Approaches to Dialogue Systems and Dialogue Management

David Traum
Institute for Creative Technologies
University of Southern California
traum@ict.usc.edu
http://www.ict.usc.edu/~traum

Outline

- Basic Terms & Overview of Dialogue Structure
- Example Systems
- Dialogue System Components
- Dialogue Management Approaches
  - Structural
  - Principle-based

NL Dialogue Overview

- Communication involving:
  - Multiple contributions,
  - Coherent Interaction
  - More than one participant
- Interaction modalities:
  - Input: Speech, typing, writing, menu, gesture
  - Output: Speech, text, graphical display/presentation, animated body

Dialogue terms

- Dialogue Modelling
  - Formal characterization of dialogue, evolving context, and possible/likely continuations
- Dialogue system
  - System that engages in a dialogue (with a user)
- Dialogue Manager
  - Module of a system concerned with dialogue modelling and decisions of how to contribute to dialogue
  - Cf speech recognizer, domain reasoner, parser, generator, tts,…

Roles for Dialogue Systems

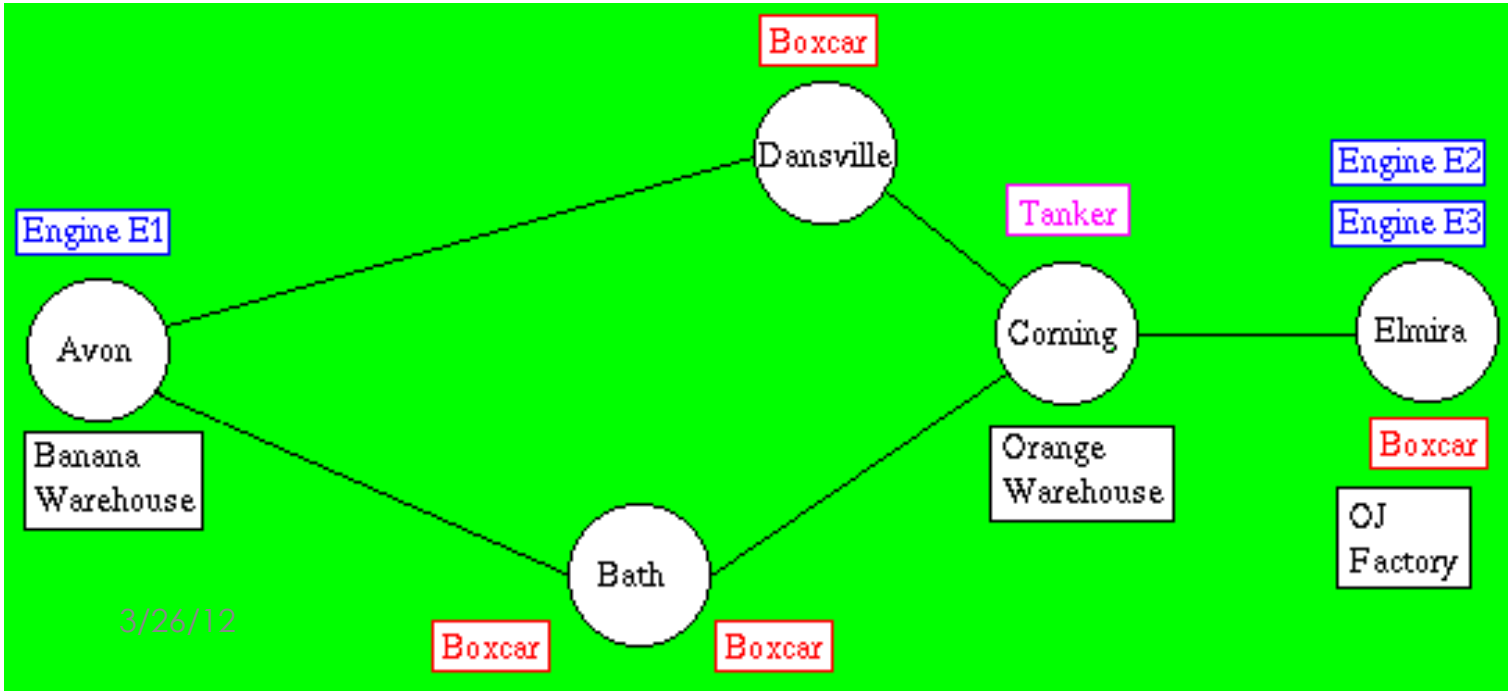
- Information provider
- Advisor
- Service provider
- Collaborative partner
- Tutor
- Instruction-giver
- Conversational Partner
- Competitor
- Antagonist

Dialogue Systems: State of the Art

- Deployed Commercial Systems
  - Call routing/call center first contact
  - Voice menus
  - Simple information tasks (Siri)
  - in car navigation & services
- Useful systems
  - Command & control
  - Language tutoring
  - Immersive Training
- Advanced Research Prototypes
  - Collaborative systems
  - Adaptive systems
  - Multi-modal & Robot systems
  - Companion Systems
  - Non-cooperative role-playing agents

Rochester: TRAINS-93
(Allen et al 1994)

```
Manager: We better ship a boxcar of oranges to Bath by eight a.m.        (1.1)
System: OK                                                               (2.1)
Manager: So we need to get a boxcar to Corning, where there are oranges. (3.1)
         There are oranges at Corning.                                   (3.2)
         Right?                                                          (3.3)
System: Right                                                            (4.1)
Manager: So we need an engine to move the boxcar.                        (5.1)
         Right?                                                          (5.2)
System: Right                                                            (6.1)
Manager: So there's an engine at Avon.                                   (7.1)
         Right?                                                          (7.2)
System: Right                                                            (8.1)
Manager: So we should move the engine at Avon, engine E1, to Dansville
         to pick up the boxcar there                                     (9.1)
System: Okay                                                             (10.1)
Manager: and move it from Dansville to Corning                           (11.1)
         load up some oranges in the boxcar                              (11.2)
         and then move it on to Bath                                     (11.3)
System: Okay                                                             (12.1)
Manager: How does that sound?                                            (13.1)
System: That's no problem                                                (14.1)
Manager: Good                                                            (15.1)
```

