NATURAL LANGUAGE INTERFACES AND DIALOGUE SYSTEMS

Natural Language Interfaces and Dialogue Systems

- **Introduction**
- Tasks of the dialogue systems
  - Interpreting user intervention
  - Dialogue Management
  - Generating system's intervention
- Architecture
- Development and evaluation

Introduction

- **Main goal in humam/machine communication**

Help users perform specific tasks according their objectives

- **Two metaphores**

  - Human conversation

    question/answering, menu systems, forms, command language dialogue

  - World model

    Direct access to the domain objects (the icons representing them). *WYSWYG*

Using the Natural Language Mode

- Advantages
  - Human Language (natural,friendly)
  - Complex ideas can be expressed
  - References to other entities are easy to express
- Disadvantages
  - High cost
  - Ambiguity -- mistakes
  - Limitations for accessing several applications (such as graphics)
- Appropriate for occasional access to applications that need to express complex operations (especially when domain can be restricted)

Using the Natural Language Mode

- Advantages of the speech mode
  - Convenient
  - Typing is not needed
    - Small devices
    - Hands cannot be used (i.e.: driving cars, airplanes)
    - Handicapped
- Disadvantages of the speech mode
  - Technical limitations to recognize voice
  - Ambiguity, mistakes
  - High cost
  - Not appropriate for specific applications
- **Appropriate for expressing simple operations on restricted domains**

The Need for Dialogue Systems

- User cannot express a requirement in only one intervention
- User need the system intervention to make the communication friendlier

