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

Course web page:

http://www.ict.usc.edu/~traum/ESSLLI08

Outline for Course

- Yesterday: Introduction, Architecture of Dialogue Systems, Example Systems
- Today: Simple structures: S-R, IR, finite State
- Tomorrow: Frame-based and Information State
- Thursday: Plan-based and Logic Based
- Friday: Advanced Topics

Outline for Today

- Disalogue Structure
- Simple Control structures
  - Scripts
  - Keyword-based
- Eliza & AIML Chat systems
- Information-Retrieval based systems
- Speech Acts
- Finite State systems
  - Example
  - Toolkits
    - CSLURP RAD
    - Voice-XML

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

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

Other Local organizations

- Exchange
- IR Unit
- Game
  - Can be more than two utterances in sequence

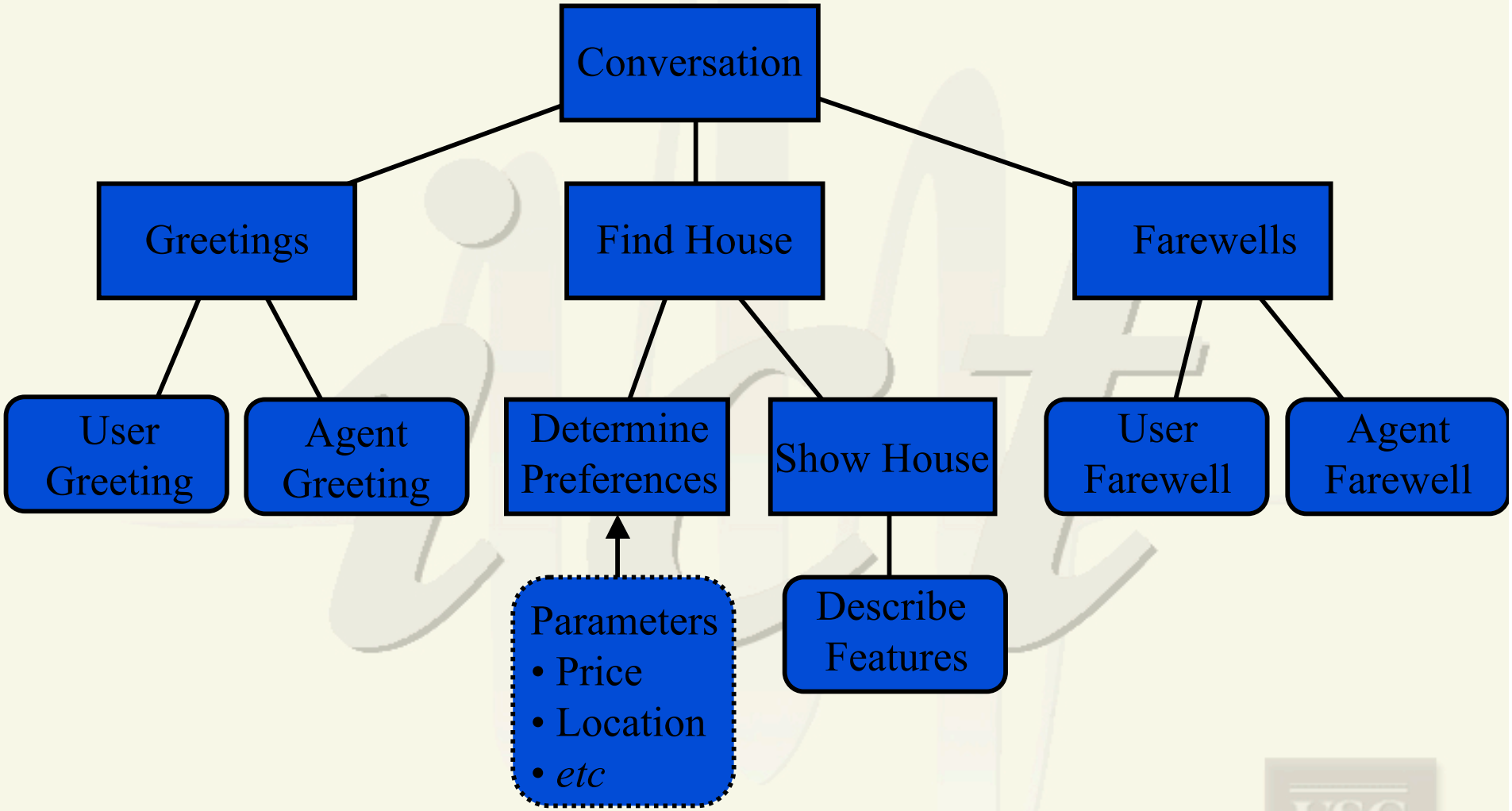
Global dialogue structure

- 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

Example Script: Scene 1 Monty Python & the Holy Grail

- ...
- **ARTHUR:** Well, it doesn't matter. Will you go and tell your master that Arthur from the Court of Camelot is here?
- **SOLDIER #1:** Listen. In order to maintain air-speed velocity, a swallow needs to beat its wings forty-three times every second, right?
- **ARTHUR:** Please!
- **SOLDIER #1:** Am I right?
- **ARTHUR:** I'm not interested!
- **SOLDIER #2:** It could be carried by an African swallow!
- **SOLDIER #1:** Oh, yeah, an African swallow maybe, but not a European swallow. That's my point.
- **SOLDIER #2:** Oh, yeah, I agree with that.

Eliza

- 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

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

EXAMPLES:

```
<category><pattern>BYE</pattern>
<template><random>
<li>See you later <get name="name"/>.</li>
<li>Goodbye <get name="name"/>.</li>
<li>Until next time <get name="name"/>.</li>
<li>Thanks for chatting, <get name="name"/>.</li>
<li>See you later <get name="name"/>.</li>
</random></template> </category>

<category><pattern>AU REVOIR</pattern>
<template>
<srai>BYE</srai>
