-communication,
- multi-task, multi-domain, multi-device, multi-party, ...

Common system categories
- Interactive voice response systems
- Question-answering systems
- Interactive guides
- Natural-language interfaces
  - to databases, (information seeking)
  - to expert systems (problem-solving or collaborative systems)
- Spoken dialogue systems
- Conversational agents
- Conversational systems / Chatbots
- Dialogue support (translation) systems

Interactive guides

Question-answering systems

History in brief
- In the beginning ...
  - The Turing Test
  - Conversational partners
    - ELIZA (1966), PARRY (1971), ...
- Database interfaces
  - BASEBALL (1963),
  - LUNAR (1972)
- Commercial systems: INTELLECT, Q&A
History in brief

- Natural-language interfaces 1975-1990
  - Syntax-driven semantic interpretation
  - LUNAR (Woods et al. 1972, Woods 1977)
  - CHAT-80
  - Definite-Clause Grammars (Warren & Pereira, 1982)
  - The Core Language Engine (1992)

- AI research 1970-1990
  - SHRDLU (1972)
    - varied types of input (procedural semantics),
    - knowledge-based (blocks world, language)
  - Grosz (1977)
    - dialogue structure as based in task structure
  - Allen, Cohen, Patnaik, Carberry, Traum, ...
    - Agents act because they have beliefs, desires and intentions,
    - Dialogues are goal-directed,
    - Speech acts contribute to goal fulfillment

History in brief

- Spoken dialogue systems (1990- )
  - Dialogue systems as applications of speech technology (ASR, TTS)
  - Emphasis on system building
  - GALAXY (MIT, 1994)
  - Circuit-Fix-It Shop (1994)
  - Philips train timetable information system (1995)
  - Waxholm (1995)
  - TRAINS
  - GoDiS
  - ...

- The shift to empiricism in CL/AI (1990-)
  - Data collection and analysis
  - Robust methods for understanding
  - Design for usability
  - Methods for system development
  - Evaluation methods, and
  - The deployment of commercial systems
  - The launching of the Loebner Prize

The Loebner Prize

History in brief

- Multimodal systems (1980 - )
  - "Put that there" (Bolt, 1980)
  - Multimodal input guidance
  - NLMenu (Tennant, 1983)
  - Multimodal input interpretation
  - CUBRICON (Naal et al., 1989)
  - XTRA (Wahlster, 1989)
  - Multimodal generation
  - COMET (Fainner & McKeown, 1991)
  - WIP (Wahlster et al., 1993)
Merits of (practical) dialogue systems

- Natural interaction
- Allows for fragmentary input
- Clarifications
- Complex requests
- Contextual interpretation
- The language can be adapted to the user

Levels of Complexity

- Strict Policy
  - User can only specify information relating to current goal/subgoal
  - Context is easier to determine
- Free Policy
  - Handle unintended requests or requests that deviate from the task
  - Context more difficult to determine
  - Can lead to confusion/errors

Initiative

- System-initiative: system always has control, user only responds to system questions
- User-initiative: user always has control, system passively answers user questions
- Mixed-initiative: control switches between system and user using fixed rules
- Variable-initiative: control switches between system and user dynamically based on participant roles, dialogue history, etc.

Components/Architecture

- Dialogue manager
  - Speech recognition
  - Language understanding
  - External communication
  - Response generation
  - Speech output

Knowledge sources

- Dialogue history
  - A record of the dialogue so far
- Task record
  - Representation of the information to be gathered
- World knowledge
  - General commonsense background information
- Domain model
  - Information for the specific domain
- Conversational model
  - Principles of conversational turn-taking
- User model
  - User specific information

Dialogue management

- Focus handling
  - Contextual interpretation considering user input, dialogue history and domain
- Dialogue control
  - Task control
  - Sub dialogue control
  - Clarifications
Dialogue and task complexity (Allen)

<table>
<thead>
<tr>
<th>Technique used</th>
<th>Example Task</th>
<th>Dialogue phenomena</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite state script</td>
<td>Operator phone</td>
<td>User answers questions</td>
</tr>
<tr>
<td>Frame based models</td>
<td>Time-table</td>
<td>User asks questions, simple clarifications by system</td>
</tr>
<tr>
<td>Sense of contents</td>
<td>Travel looking</td>
<td>User session preconditions</td>
</tr>
<tr>
<td>Plan-based models</td>
<td>Design consultant</td>
<td>Dynamically generated topic structures, collaborative negotiation</td>
</tr>
</tbody>
</table>
| Agent-based models      | Disaster relief  |                                      |}

Finite State Dialogue Modeling

- System asks a series of questions that the user answers: “What number would you like to call?”, “Is this a Linköping number?”
- Initiative always with the system.
- Context is fixed by the question being asked.

Finite State Dialogue Grammar

```
S / Qn U / Ans S / Bye
```

Frame Based Dialogue Modeling

- System acquire enough information in order to perform a specific action.
- Single fixed context
- Form filling where the form specifies all relevant information for an action:
  - Monitor the form for completion.
  - From user utterances extract relevant elements.
  - Use empty slots as triggers for questions to the user.

Menus, Forms and Links

- Welcome to the Travel agency.
  - Book a trip.
  - Check arrivals or Talk to an operator.
- Book a trip.
  - From where do you want to go?
- From Stockholm
  - At what time?
- Three o'clock
  - You want to travel from Stockholm at three o'clock
- Abort
- Aborting. Welcome to the travel agency…

VoiceXML: Forms

```
<field name="to">
  <prompt>Where do you want to go?</prompt>
  <option>stockholm</option>
  <option>göteborg</option>
  <option>malmö</option>
</field>

<field name="when" type="time">
  <prompt>At what time?</prompt>
</field>

You want to travel to <value expr="to"/> at <value expr="when"/>

<form next="booking_to.vxml"/>
```

FIA: Form Interpretation Algorithm

- Go through each field in the form.
- Stop at the first field that is not filled.
- Play the prompt, start the speech recogniser.
- Fill the fields in the form that are specified by the matching grammar.
- If all fields are filled: Leave the form.
  - Otherwise: start again.
- Also possible to control the execution by scripting!