An Example of Conversational System

System: **Welcome to the informaton service, what do you want?**

User: **I want to go from Barcelona to Valen**cia

System: **When do you want to go?**

User: **Next Tuesday**

System: **At what time, morning or afternoon?**

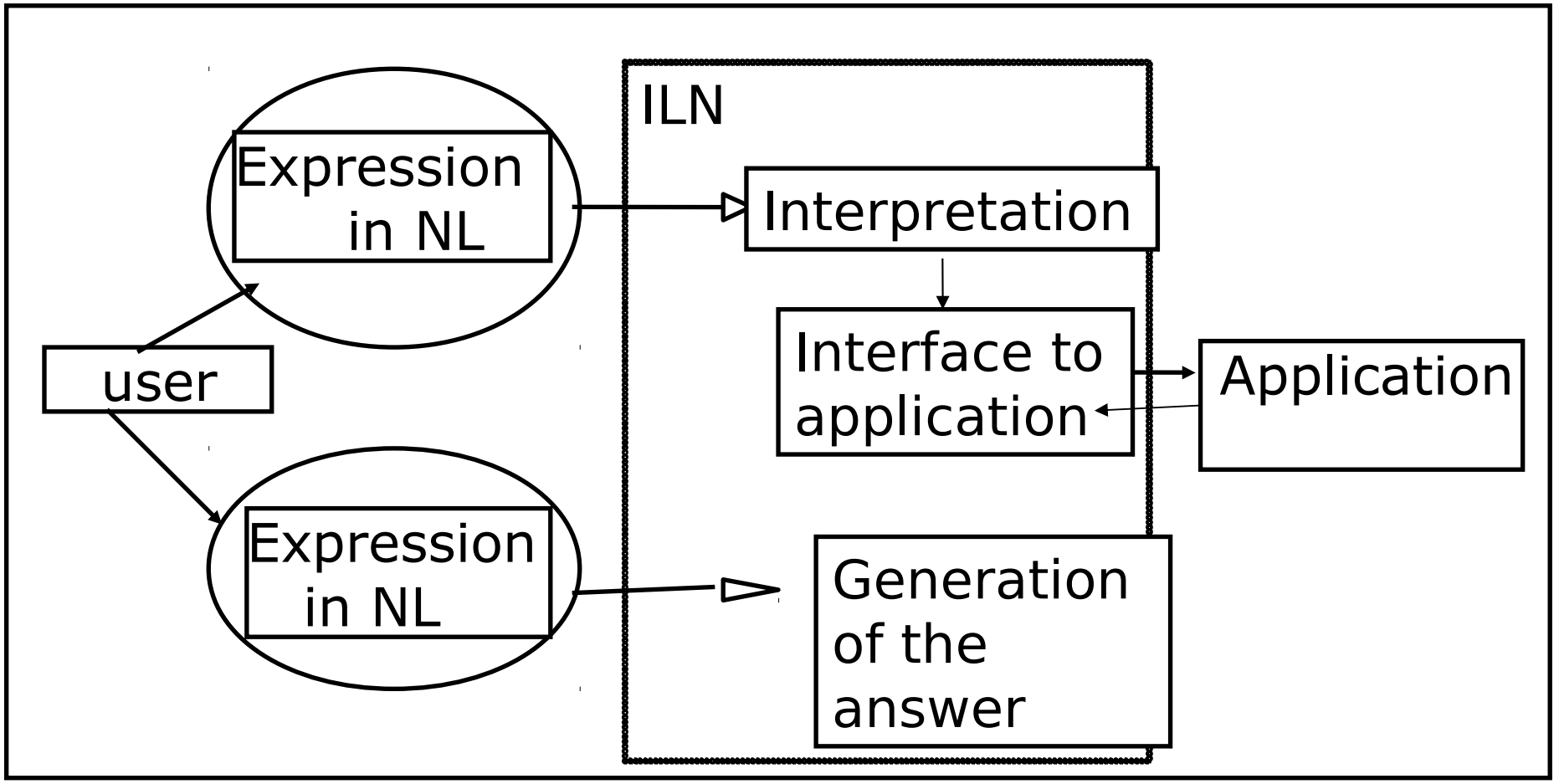
User: **Morning, please**

System: **There are 3 Euromed trains on Tuesday morning**

The Terms

- **Human Computer Interfaces**
- **Natural Language Interfaces**
- **Speech Based Interfaces**
- **Conversational Interfaces**
- **Conversational Agents**
- **Dialogue Systems**
- **Spoken Dialogue Systems**
- **Spoken Language Systems**

Basic cycle in an interaction question/answering



Main trends in Natural Language Interfaces

- Evolution parallel to Language Engenieering
- Improving adaptability and friendliness
- Portable (or adaptable) interfaces.
  - They are usually developed for a specific domain or application.

    - They use domain (or application) restricted knowledge (domain model)
  - They are usually developed for one (or more) specific languages

Main trends in Natural Language Interfaces

- Improving portability: Reusability of basic components
  - Speech recognizer
  - Language interpretation and generation: lexicon, parsers, grammars
- Improving friendliness
  - Multimodality
  - Friendly systems