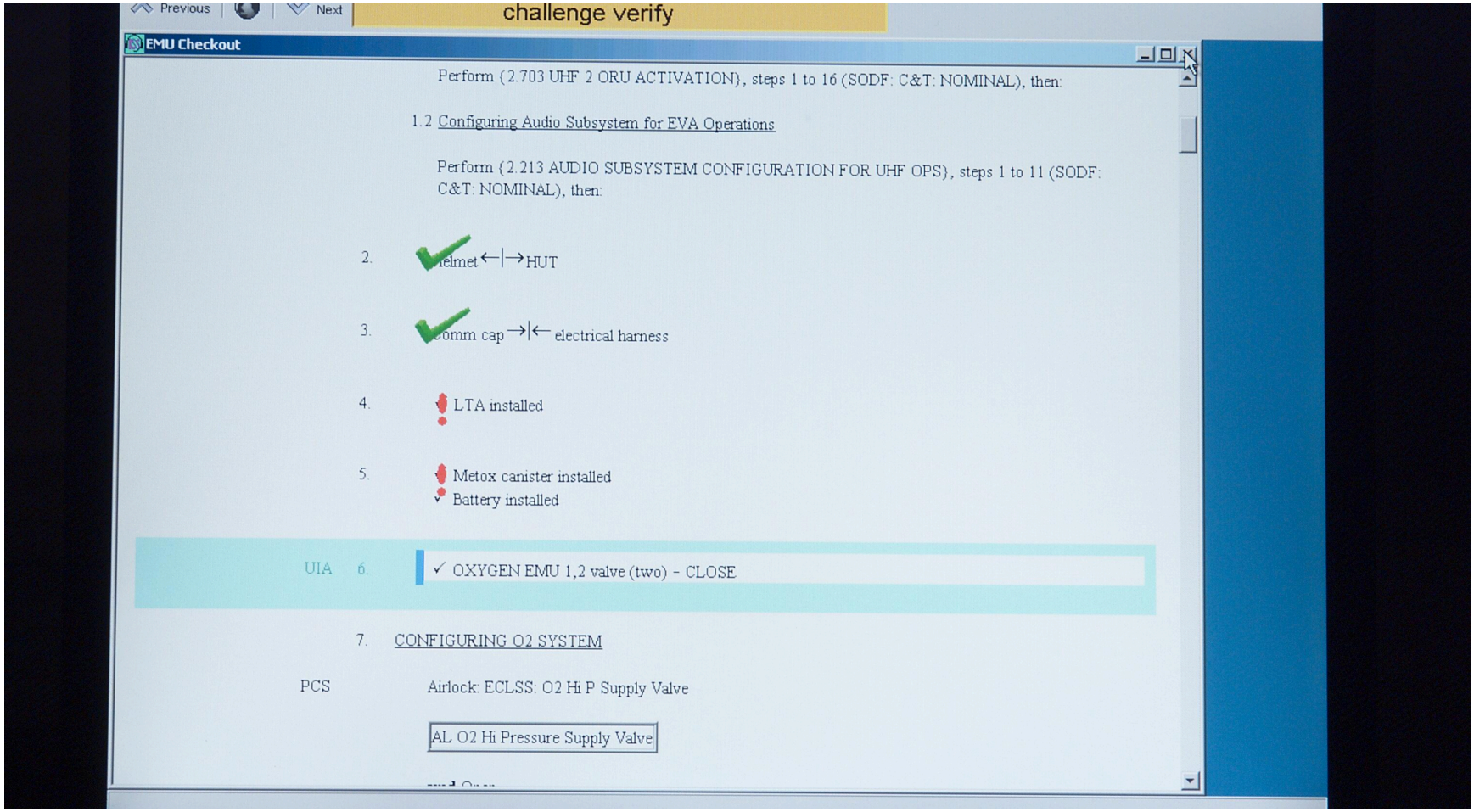


MERL: Sample Collagen System: VCR Help

(Rich, Sidner & Lesh 2001)