</template></category>

<category><pattern>FAREWELL *</pattern>
<template>
<srai>BYE</srai>
</template></category>
```

AIML: Advanced

- Srai operator
  - Synonyms
  - Splitting patterns
  - conditionals
- Context
  - That
  - Topic
  - set
- System calls
- Get & set variables

Statistical Retrieval-based Dialogue

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

Some Word-based Classification Approaches

- Latent Semantic Analysis (LSA)
- Support-Vector Machines (SVM)
- Relevance Model Retrieval
- Cross-language Relevance Model

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

Text as vectors

Tell me about yourself?

Term	tf
tell	1
me	1
about	1
…	…

- "Bag of words"

Tell	me	about	yourself

- Stopping
- Stemming
- N-grams

Tell me	Me about	About yourself	
Tell me about	Me about yourself		

Term Weights

Term	tf	df
tell	1	5
me	1	100
about	1	10
…	…	

$$w_{i,j} = \begin{cases} 1 & \text{word } i \text{ is present in string } j \\ 0 & \text{otherwise} \end{cases}$$

$$w_{i,j} = tf_{i,j}$$

$$w_{i,j} = tf_{i,j}/df_i$$

$$w_{i,j} = tf_{i,j}/\log df_i$$

$$w_{i,j} = \frac{tf_{i,j}}{tf_{i,j} + 0.5 + 1.5\frac{doclen}{avgdoclen}} \cdot \frac{\log(\frac{colsize+0.5}{docf_i})}{\log(colsize+1)}$$

SVM

- Text string tokenized
- unigram, bigrams, trigrams
- term vectors

$$w_{i,j} = \frac{tf_{i,j}}{tf_{i,j} + 0.5 + 1.5\frac{doclen}{avgdoclen}} \cdot \frac{\log(\frac{colsize+0.5}{docf_i})}{\log(colsize+1)}$$

Relevance Model

- Relevance Model: P(w|R) - prob that a random word from appropriate answer is w. P(w|R) ~apprx P(w|Q)
- Estimate P(w|Q) - prob observing word w in an answer given question
- Estimate P(w|A) - prob observing word w in an answer given answer
- Compare two probabilities

Compare P(w|Q) and P(w|A)

- Compute cross-entropy
  - Kullback-Leibler divergence
  - Minimize $D(p_q||p_a)$

$$D(p_q||p_a) = \sum_{w \in V} P(w|Q) \log \frac{P(w|Q)}{P(w|A)}$$

Estimate P(w|A)

- Estimate from known data

$$P(w|A) = \pi_A(w)$$

Maximum likelihood estimator (MLE)

$$\pi_s(w) = \lambda_\pi \cdot \frac{\#(w,s)}{|s|} + (1-\lambda_\pi) \cdot \frac{\sum_s \#(w,s)}{\sum_s |s|}$$