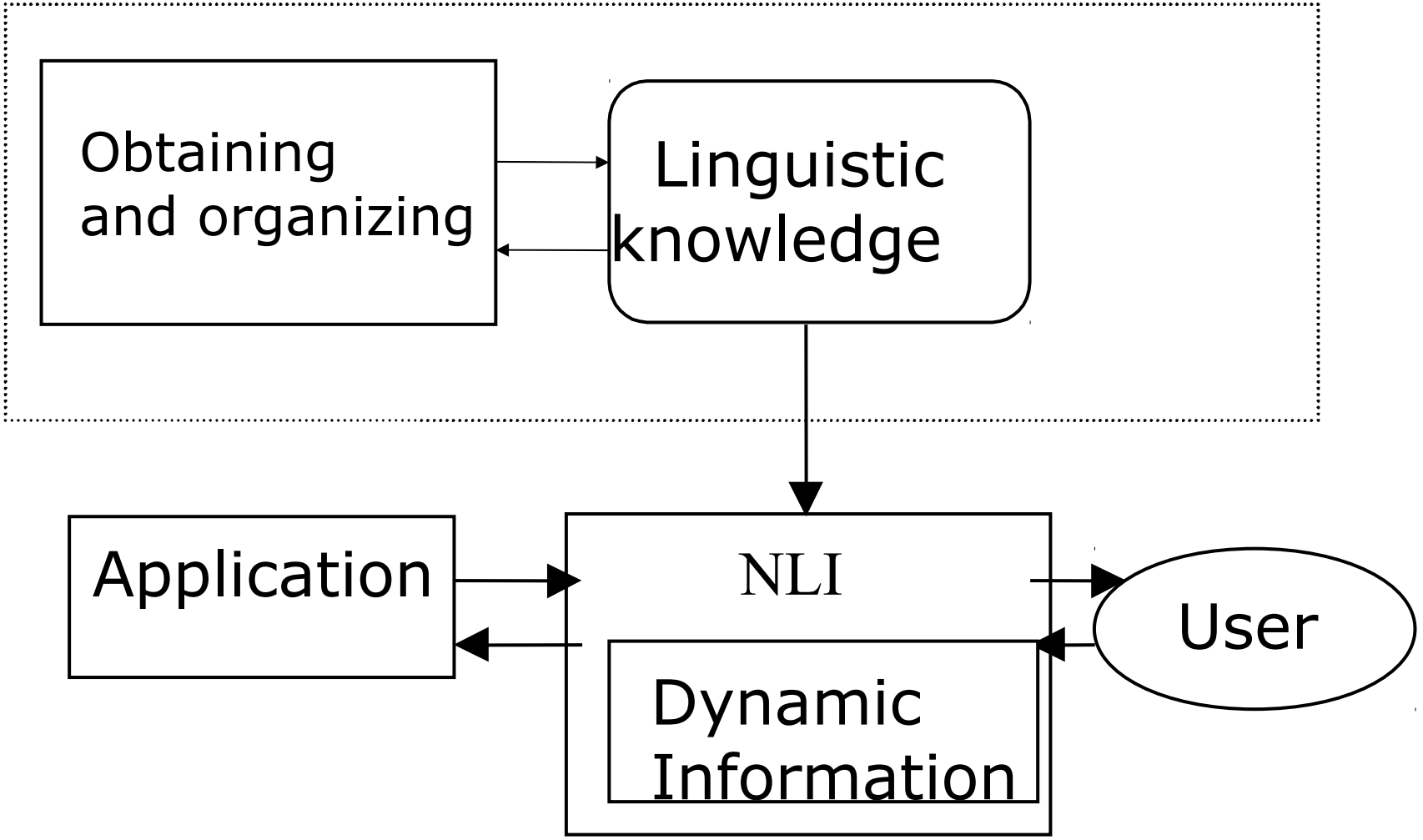
Improving friendliness

- Quality of the interaction
- Supporting possible mistakes in the user intervention
  - Spelling and grammar mistakes
  - Short names, colloquial expressions, other language words
- Mixed and user initiative
- User models

Multimmodality

- Integration of graphics and language for input and output expressions.
- Supporting complex references across graphics and text.
- Integration of gestures and text
- Considering the content, the context and the user model to choose the best presentation of the output (and the best way to combine different modes of interaction)