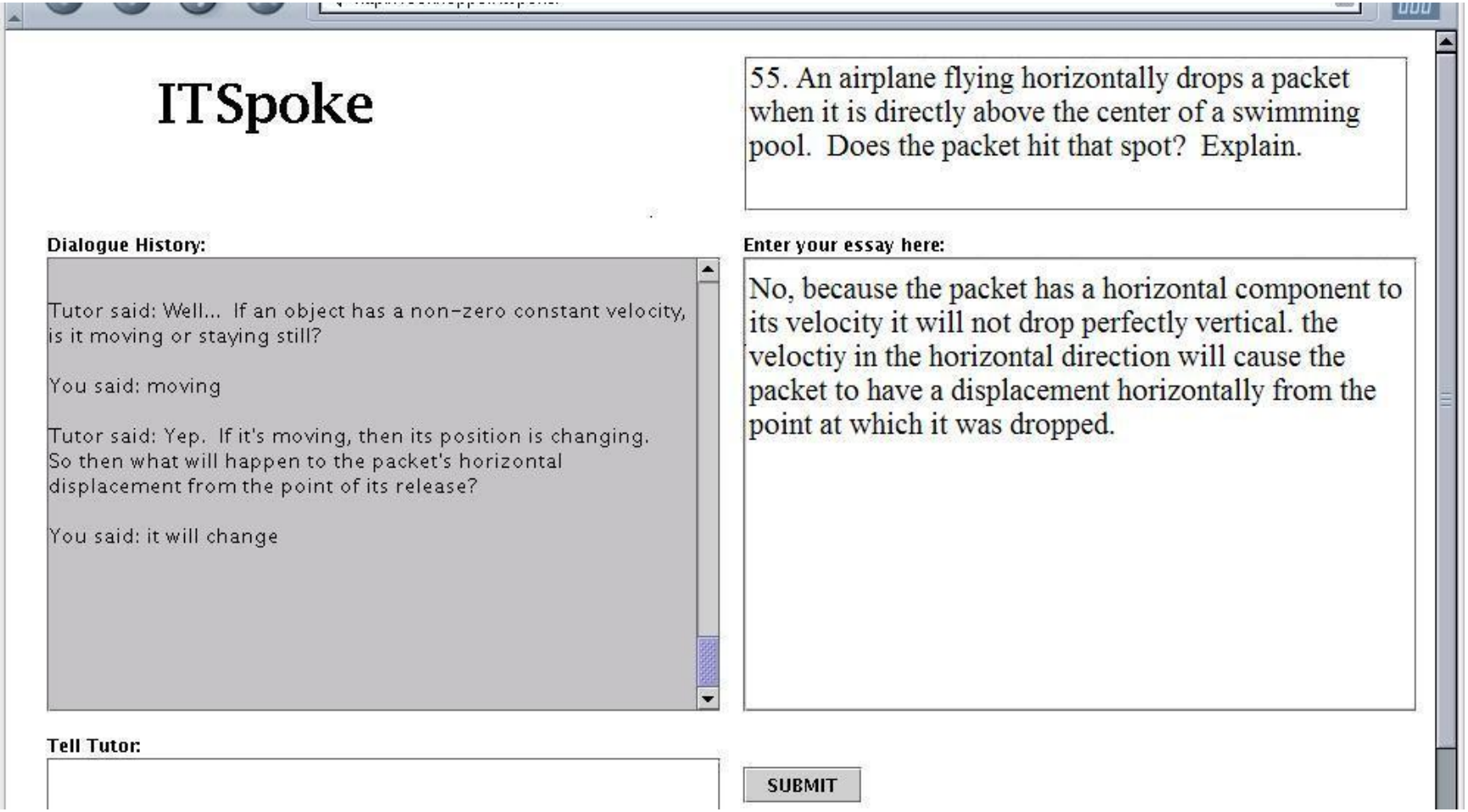
1 User says "Hello."
2 Agent says "Hello, what's your name?"
3 User says "Candy."
4 Agent says "Hello, Candy!"
5 User says "Let's set the clock."
6 Agent says "What is the time?"
7 User says "The time is 11:00 AM."
8 Agent sets the clock to 11:00 AM.
...
9 User says "How do I play back a video from my camcorder?"
10 Agent says "First, plug in the camcorder cables."
11 User says "How do I plug in the camcorder cables?"
12 Agent points to where you plug in the camcorder cables.
13 Agent says "The red cable plugs into the red terminal, the black cable plugs into the black terminal, and the yellow cable plugs into the yellow terminal. Make sure the other ends of the cables plug into the corresponding colors on the camcorder."
14 User says "What next?"
15 Agent says "Now press Play on your camcorder. You should see a picture on the TV screen."
16 User says "Ok."

NASA Clarissa System
(Rayner et al 2005)

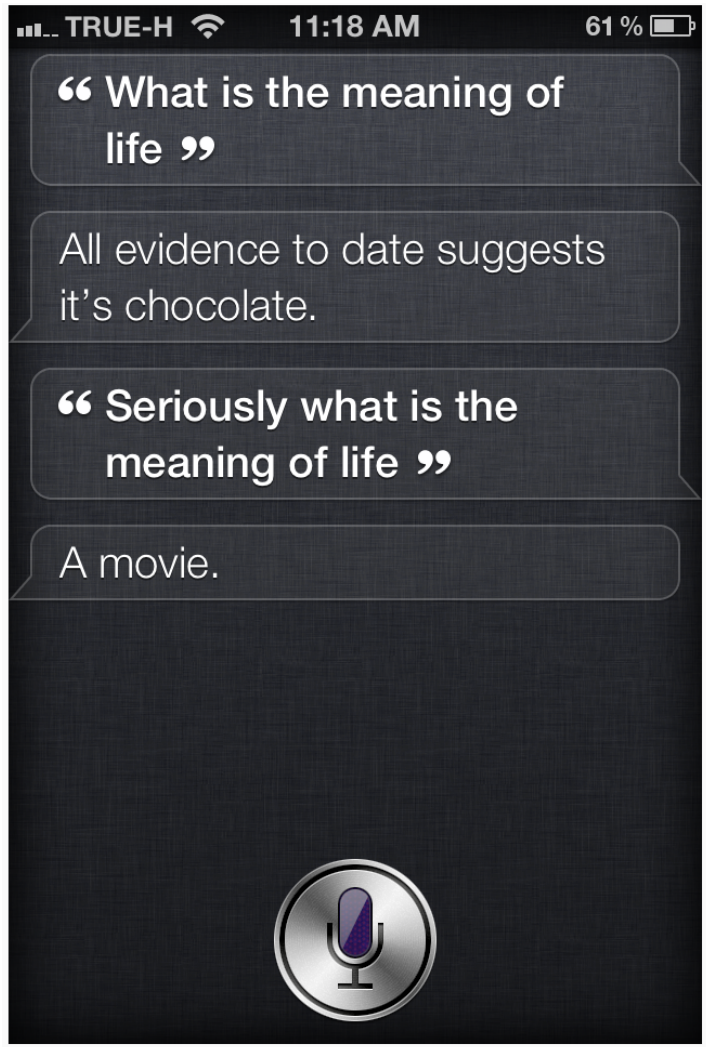


U Pittsburgh: ITSpoke

(Litman & Silliman 2004)

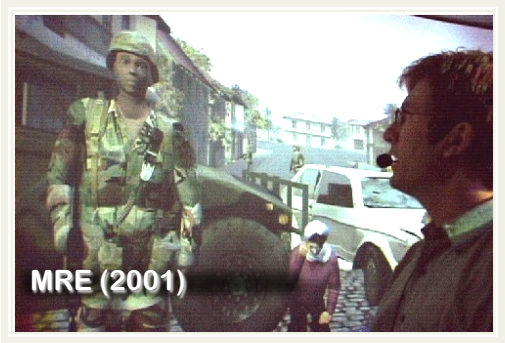


Siri


TRUE-H
11:18 AM
61 %
“What is the meaning of life”
All evidence to date suggests it’s chocolate.
“Seriously what is the meaning of life”
A movie.