Estimate P(w|Q) v. 1

- Approach 1: consider answers as class labels (ignore content of answer)
- combine together all questions for an answer into a pseudo-answer
- Compare a test question to each pseudo-answer and select the best match

Estimate P(w|Q) v. 1

$$P(w|Q) = \frac{\sum_{s \in S} \pi_s(w) \prod_{i=1}^{m} \pi_s(q_i)}{\sum_s \prod_{i=1}^{m} \pi_s(q_i)}$$

$$P(w|A) = \pi_A(w)$$

- w is any word we ever see in questions
- q is a word in test question
- S is the set of all training questions
- A is all questions for an answer combined
- P(w|Q) is an average for a w over all s and q

Estimate P(w|Q) v. 2

- Approach 2:
  - Answer text matters!
  - Questions and answers are two different languages

Estimate P(w|Q) v. 2

$$P(w|Q) = \frac{\sum_s \alpha_{A_s}(w) \prod_{i=1}^{m} \pi_{Q_s}(q_i)}{\sum_s \prod_{i=1}^{m} \pi_{Q_s}(q_i)}$$

$$P(w|A) = \alpha_A(w)$$

- s iterates over all {Q,A} pairs of training data
- alpha is like pi, but on the answer domain

$$\alpha_x(w) = \lambda_\alpha \frac{\#(w,x)}{|x|} + (1-\lambda_\alpha) \frac{\sum_s \#(w,x)}{\sum_s |x|}$$

[Victor Lavrenko. 2004. *A Generative Theory of Relevance*.]

Sgt Blackwell Evaluation

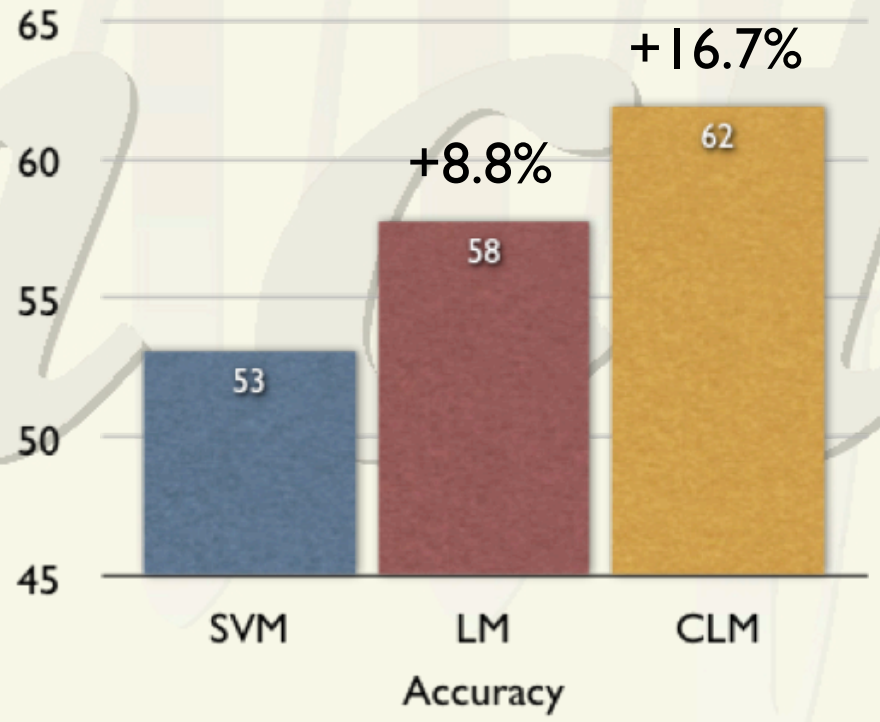
Questions:

1. What are the best classification techniques?
2. How much do speech recognition errors affect performance?

Question 1: Which classification methods are best?

- Method:
  - Use off-line training set & test set paradigm
    - Set of appropriate answers labelled for each question in corpus
    - 1261 questions, 60 answer-classes
    - 10-fold cross-validation
  - Consider top answer: is it appropriate?