General Schema of a Natural Language Interface



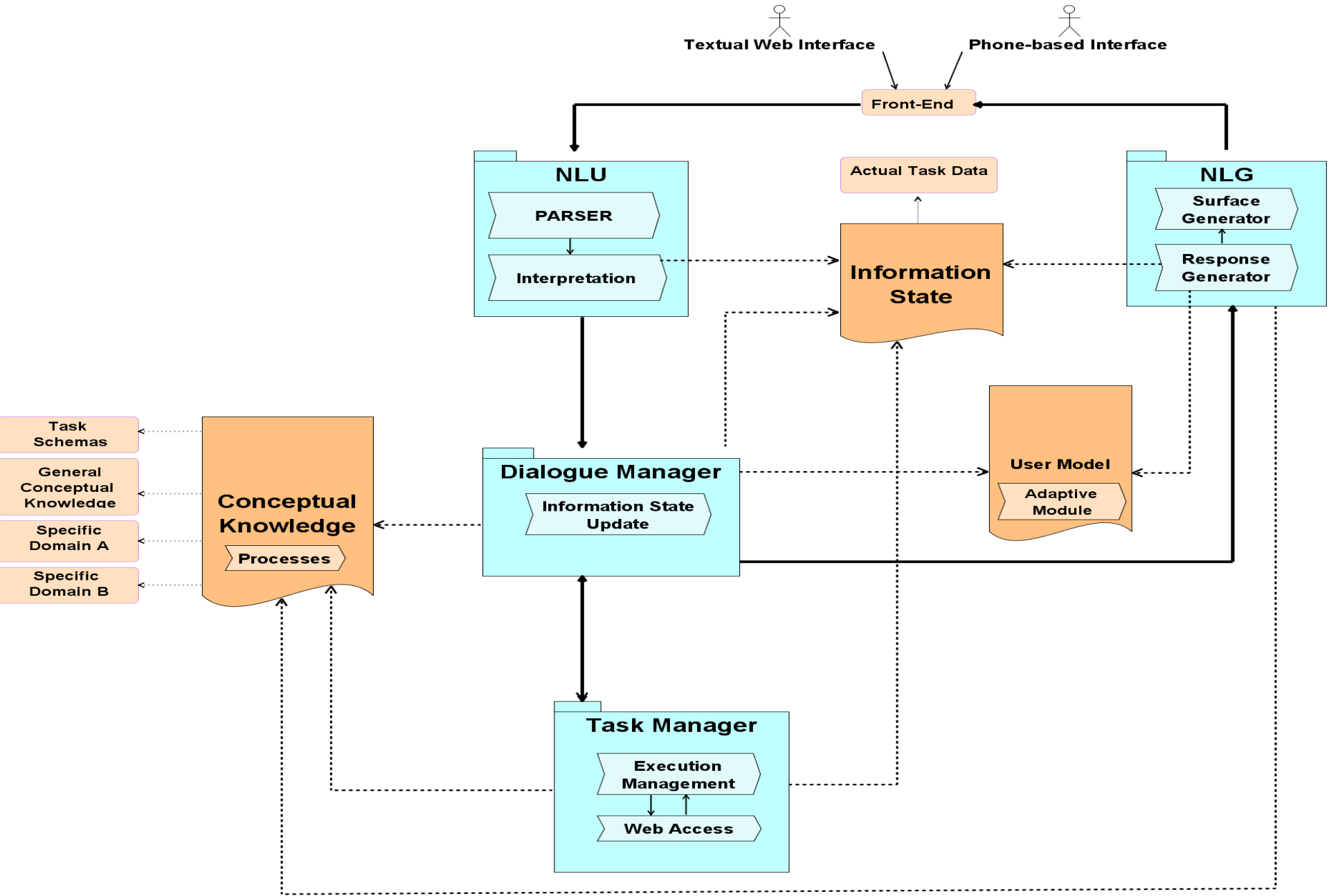
First Natural Language Interfaces

- NLI to Databases
  - 70s : LUNAR (Woods), RENDEZVOUS (Codd, 1974)
  - 80s: LADDER (Hendrix), TEAM (Grosz), CHAT-80 (Warren, Pereira) , DATALOG (Hafner)
- Accessing other applications
  - Operating systems, experts systems, tutoring systems, consulting systems, ...
  - Research systems
    - LDC, TELI, XTRA, INKA, XCALIBUR, UC, GISE
  - Comertial systems
    - INTELLECT, PARLANCE, HAL, Q&A, LOQUI, NAT, NL, SESAME

Complex Natural Language Interfaces

- Multimodal systems
  - MMI2 (Multimodal Interface for Man Machine Interaction), MATIS (Multimodal Airline Travel Information System), ALFRESCO
- Using voice
  - Dictation systems, telephone interfaces
  - VOYAGER (MIT), Office Manager (CMU), MASK (Multimodal Multimedia Automated Service Kiosk), ATIS (MIT, CMU), Railtel, Sundial, Verbmobil
  - VoiceXML

Architecture of a Dialogue System



Tasks of the Dialogue Systems

- Interpreting the user's intervention
  - Using dialog and domain knowledge
- Dialogue Management
  - Determine next system actions considering user's intention
- Answer Generation
  - Generate the system's sentences most appropriate at each state of the dialogue

Interpreting the user intervention

- Goal: understanding user's intention
- Knowledge involved
  - Phonetics and phonology
  - Morphology
  - Syntax
  - Semantics (lexical and compositional)
  - Pragmatics
  - Discourse

Interpreting the user intervention

- Goal: understanding the user's intention
- Precise information from the user is required
- The complexity of this process depends on the system
  - Complete (deep) syntactic and semantic analysis
  - Partial (shallow) syntactic and semantic analysis
  - *Processing key words*

- This process is restricted by considering limited applications tasks

Interpreting the user intervention

- Main tasks
  - Reference resolution
  - Intention recognition
- The use of the context
  - Dialogue history
  - Domain knowledge

Interpreting the user intervention

- Main problems
  - Processing syntactic relations
  - Quantification
  - Coordination and subordination
  - References
  - Ellipsis
  - Ungrammatical expressions

Reference resolution

U: Where the movie Heroes is shown in Sant Cugat?
S: Heroes is shown at Cinema Cinesa in Sant Cugat
U: At what time is it shown?
S: It is shown at 8:30pm, 10pm and 11:30pm.
U: I want 2 tickets for adults and 2 for children for first session
How much is it?

- Knowledge Sources:
  - Domain Knowledge
  - Dialogue Knowledge
  - Domain (world) knowledge

Reference resolution (real systems)

- There is no reference resolution
- Only simple references are processed
  - A stack with the entities that are the possible focus is used
  - No discourse structure is used

Reference resolution

- Central elements of the sentences have to be selected
  - They are grammatically related to the main verb (subject, object,…)
  - They can connect a sentence with previous
  - They can connect a sentence with next
- When pronouns are found several rules are used to rank and filter the possible central elements

Reference resolution

- Most references are solved using knowledge discourse
- Central elements (focus) are stored in a stack
    - Only lasts nominal groups are stored
- Objects satisfying syntactic, semantic and pragmatic restrictions are selected
  - Starting by the stack top
    - “There ” is a place
  - Considering discourse structure
    - Relating objects and subdialogs

Intention Recognition

- User's interventions are interpreted as one (or more) **dialogue act** (*speech act* or *dialogue move*)
- Examples of dialogue moves
  - **Switchboard DAMSL**
    - Ini/final conventional
    - Opinion
    - Confirming/Accepting
    - Recognizance
    - Question/Answer/Yes-No
    - No-verbal
    - Quit
  - **Verbmobil**
    - Greet/Thank you/Goodbay
    - Suggestion
    - Accepting/Not accepting
    - Confirmation
    - Question/ Clarification/Answer
    - Giving the reason
    - Thinking
- Efforts for standard definition

Intention Recognition

- Dialogue grammars (finite state machine)

Greet → Question ⇄ Answer → Thankyou → Goodbay

- Plans
  - Receipts: General frames to perform actions
  - Inference rules
    - Planification rules (Artifitial Intelligence)

Intention Recognition (Real systems)

- The system infers the application task the user is asking for
  - Application: Giving information on cultural events
    - *Time or place where a specific event takes place*
    - *Events that take place in a specific place*
  - Application: Giving information on trains
    - *Schedule for a specific train*
- The system asks the user the information the application needs
  - The system ignores the information not useful for the application

Intention Recognition (Real systems)

- System initiative
- User initiative very limited
  - Not allowed in complex acts such as confirmation, clarification and indirect answers

S1: Which is your account number?
U1: My account number in Online Bank?