ICT Virtual Human Negotiation:

Capability-advancing prototypes



Decision-making

Multi-party Interaction

Non-cooperative Negotiation

Multi-party negotiation

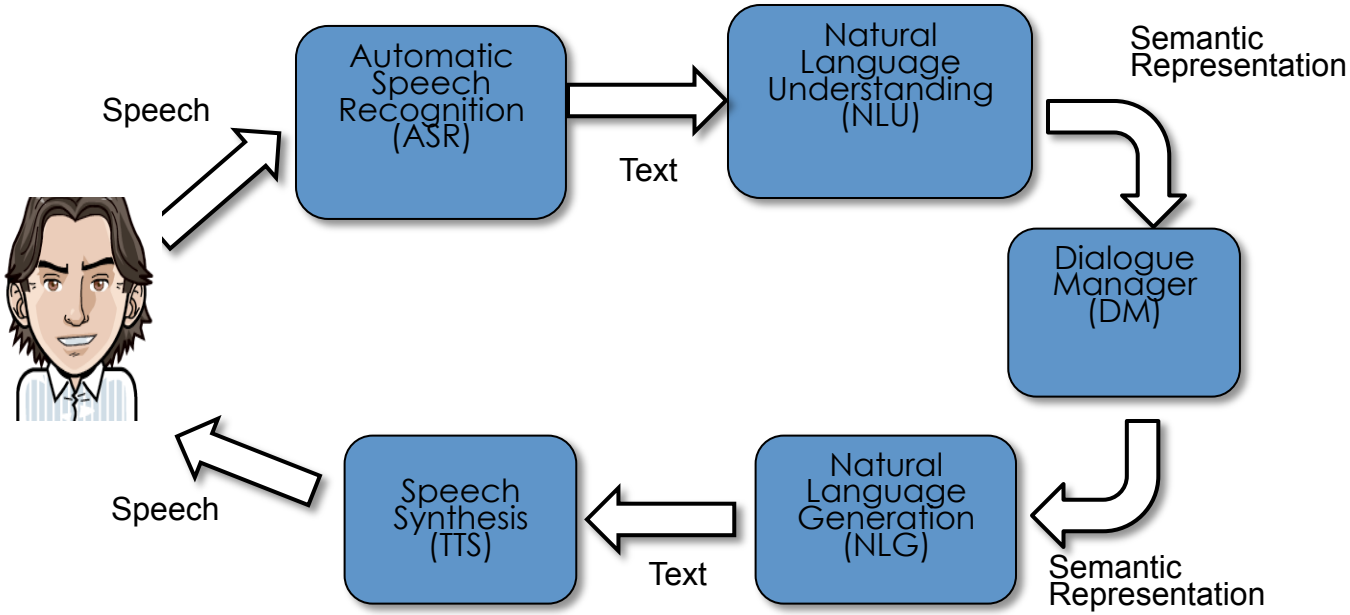
Persuasion and Conflict resolution

SDS Components

- Architecture
- Back-end/Domain Reasoners
- Input Interface (Audio, Keyboard,etc)
- Interpretation (internal representation)
- Dialogue Management
- Generation
- Output Interface

Dialogue Modules & Architecture

- Standard Pipeline Architecture



Interpretation: Speech Recognition

- Phases
  - Signal Processing
  - Acoustic Model, tri-phones
  - Language Model (N-grams)
- Issues
  - Small or large vocabulary
  - N-gram or grammar-based language model
  - Integrated or pipelined understanding
  - Output (concepts, n-best word list, lattice)
  - Unified or State-specific recognizers

Interpretation: Parsing/Semantic Representation

- Tasks
  - Retrieval/Classification
  - Understanding/Extraction
- Output
  - Response
  - (aspects of) Meaning (e.g., semantic roles, speech acts, parse)
- Styles
  - Key-word
  - Language model
  - Grammar-based
  - Concept-based (semantic parser)
  - Expectation-driven
- Spoken Dialogue vs. Written text
  - Utterance length, grammaticality, interactivity, repairability, transience, …

Generation & Synthesis

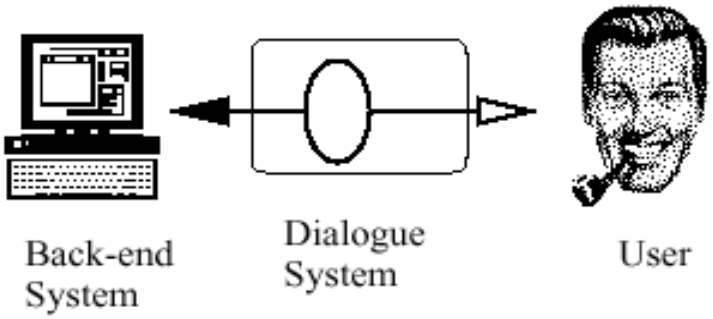
- Generation
  - Output
    - Text
    - Prosodic cues
    - multimodal generation
  - Method
    - Fixed text
    - Template-based
    - Sentence Planning & Realization
      - Grammar-based
      - Statistical Language model
- Synthesis
  - Voice Clip, or TTS
  - TTS or Concept to Speech

Dialogue Management Tasks

- Maintaining & Updating Context
- Deciding what to do next
- Interface with back-end/task model
- Provide expectations for interpretation

Two Approaches to Dialogue Systems:

front-end



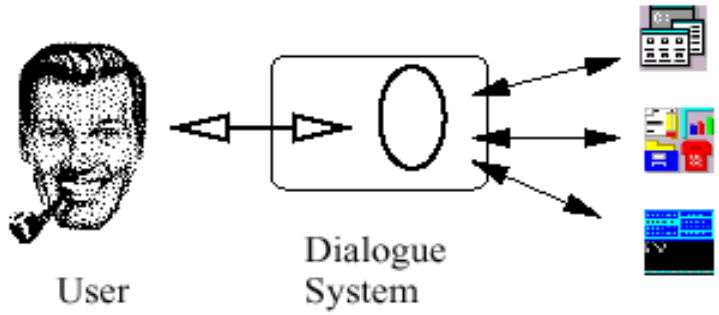
System as "translator" between user and backend system

Examples: Sundial (European Train Info), MITRE systems, MIT Galaxy

Key desgin question: how to provide back-end with understandable messages (in a manner natural to the user)?