Sgt Blackwell: Answer retrieval results



Sgt Blackwell Evaluation

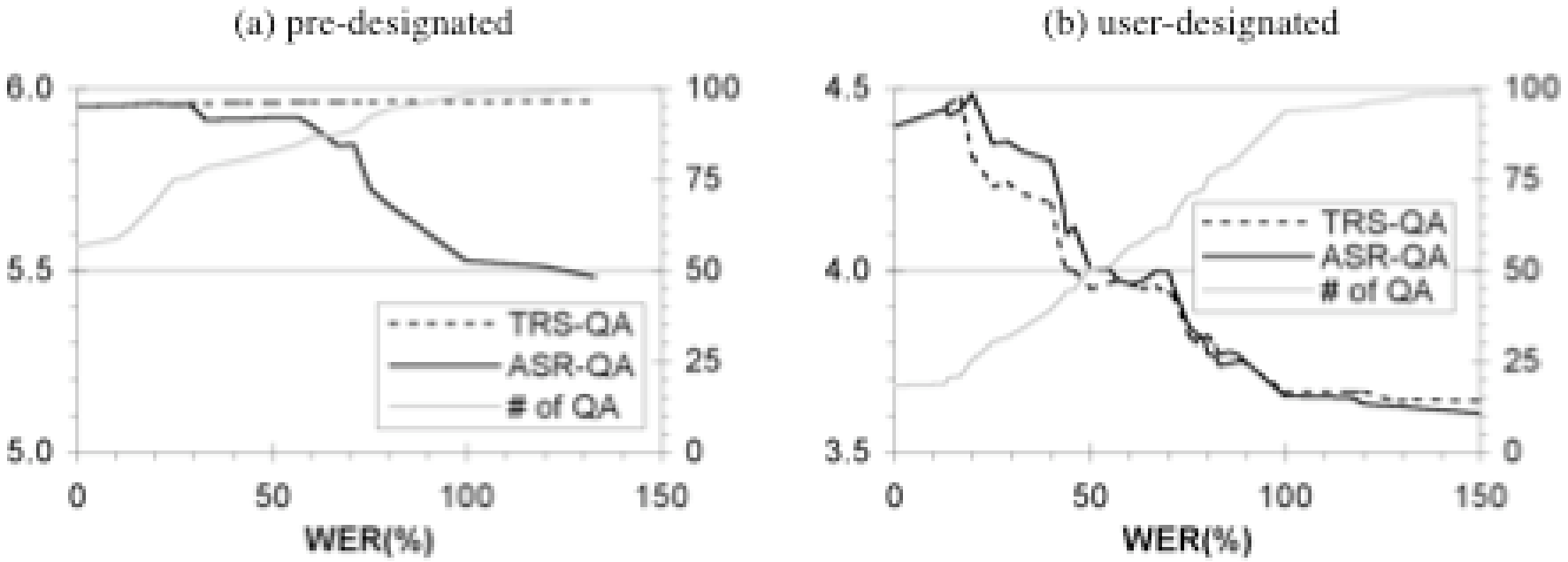
- Questions
  - What are the best classification techniques?
  - How much do speech recognition errors affect performance?
- Metrics
  - Accuracy of Speech recognizer & classifier
  - Appropriateness of replies (including to unseen and out of domain questions)
  - Answers rated for relevance (scale from 1-6)
- Experiment: 20 users, asking 20 questions: 10 given, 10 user-defined (Leuski et al IUI 2006, to appear)

Gandhe et al 2004 Response coherence coding

1	Response is not related in any way the question
2	Response contains some discussion of people or objects mentioned in the question, but does not really address the question itself
3	Response partially addresses the question, but little or no coherence between the question and response
4	Response does mostly address the question, but with major problems in the coherence between question and response; seems like the response is really addressing a different question than the one asked.
5	Response does address the question, but the transition is somewhat awkward
6	Response answers the question in a perfectly fluent manner

Table 1: Dialogue Quality Scale

Sgt Blackwell Evaluation Results



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?

Speech-act related Adjacency Pairs

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

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

Dialogue Acts in FSM

- Inform - convey information
- Question - set context for inform
- Answer - inform after question
- Confirm - show understanding
- Directive - ask for an action to be done
- Action - do an action

Example: ‘Bridge of Death’ Scene from Monty Python & the Holy Grail

Example: ‘Bridge of Death’ Scene from Monty Python & the Holy Grail

- BRIDGEKEEPER: Stop! Who would cross the Bridge of Death must answer me these questions three, ere the other side he see.
- ARTHUR: Ask me the questions, bridgekeeper. I am not afraid.
- BRIDGEKEEPER: What... is your name?
- ARTHUR: My name is Sir Arthur, King of the Britons.
- BRIDGEKEEPER: What... is your quest?
- ARTHUR: To seek the Holy Grail.
- BRIDGEKEEPER: What... is your favourite colour?
- ARTHUR: Blue. No, yellow!
- BRIDGEKEEPER: Hee hee heh heh heh heh!
- ARTHUR: Thank you. Thank you very much.