S2: Would you want to transfer 1500 euros to your new account?
U2: If I have this amount, ok

Intention Recognition (Real systems)

- Content obtained from the user's intervention
  - The application task that has to be performed
    - *Information on classical music concerts in Barcelona*
  - The information needed to perform the specific task
    - *The specific date and place*
    - *The next Saturday on the Auditori*

Intention Recognition (Real systems)

Several methods can be used

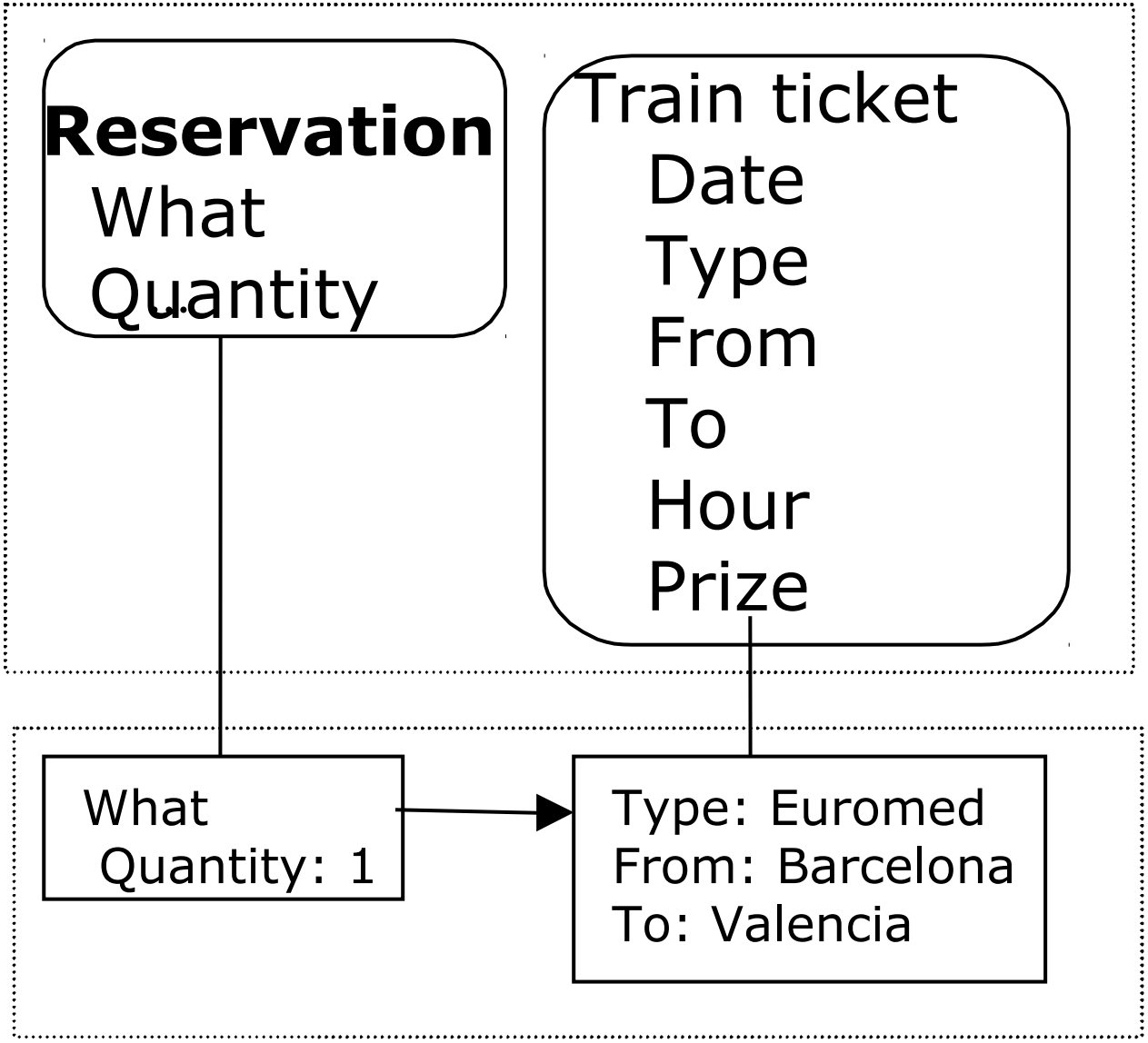
- *Frames* representing the information needed for each task
  - Trains schedule: departure and destination
- Similarity mesurement based on vectors