Key run-time question: what messages should be sent to back-end (or generated to user)?

agent



System as "homunculus", with access to task-specific functionality

Examples: TRAINS (Rochester), Circuit-Fixit (Duke), Artemis (France Telecom)

Key design question: how to coordinate with the user to accomplish a task?

Key run-time question: what should be done now (given context, inputs, goals)?

(Cf. Larsson 2005: Interfaces vs Simulations)

Using Data

- Corpus Collection
  - Human-Human
  - Wizard of OZ
  - Human-System
- Annotation
  - Coding Scheme
  - Coding
    - Automatic
    - Tool-assisted
    - Inter-coder Reliability (Kappa)

Evaluation

- Objective Metrics
  - Task success
  - Resources used (time, turns, attention,..)
- Subjective Evaluation
- Issues
  - On-line vs off-line
  - Black Box vs. Glass Box
  - Class of User (Expert, Novice)
  - Feedback into system design

Dialogue Manager Organizing Principles

- Structure-based
  - Script
  - Local
    - Exchange
    - Word-based
      - Keyword spotting
      - Advanced techniques
        - » AIML recursion
        - » Statistical Language model
    - Meaning-based
      - Speech acts
  - Grammar
  - Tree/FSM
- Principle-based
  - Frame
  - Logic
  - Plan
  - Information-State

Speech Acts

- How to "Do things" with words
  - Look at actions & effects of utterances rather than truth-conditions
  - Types of acts
    - Locutionary
    - Illocutionary
    - Perlocutionary

Searle's Types of Illocutionary acts

- Representatives
- Directives
- Commissives
- Expressives
- Declarations

Issues for computational theory of speech acts

- When can an act be recognized (as sincere and successful)?
- What are the effects of performance of an act
  - On state of hearer and speaker
  - On state of dialogue
- When should act be performed?
- How should act be performed?

Dialogue Structure

- Local
  - What binds utterances together?
  - How is one utterance (in)coherent when following another?
- Global
  - What is the structure of a conversation?
  - What is the structure of a task (that a conversation is "oriented" to)?
- How many levels of structure are there?

Local Dialogue Structure

- Utterances organized in turns
- Coherence between turns (or utterances)
  - Adjacency pairs
  - Exchange structure
  - IR(A) units
  - Games

Turn-taking (Sacks & Schegloff)

- Turns composed of one or more smaller utterance units (Turn Constructional Units = TCUs)
- Transition relevance places (TRPs)
- Signals of TRP (and pre-trp)
- Self and other selection

Adjacency Pairs (Schegloff & Sacks)

- Sequences with features
  - Two utterance length
  - Adjacent
  - Produced by different speakers
  - Typology in production
    - Pair type including First part & second part
      - E.g., Q&A, greeting-greeting, offer-acceptance
- Orientation towards Adjacency Pair
  - Conditional relevance
  - Preferred & dispreferred 2nd parts
    - Hesitations, apologies & qualifications
  - Repairs and apologies

Speech-act related Adjacency Pairs

- Question-Answer
- Propose-Accept/reject/challenge,…
- Offer-accept/decline
- Compliment-refusal/thanks
- Greeting-greeting

