CMU Communicator

Overview
Content

- Gentner/Gentner Emulator
- Sphinx/Listener
- Phoenix
- Helios
- Dialog Manager
- Datetime
- ABE
- Profile
- Rosetta
- Festival
Gentner/Gentner Emulator

- Assistive Listening Systems (ALS)
- Generates Session
- Listens to call or application
Sphinx/Listener

- Produces Hypothesis
- Adds Confidence
- Interaction with other components
- Stops Speech output on Barge-in
- Besides that: works as known
Phoenix

- Parses Hypothesis in semantic grammar
- Multiple Parses
- Domain specific
Helios

• "post-parser".
• Parse the level of confidence to the dialog levels of the system.
• Send annotated parse to dialog manager.
• makes use of learned parameters to assign confidence
• no domain-specific components
• locus for multi-modal integration
(Thus the Communicator system does not use multi-modal input)
Dialog Manager

- Carnegie Mellon AGENDA dialog manager.
- Two parts:
  - execution engine
  - Domain specific handler library
- assembles a dynamic dialogue tree over the course of a session.
- manages the dialog agenda
- Handlers are implemented as C++ objects
- incorporate logic for interpreting particular inputs
- Interacting with domain agents
- system does not follow a dialog "script"
- The Dialog Manager focuses on the task and discourse aspects
- performs minimal domain-specific reasoning,
Datetime

- interprets temporal expressions
- knowledge of holidays in the United States.
- no full coverage for religious holidays
- operates on fragments of the input parse
- makes use of context information (current time and date)
- was developed specifically for the travel domain
- date and time appear to be largely domain-independent
ABE

- ABE performs domain-specific functions
- Application to access information:
  - system database,
  - retrieval of information on the web
  - domain-specific reasoning.
- Live information retrieval
- domain-specific reasoning ("Is that Portland in Maine or Portland in Oregon?")
- managing solution sets (ranking flights on "desirability")
- Contacts database of world knowledge (e.g. meaning of yes and no)
Profile

• Manages user profiles
  – notes preferences
    (for airlines or hotels; where to email confirmation of a itinerary, calling frequency, etc.)
  – used to manage personalization
• Inactive on default
Rosetta

- language generation module
- receives semantically coded requests from the dialog manager
- computes a corresponding word string
- two generation strategies:
  - templates
  - stochastic
- Templates make use of particular domain entities (for example airport names).
- stochastic generation makes use of language models
- greetings or error notifications are handled by template.
Festival

• Speech synthesis
• domain mode (limited Vocabulary, recorded for this domain)
• concatenative synthesizer
  – database of recorded human speech
  – create requested outputs by selecting appropriate units
  – combining these by splicing.
Weiteres zu VXML

• **OpenSource Projekte zu VXML**
  – PHP Voice
  – LSRHS VXML Browser (early stage)
Quelle