Intention Recognition (Real systems)

Using frames

- Representing tasks as frames described by attributes that correspond to input and output task parameters
  - Representing the tasks of giving information about a specific train as a frame where
    - departure and destination station are represented as attributes which value has to be provided by the user
    - time and price are the attributes obtained from the application and presented to the user

Representation of a user intervention asking for a ticket



Semantic Grammar

Ticket_re ->initial, quantification, ticket,[mods]

initial -> “I want”| “I want to make”

quantification -> “one”|”two”|1

ticket -> “tickets”|”tickets”|”reservation”

mods -> mod,[mods]

mod -> “from”, city| “to”, city|”on”, typetrain

city -> barcelona|valencia

typetrain ->euromed

Intention Recognition (Real systems)

U: At ***what time*** ***Madagascar*** *is shown in* ***Sant Cugat****?*

Intention: asking for information about a particular movie

Frame: Movie_information

Attributes which value is given by the user

Movie:Madagascar

Place: ***Sant Cugat***

Attributes which value is obtained by the application: ***time***

Intention Recognition (Real systems)

- Using keywords and mesures based on similarity vectors
- Representing questions and goals (answers) as vectors
- Searching the goal vector most simlar to the question
- Similarity metrics
  - cosinus of the angle between the two vectors

Tasks of the Dialogue Systems

- Interpreting user intervention
  - Using dialog and domain knowledge
- **Dialogue Management**
  - Determine next system actions considering user's intention
- Answer Generation
  - Generate the system's sentences most appropriate at each state of the dialogue

Dialogue Management

- Controlling dialog to help the user to achieve his goals
  - At each step of the conversation
    - **Who can speak**
    - What can be said
  - Used information
    - Interpretation of the user intervention
    - Application (domain) knowledge

Dialogue Management

- Determine the next system's action(s)
  - Answer user's questions
  - Ask the user for more information
  - Confirm/Clarify user's interventions
  - Notify problems when accessing the application
  - Suggest alternatives
- Generation of the system's messages
  - The content
  - The presentation

Dialogue Management

Example: Application giving information on flights

Attribute	Optional question
DepartureAirport	"From which airport do you leave?"
ArrivalAirport	"Where do you want to flight?"
DepartureTime	"At what time do you flight"
ArrivalTime	"At what time it arrives?"
Class	
Company	

Dialogue Management

Preparing next system's intervention

Obtaining information from the user:

- Ambiguities because recognition problems
  - *Did you said Barcelona o Badalona ?*
- Uncompleted specification
  - *On which day do you want to travel ?*

- Giving an answer to the user's question
  - Asking the user to restrict the question when many results are found :
    - *I found 10 flights, do you prefer any special flight?*
  - Asking the user to relax the question when no results are found
    - *The are flight on the morning, would you mind flying at night?*

Dialogue Management

- Guiding the user about accessible information
  - Presenting new goals

    Do you want to know the flight price?

    Presenting alternatives

    There is no information about Girona airport, only about Barcelona airport
- Guiding the user about the system's limitations:
  - When there is user's initiative there are more problemes caused by the lack of information of:
    - Application tasks
    - Domain information
    - Language recognized by the system
- Helping guides: initial indications, system's messages

Dialogue Management

- Errors recovery
  - Different causes: noise envoirment, accents, vocabulary
  - Several strategies to deal with problematic input
    - Directed dialog
    - Explicit confirmation: asking to confirm only what has not been completly understood
    - Using statistics
    - Others
- Dealing with interruptions

Dialogue Management

- Main design decisions
  - Functionality: The tasks the system has to perform
  - Processing: How these tasks have to be performed
- Considerations
  - Task complexity
  - Dialogue complexity: Which dialogue phenomena will be allowed
    - User initiative ?
- Results
  - Robustness
  - Natural

Dialogue Management

- Content of dialogues
  - Restricted to the information related to the application
  - Subdialogs: clarification, confirmation
  - Meta-dialogs (about the dialog)
    - *Are you still here?*
- Who can initiate different types of dialogues
  - Only the system
  - Only the user
  - Both

Functionality of dialogue manager

- Determine the set of possible goals the system can select at each turn
- Conditioned
  - Task complexity
  - Dialogue complexity

    Are subdialogues allowed? Who is allowed to use them?

    Are meta-dialogues allowed? Who is allowed to use them?