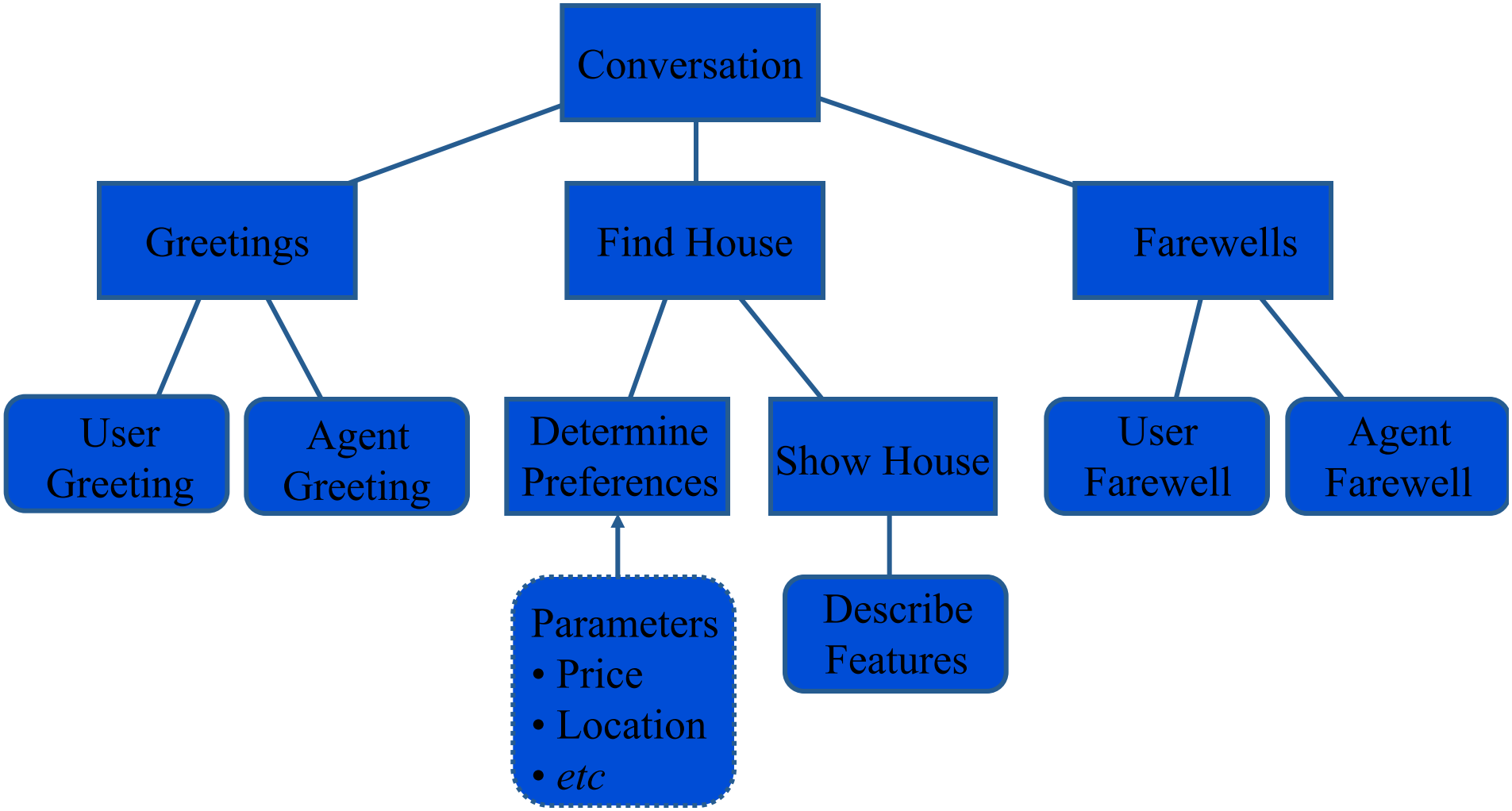
Global dialogue structure

- Conversation phases
  - Opening
    - Engagement
    - Greetings
    - Preambles/agendas
  - Body
    - Topics
      - Topic relations
  - Closing
    - Pre-closings
    - Termination bids
    - farewells

Task Structure (Grosz & Sidner '86)

- Hierarchical & sequential tasks
  - Linear precedence
  - Immediate dominance
- Topic stack
- Topic transitions
  - Push
  - Pop
  - Pop-push

Plan Tree for REA (Cassell et al)



Topic changes when focus stack changes.

Simple Organizational Structures

- Script
- Local
  - Exchange
    - Word-based
      - Keyword spotting
      - Advanced techniques
        - » AIML recursion
        - » Statistical Language model
    - Meaning-based
      - Speech acts
- Grammar
- Tree/FSM

Eliza

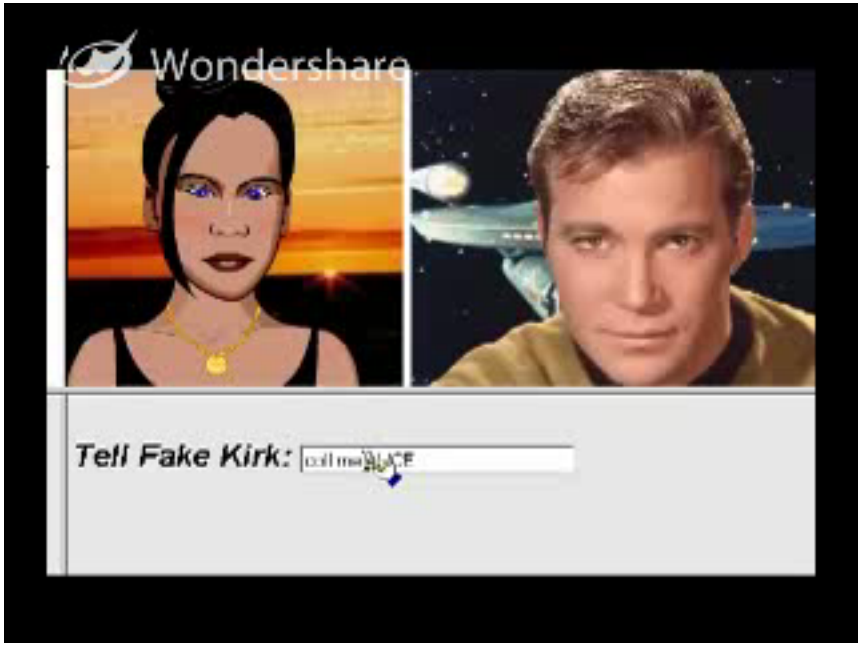
Weizenbaum, CACM 1966

- Local organization
- Produce response based on analysis of input
  - Keyword spotting
  - Pattern recognition
  - Pattern selection
  - Transformation rules
- Example: emacs Doctor program
- Example 2: CL simple-eliza rules
- http://hampshire.edu/lspector/courses/eliza-simple.lisp

Advanced Patterns: AIML

http://www.alicebot.org/aiml.html

- XML Syntax
- Stimulus-response interaction
- Categories
  - Pattern
  - Template



Statistical Retrieval-based Dialogue

- Basic idea: use IR-like techniques to find the correct response to an initiative
- Applications
  - Call routing (Chu-Carroll & Carpenter 1999)
  - Question-answering character (Leuski et al 2006)
  - FAQ systems (Marom & Zuckerman 2007, Selberg & Jonsson 2008)

Example: Sgt Blackwell (Leuski et al 2006)

