SoSe 06 - Project „Spoken Dialogue Systems“

Dialogue Management

Cristina Vertan, Walther v. Hahn

Contents

- Role of Dialogue Manager in a Dialogue Systems
- Types of dialogues
- Models for dialogue management
- Issues in controlling user‘s input

Role of Dialogue Manager

- Central component of a dialogue system
- Role: controls the dialogue flow, namely:
  - Accepts spoken input (transcripts) from the user (via the speech recognizer)
  - Interacts with external knowledge sources (e.g. DB with informations like train routes)
  - Produces messages which will be communicated to the user via the TTS.
- The complexity of the dialogue manager depends on the complexity of the „natural language“ to be used

Types of Dialogue Systems /Dialogue Managers

- According to the input type:
  - Only spoken (especially important for disabled users)
  - Only text (quite rare)
  - Spoken + DTMF(Dual Tone Multiple Frequency) (very frequent)
- According to the output type:
  - Spoken
  - Text
  - Text with visual elements (graphics, images, tables)
- According to the initiator of the dialogue:
  - System-directed (system asks a sequence to acquire the required parameters of the task from the user)
  - User-directed (user controls the dialogue and asks questions i order to obtain some information)
  - Mixed-initiative (dialogue control is shared, user can ask questions at any time but system can take control and ask for clarifications)

System-directed Dialogue

- Most part of the current existent systems.
- User‘s input is restricted to single words or phrases that provide answers to carefully designed prompts
- Advantage: the required vocabulary and grammar for each answer can be specified in advance
- In this way Speech Recognition and natural language understanding are constrained and consequently very accurate
- Disadvantage:
  - the user has a lot of restrictions
  - Corrections of missrecognized items is difficult
  - It is also difficult to inhibit the user to take the control of the dialogue (ask questions or introduce new topics)

System-directed Dialogue -Example

- System: Where are you travelling to?
- User: London
- System: What day do you wish to travel?
- User: Friday
- System: At what time?
- User: 9 a.m.

User-directed Dialogue

- Very similar to NL-interfaces to DB
- The user determines the questions to be asked
- In some systems the system may ask clarifications
- Advantage: user is not restricted
- Disadvantage: the system need comprehensive speech and language processing capabilities (large lexicon, grammar, discourse processing).
- It is difficult to make clear to the user which words and syntax the system knows

User-directed dialogue -Example-

- How many employees living in the London area earn more than 50.000 pounds?
- System: Fifty four.
- User: How many are female?
- System: Eleven
- User: And managers?
- System: Nine

Keep track of entites previously mentioned. Extend to: *„How many female employees living in the London area earn more than 50.000 pounds?“*

Extend to: *„How many managers living in the London area earn more than 50.000 pounds?“*

Mixed initiative Dialogue

- User can take initiative to ask questions, intiate topics, request clarifications, etc.
- In VoiceXML context the term is used in a more restrictive sense: to describe dialogues in which the system has overall control of the dialogue but the user has some flexibility to volunteer more information.

Mixed initiative Dialogue -Example-

- System: Where are you travelling to?
- User: I want to fly to London on Friday
- System: At what time do you want to fly to London?
- User: Are there any cheap flights?

Shift intiative

Mixed initiative Dialogue
-VoiceXML Example-

- System: Where are you travelling to?
- User: I want to fly to London on Friday
- System: At what time do you want to fly to London?
- User: 9 a.m.

Over-informative utterance, not rejected by the system

Dialogue control

There are three methods for representing and implementing the flow of the dialogue

1. **Finite state-based systems** - the dialogue flow can be determined in advance and represented in terms of a dialogue network or grammar
2. **Frame-based systems** - implement the limited mixed-initiative as in VoiceXML
3. **Agent-based systems** - the dialogue flow is determined dynamically through a process in which the dialogue manager performs some reasoning in order to determine the next actions