Funtionality of dialogue manager(Research)

- Task complexity: from medium to complex
  - *Travel planning*
- Dialogue complexity
  - Subdialogues supporting complex features
    - Mixed initiative
    - Collaboration to solve tasks
  - Meta-dialogues

Functionality of dialogue manager (Real systems)

- Tasks complexity: from simple to moderate
  - *Information about the weather*
  - *Information about train schedule*
- Dialogue complexity
  - Tasks restricted
  - System initiative
    - Limited subdialogues

Processing of dialogue manager

- Initiative strategies
  - System
  - User
  - Mixed
  - Variable
- Mechanisms for modeling initiative
  - System and mixed initiative: finite state machine
  - Variable initiative, depending on:
    - Dialogue history
      - Understanding errors
      - Others

Knowledge sources

- Dialogue Manager can use
  - Dialogue models
    - Define general dialogue phenomena
  - Tasks models
    - Define specific application tasks
    - *Frame based systems*. Obtain parameter values

      Flight information: departure, arrival, date
  - Domain models
    - Concepts and relations in a specific domain
    - Apropriate for complex domain
  - These models can be implicit
    - Finite state systems

Dialogue Model

- Dialogue model
  - Define the framework under which the user interventions have to be interpreted
- Dialogue state
  - Reference entities, relationship between them

Dialogue Management (Abstract)

- Decide system’s respond to user's intervention
  - Inferring user's needs
    - Dealing with ambiguity and not complete information
  - Accessing the application (or knowledge source)
  - Presenting the answer to the user

Dialogue Management (Abstract)

- Research systems
  - Focused on the development of models and algorithms for supporting several dialogue phenomena for complex tasks
- Real systems
  - Focused on the development of robust strategies, to deal efficiently with most common dialogue phenomena for simple applications

Tasks of the Dialogue Systems

- Interpreting user’s intervention
  - Using dialog and domain knowledge
- Dialogue Management
  - Determine next system actions considering user's intention
- Answer Generation
  - Generate the system's sentences most appropriate at each state of the dialogue

Answer Generation

- Generation of sentences to achieve the goals the dialogue manager has selected
- Tasks
  - Content selection: what has to be said
    - Belongs to the discourse plan
  - Superficial realization: how has to be said
    - Presenting content correctly

Content Selection

- Determine the content of the system sentences in order to achieve the goals
- Examples:
  - *Madagascar is not shown in Sant Cugat [Nucleus]*
    - *It is shown in Barcelona [Satellite]*
  - *Would you like a suite? [Nucleus]*
    - *It is the same price than the doble room [Satellite]*
  - *Magic Flaute is not shown this year at Liceu [Nucleus]*
    - *But Figaro Wedding is [Satellite]*

Content selection (Research)

- Knowledge Bases
  - Domain knowledge
  - User believes
  - User model: preferences, language.
  - Dialogue history
- Mechanisms for content selection
  - Schemes - patterns
    - First object name, then attributes
  - Rules
  - Plans
  - Reasoning

Content selection (Real systems)

- Knowledge sources
  - Domain knowledge
  - Dialogue history
- Strategies pre-defined for content selection
  - Only nucleus, not satellite
  - Nucleus + satellite fixed

Superficial realization

- Goal: to determine **how** content selected is presented
- Examples:

  Madagascar is shown at CINESA cinema in Sant Cugat

- Tasks
  - Construction of phrases
  - Lexical selection

Superficial realization (Research)

- The generator input is
  - Semantic representation
  - Phrase structures
- The generator uses a grammar and a lexicon for generating the sentence

Superficial realization (Real systems)

- Predefined (canned) sentences
  - Sentences to achieve specific goals
    - Initial and final sentences
    - Ask the user to repeat
  - Specially appropriate for speech
- Patterns
  - Patterns for goals
    - Notification: You have been assigned number X.
    - Information: A,B,C and D are shown at cinema E.
    - Clarification: Did you said X or Y?

Components of spoken dialogue systems

Voice input: From acustic signal to meaning

Conversion of the signal to a set of words

Obtention of meaning from words

Voice output: From content to acustic signal

Conversion of content to text sentences

Conversion from text to signal

Dialogue Manager

Voice input: From acustic signal to meaning

Transforming the signal to a set of words

- Disfluences
  - Pauses, expressions like: *umm*, *aah*
  - Fragments of words
  - Models using pauses
- Words that do not appear in the lexicon used
  - Models that can learn new words
- Mobile: more difficult

Voice input: From acustic signal to meaning

Obtention of meaning from words

Probabilistic model (i.e. n-gram)

It specifies the probability of a particular word once previous word has been recognized

- It is not understanding
- Probability of previous word is not one : many alternative hypothesis
  - Example: *Euthanasia* o *youth in Asia*
- Not gramaticaly correct, fragments, disfluencies