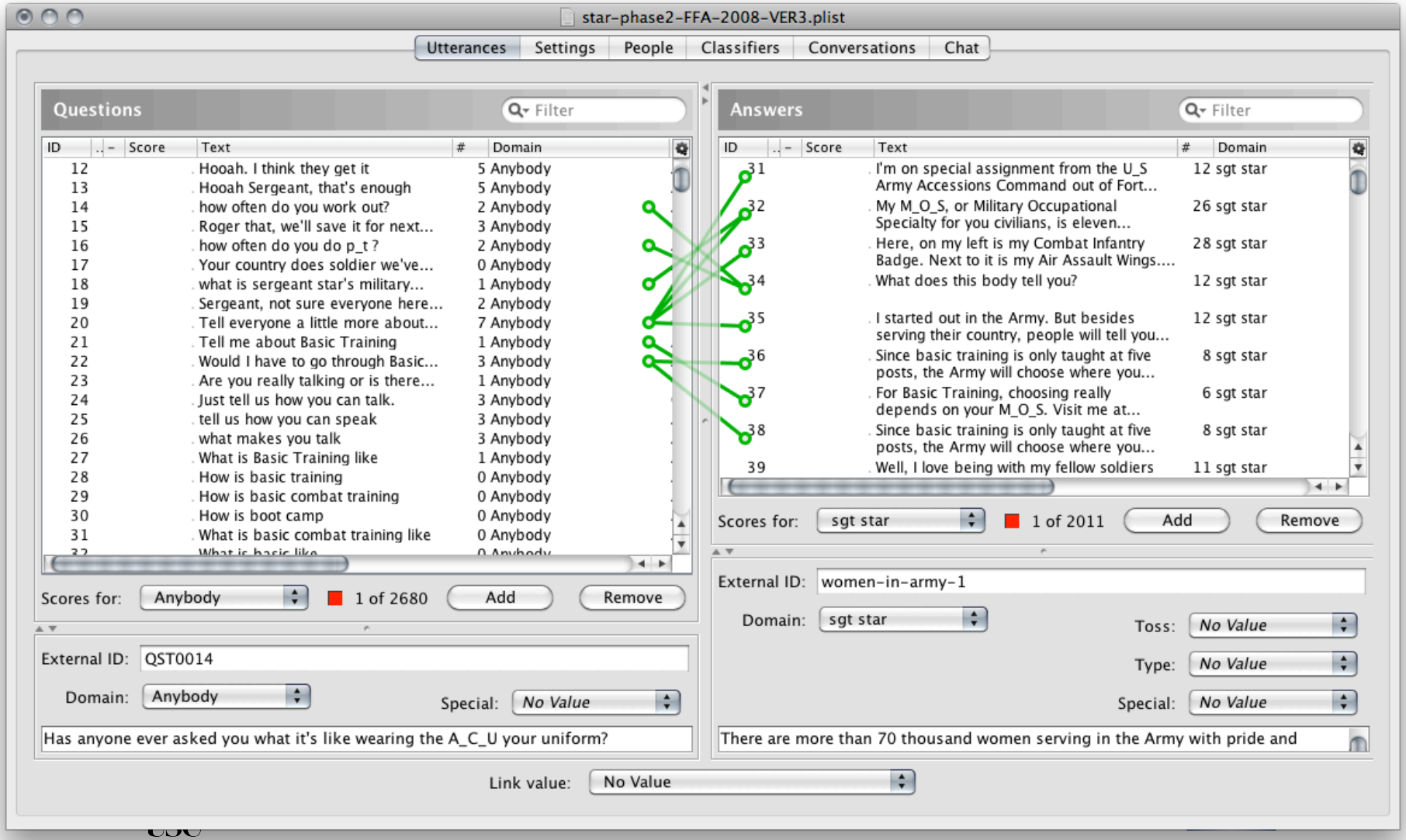
- Focus: technology demo
- Highlights:
  - Life-sized, mixed reality
    - Trans-screen
  - High-production quality
    - Rendering (> 60K polygons)
    - Voice
    - Authored Text
    - Robust responsiveness
      - Speech recognition and speech and non-verbal reply
  - Limited domain of interaction: responding to interview/Q&A

Sgt Blackwell Video

Sgt Blackwell "Dialogue Model"

- Set of pre-constructed answers
  - In domain
  - Off-topic
  - Prompt
- Local history
- IR-based classification
  - Given possibly unseen question, map to best answer

NPCEditor: Authoring Visualization Links from questions to Answers



NPCEditor: New runtime question answer visualization

twins-combined.plist

Utterances | Settings | People | Classifiers | Conversations | Dialog Manager | Chat

2. Question

Anybody asks The Twins

what do you do here girls

Clarity A: 1.4004 Clarity Q: 1.3544

Enter Question (^+⏎)

3. Similar Questions

Similarity	Text
0.9189	What do you do here
0.0203	do you live here
0.0166	do you like boys or girls
0.0139	where comigo eporimimo blah you...
0.0086	Do you like working here

4. Linked Responses

Importance	Text
0.9199	We're museum guides! We help visitors like you understand more about computers, robots, and...
0.0204	One of the computers behind us is our current home, but we move a lot. Actually, we can be many...
0.0168	Why?? Is there another virtual human that you think we should meet? I hope they're into neuro networks...

5. Answer Words from Respo

Importance	Term
0.0734	guide
0.0696	understand
0.0667	communications
0.0592	visitor
0.0589	we

6. Responses Matching the Answer Words

Score	Text
-4.0	We're museum guides! We help visitors like you understand more about computers, robots, and communications.
-5.5	My name's Ada. And I'm Grace. We're your Virtual Museum Guides. With your help, we can suggest exhibits that will get you thinking! Or answer...
-5.7	Cahners ComputerPlace is named after Norman Cahners, a publisher

7. Selected Response

We're museum guides! We help visitors like you understand more about computers, robots, and communications.

1. Conversation Log

<import id="artist_animators_our_existance_2_A"/>

The Twins: The Robot Arena is where the white car is. Just look to your right.

The Twins: Computer World is just to our left.

The Twins: At Computer World, young learners can explore a computer by walking through it.

The Twins: Well, I know that in eightteen seventy four Alexandra Graham Bell invented the telephone here in Boston. And now we have cell nation here for you to explore.

Anybody: what do you do here girls

Ada: Hey there!

Grace: How can we help you?

Anybody: what do you do here girls

Grace: We're museum guides!

Ada: We help visitors like you understand more about computers, robots, and communications.