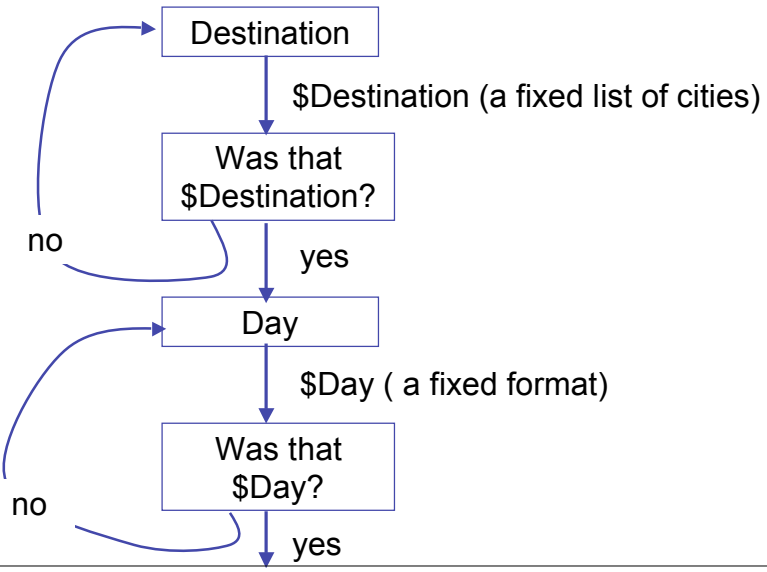
1. Finite State-based Dialogue Control -1-

- The dialogue structure represented as a transition network where:
  - The nodes represent the systems questions
  - The transitions between nodes represent all possible paths through the network
- The graph specifies all legal dialogues
- Each state represents a stage in the dialogue in which
  - some information is elicited from or confirmed with the user
  - Or some action is performed by the system

1. Finite State-based Dialogue Control -2-

- Dialogue Networks can be used both for simple and complex dialogue systmes
- Sub-dialogue networks can be used to model subtasks in a dialogue
- One can insert always a kind of „default-repair“ sub-dialogue to handle situations where the system cannot detect any input from the user or the reccognition accuracy is under an acceptable treshhold

1. Finite State-based Dialogue Control -Example-



1. Finite State-based Dialogue Control Pros and Cons

- Advantages:
  - Vers simple model
  - Natural way of modelling system-directed dialogue systems
    - that involve well-structured tasks and
    - Information is exchanges in a well defined sequence
- Disadvantages:
  - Inflexible
  - Dialogue paths are specified in advance, there is no way to deviate from the paths
  - User cannot introduce information which were maybe not previewed when designing the system
  - E.g. planning a journey requires dynamic contraints introduced by the user

Frame-based Dialogue Control

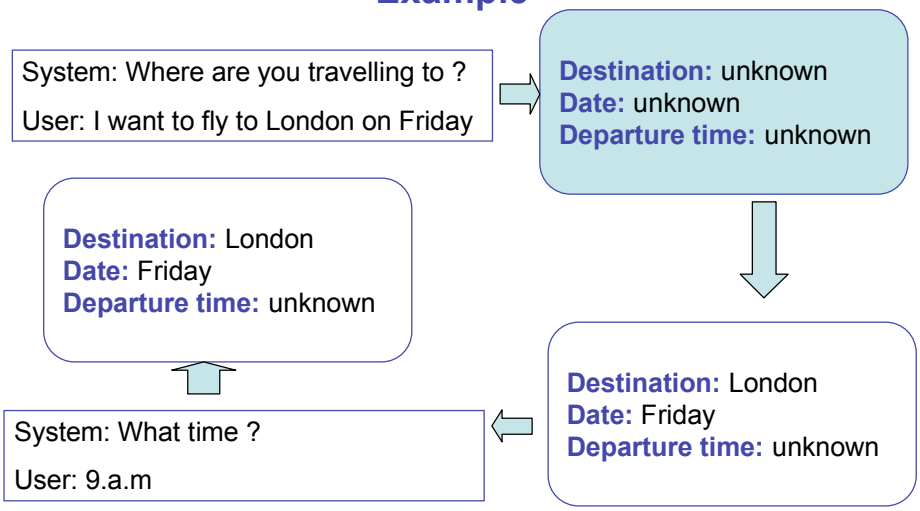
- Suitable for form-filling tasks in which the system asks the user a series of questions to gather informaion, and then consults the external knowledge source
- There is no predefined sequence of form completion
- Mainly used to implement the limited mixed initiative supported by VoiceXML

Frame-based Dialogue Control -Components-

1. A frame (template) which keeps track of the items of information which the system has to elicit from the user
2. A more extensive recognition grammar
3. A dialogue control algorithm that can detremine the system‘s next actions based on the contents of the frame.