Voice input: From acustic signal to meaning

Obtention of meaning from words

- NLP is based in a complete syntactic analysi
- The approach in voice is different
    - Many recognition mistakes, unknown words
- Semantic approach (in restricted domain systems)

VOICEXML: A VOICE STANDARD

- System development easy and fast
  - Low level details are transparent
- Internet access
  - Client/Server architecture
- Multilinguality
  - More than a language

Example: Català and Castellà can be mixed

“ Plaza Sants”

“ Calle Manuel Girona”

VOICEXML: A VOICE STANDARD

Limitations

- Only voice
  - Touch-tone DTMF
- The dialog has to be defined for each application
- System initiative
  - No user initiative

Design of dialogue systems

Gould and Lewis principles (1985)

- Study users and application tasks
  - Interviews to users
  - Obtaining person-person dialogues
- Development of a prototype
  - Using ***Wizard-of-Oz*** method. A person substitutes the machine
- Interactive design
  - Users have to prove the system. Incorporation of new information

Evaluation

- Goal: to determine how well the system is working
- Difficulties
  - Determine correct and incorrect dialogues
  - Comparing strategies and dialogs
  - Metrics selection
    - Efficiency versus correctness
  - Determine the relationship between different metrics
    - Long or short dialogues?
  - High cost

Evaluation

- Only system’s initiative
  - More successful dialogues
  - Less recognition errors
- Only user’s initiative
  - More natural
  - Shorter (advantage?)
  - More subdialogs for detecting errors

Evaluation

- Evaluation paradigm
  - Evaluation of the final result success
  - Evaluation of the final result success and also the process
    - Different metrics for different components
    - Only one function to evaluate the set

Evaluation

Evaluation only of the success of final result

- Appropriate for question/answering systems
- Easy to define correct answers
- For each question
  - Obtaining the correct answer
  - Obtaining system's answer
  - Comparing answers
  - Quantification of the processing of the system
- Advantages: simple
- Disadvantages: ignore other important

Evaluation

Evaluation of the final result and the process

Different metrics for different components

- Voice recognition
  - errors in word recognition (WER)
- Interpretation: attribute-value matrix
- Dialogue Manager
  - Quality of system's responds
  - Strategies for recovery of errors
- Dialogue system
  - Success of final result
  - Number of turns
  - Time

Evaluation

Different metrics for each component

- Advantages:
  - Considering all the process to complete the task
- Limitations:
  - The metrics may not be independent from each other
  - Difficulties for comparing different dialogue systems

Evaluation

Only one function to evaluate all the process

- PARADISE [Walker et al]
- Maximizes user's satisfaction
  - Maximizes task success
  - Minimizes cost
- Efficiency measurements
  - Number of interventions
  - Waiting time
- Quality measurements
  - Ratio of errors recovery

Evaluation

PARADISE

- Function of measurements
  - Values on user satisfaction
    - Questionnaires
  - Values of several metrics
  - Applying multiple linear regression to obtain a function that relates user satisfaction and other metrics

Evaluation

- Advantages
  - It compares different systems
  - It specifies the contribution of each system component to the global performance
  - It can be used for predicting future versions
- Disadvantages
  - High cost for obtaining the function
  - High cost for adapting the function to other systems.

Formas lógicas utilizadas en SISCO

¿Cúal es el caudal del rio Ebro?

$\longrightarrow$

preg(Y, ex(rio1(X)&X=ebro,caudal(Y,X)))

¿Dónde desemboca el rio Ebro?

$\longrightarrow$

preg(X, ex(rio1(Y)&Y=ebro,desemboca(Y,X)&lugar(X)))

An example of conversation in a Dialogue System

C		
D	{tema: bienvenida }	
Interc	{objetivo: saludar}	S> Welcome to the informaton service, what do you want
Interv		
D	{tema: viaje en tren de Barcelona a Valencia}	
Interc	{objetivo: petición de una información}	
Interv		U> Quisiera ir de Barcelona a Valencia
Sub	{tema: fecha del viaje}	
Interc	{objetivo: precisión}	
Interv		S> ¿En qué fecha?
Interv		U> el martes que viene
Sub	{tema: hora del viaje}	
Interc	{objetivo: precisión}	
Interv		S> ¿qué horario, mañana o tarde?
Interv		U> a primera hora
Interv		S> Hay 3 trenes, el Euromed ....
Interc	{objetivo: petición de una información}	
Interv		U> ¿Cuánto vale el billete en segunda?
Interv		S> 8000 pesetas
D	{tema: despedida}	
Interc	{objetivo: despedirse}	
Interv		U> Gracias, buenas tardes
Interv		U> Gracias a Vd, buenas tardes