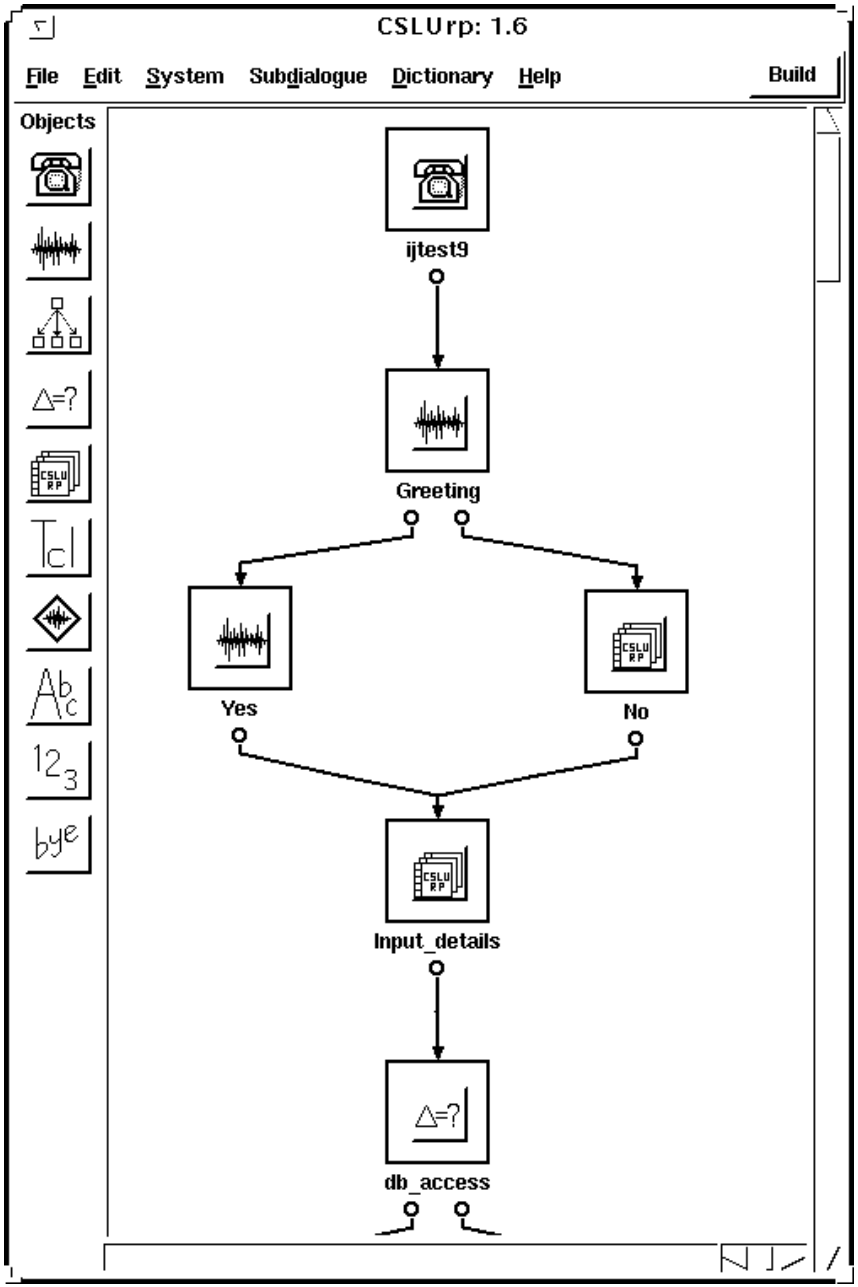
Dialogue Grammar

- Specify set of legal moves to be a “legal” dialogue
- Specify set of moves at any given point
- Specify context update

FSM Dialogue model

- Set of states
- Set of moves from each state
- Transitions to new state

CSLUrp Interface



Transaction Dialogues

- User has a request
- System needs info from user to process request
- Dialogue proceeds as:
  - User specifies request
  - System gathers necessary info
    - Q&A
    - Spontaneous assertion from user
  - System looks up information & provides response

Frame-based Approach

- Also called form-based (MIT)
- Central data structure is frame with slots
  - DM is monitoring frame, filling in slots
- Used for transaction dialogues
- Generalizes finite-state approach by allowing multiple paths to acquire info
- Frame:
  - Set of information needed
  - Context for utterance interpretation
  - Context for dialogue progress
- Allows mixed initiative

Example: MIT Wheels system

- Domain: searching used car ads
- Transaction domain + constraint satisfaction
- No slots are mandatory,
  - try to find the best set of matches
  - Try to find an appropriate # of matches

Example: MIT Jupiter System (1)

- Retrieval of weather forecast domain
  - Multiple sources
  - Content processing
  - Information on demand
  - Context
- 1-888-573-8255

Trindi: Information State Theories of Dialogue

- Statics
  - **Informational components** (functional spec)
    - e.g., QUD, common ground, dialogue history, ...
  - **formal representations** (acessibility)
    - e.g., lists, records, DRSes, …
- Dynamics
  - **dialogue moves**
    - abstractions of i/o (e.g., speech acts)
  - **update rules** - atomic updates
  - **update strategy** - coordinated application of rules

Hot Topics for Dialogue Systems

- Dialogue System Evaluation
- Grounding & Robustness
- Incremental dialogue processing
- Enculturated Dialogue Agents
- Adaptive Dialogue Systems
  - Learning
  - User Modelling
- Multi-modal dialogue
- Multi-task dialogue
- Multi-party dialogue
- Non-cooperative dialogue agents
- Dialogue "in the wild"
- Long-term Dialogue Companions

To Find out More

- Sigdial
  - www.sigdial.org
  - Annual meetings
- Lists of dialogue systems
  - http://www-2.cs.cmu.edu/%7Edbohus/SDS/
  - http://www.ling.gu.se/~sl/dialogue_links.html
- Book: "Spoken Dialogue Systems" By Kristina Jokinen, Michael McTear