- The frame is a simple data structure consisting of a number of slots to be filled with appropriate values

Frame-based Dialogue Control
-Example-



Frame-based Dialogue Control
-complex features-

- Instead of specifying for the fields only „unknown/filled“ a new value could be introduced to check if the value was verifiedor not
- The verification could be also refined with some confidence values
- This can have implications on further steps in the dialogue
- There are dialogue systems which include a user profiel. According to the user profile certain slots have assigned priorities.

Frame-based Dialogue Control Grammars

- The recognition grammar is more complex than in finite-state based systems, so that the user has some freedom in expression
- E.g for the anwer to the question „Where are you travelling to?“ the user has following options:
  - Destination
  - Destination + Date
  - Destination + Time
  - Destination + Date + Time
  - Destination + time + Date

Agent-based Dialogue Systems

- Rely on techniques from AI and focus on the modelling of dialogue as collaboration between intelligent agents
- Appropriate fpormore complex tasks as negotiation and problem solving
- The dialogue is performed by 2 agents (1 the human and 1 the machine). Each agent is capable of reasoning

Agent-based Dialogue Systems
-Example-

- User: I‘m looking for a job in Granada. Are there any servers there ?
- System: No, there is no server in Granada . However there is one in Malaga and one in Sevilla. Are you interested in one of these ?

The system instead of giving only a negative answer proposes some alternatives.

Agent-based Dialogue Systems
-main problems-

1. The system cannot guarantee that the representation of the user‘s input, that it receives from the speech recognitionand language understanding components , is accurate
2. There are many discrepancies between the information that the user requests and what is available in the external knowldge source.

Agent-based Dialogue Systems
-why erroneous user‘s input-

- The dialogue manager receives a representation of what the user said from the speech recognition and language understanding components and decides what to do with this input.
- Input may be input because:
  - The speech recognizer may have detected silence even though the user had spoken --> no words will be returned
  - Only a part of the user‘s utterance has been recognised and returned. Very often when the user has to speak after a „beep“ the first part of the user‘s utterance is not recognized.
  - All the utterance was captured but not all the words were recognized correctly
  - The utterance was recognized correctly but the NL component assigns wrong meanings to words, or a number of possible meanings (ambiguity)

Agent-based Dialogue Systems
-how to repair user‘s input -

1. Clarification subdialogues - report the problem back to the user and request a reformulation of the input

- Main problem: the user receives nno indication why there was a problem, i.e why the input may be ill-formed or incomplete

2. Address the problem explicitly. E.g if a silence ghas be detected promt something like:

„I do not hear anything. Please speak louder“

Or in case the meaning seem to be false the system could ask for reformulation

Agent-based Dialogue Systems
-how to verify user‘s input -

- Even when no error is detected it is good that the system verifies the corectness of the utterance

1. Explicit Verification: a question asking for confirmation of the input

Although the method is robust in case of longer utterances the user can miss some verifications

Ask verification for each value. The risk in this case is that the user will be rapidly anoyed

2. Implicit verification: The system embeds in his next question parts of the previous utterance.

Verification of utterances is a very important and challenging issues in spoken dialogues systems (see talk about Pragmatics)

Accessing Information from an External Knowledge Source

Problems:

- The vocabulary of the dialogue does not map directly on to the vocabulary of the application
- The data which is retrieved is ambigous or indeterminate.
- How to present the data to the user

Choices to be made in the development of a Dialogue systems

- Choice of dialogue initiative: system-directed, **mixed-initiative** or user-directed
- Choice of dialogue control strategy: finite-state-based, **frame-based** or agent-based
- **Design of system prompts**: a carefully designed system prompt can constrain the range of possible user responses and support the speech recognizer and NL componet
- **Choice of verification strategy**: compromise between ensuring accuracy at all costs and avoiding very lengthy transactions.
- **Choice of error recovery strategies** (involves also error detection)
- Designing and implementing links to external knowledge sources