Example: ‘Bridge of Death’ Scene from Monty Python & the Holy Grail

BoD: Preamble

- **GALAHAD:** There it is!
- **ARTHUR:** The Bridge of Death!
- **ROBIN:** Oh, great.
- **ARTHUR:** Look! There's the old man from scene twenty-four!
- **BEDEVERE:** What is he doing here?
- **ARTHUR:** He is the keeper of the Bridge of Death. He asks each traveller five questions--
- **GALAHAD:** Three questions.
- **ARTHUR:** Three questions. He who answers the five questions-
- **GALAHAD:** Three questions.
- **ARTHUR:** Three questions may cross in safety.
- **ROBIN:** What if you get a question wrong?
- **ARTHUR:** Then you are cast into the Gorge of Eternal Peril.

BoD: Preamble (2)

- **ROBIN:** Oh, I won't go.
- **GALAHAD:** Who's going to answer the questions?
- **ARTHUR:** Sir Robin!
- **ROBIN:** Yes?
- **ARTHUR:** Brave Sir Robin, you go.
- **ROBIN:** Hey! I've got a great idea. Why doesn't Launcelot go?
- **LAUNCELOT:** Yes. Let me go, my liege. I will take him single-handed. I shall make a feint to the north-east that s--
- **ARTHUR:** No, no. No. Hang on! Hang on! Hang on! Just answer the five questions--
- **GALAHAD:** Three questions.
- **ARTHUR:** Three questions as best you can, and we shall watch... and pray.
- **LAUNCELOT:** I understand, my liege.
- **ARTHUR:** Good luck, brave Sir Launcelot. God be with you.

BoD: Dialogue 1: Lancelot

- **BRIDGEKEEPER:** Stop! Who would cross the Bridge of Death must answer me these questions three, ere the other side he see.
- **LAUNCELOT:** Ask me the questions, bridgekeeper. I am not afraid.
- **BRIDGEKEEPER:** What... is your name?
- **LAUNCELOT:** My name is 'Sir Launcelot of Camelot'.
- **BRIDGEKEEPER:** What... is your quest?
- **LAUNCELOT:** To seek the Holy Grail.
- **BRIDGEKEEPER:** What... is your favourite colour?
- **LAUNCELOT:** Blue.
- **BRIDGEKEEPER:** Right. Off you go.
- **LAUNCELOT:** Oh, thank you. Thank you very much.

BoD: Dialogue 2: Robin

- **BRIDGEKEEPER:** Stop! Who approacheth the Bridge of Death must answer me these questions three, ere the other side he see.
- **ROBIN:** Ask me the questions, bridgekeeper. I'm not afraid.
- **BRIDGEKEEPER:** What... is your name?
- **ROBIN:** 'Sir Robin of Camelot'.
- **BRIDGEKEEPER:** What... is your quest?
- **ROBIN:** To seek the Holy Grail.
- **BRIDGEKEEPER:** What... is the capital of Assyria? [pause]
- **ROBIN:** I don't know that! Auuuuuuuugh!

BoD: Dialogue 3: Galahad

- **BRIDGEKEEPER:** Stop! What... is your name?
- **GALAHAD:** 'Sir Galahad of Camelot'.
- **BRIDGEKEEPER:** What... is your quest?
- **GALAHAD:** I seek the Grail.
- **BRIDGEKEEPER:** What... is your favourite colour?
- **GALAHAD:** Blue. No, yel-- auuuuuuuugh!
  BRIDGEKEEPER: Hee hee heh.

BoD: Dialogue 4: Arthur

- **BRIDGEKEEPER:** Stop! What... is your name?
- **ARTHUR:** It is 'Arthur', King of the Britons.
- **BRIDGEKEEPER:** What... is your quest?
- **ARTHUR:** To seek the Holy Grail.
- **BRIDGEKEEPER:** What... is the air-speed velocity of an unladen swallow?
- **ARTHUR:** What do you mean? An African or European swallow?
- **BRIDGEKEEPER:** Huh? I-- I don't know that. Auuuuuuuugh!
- **BEDEVERE:** How do know so much about swallows?
- **ARTHUR:** Well, you have to know these things when you're a king, you know.

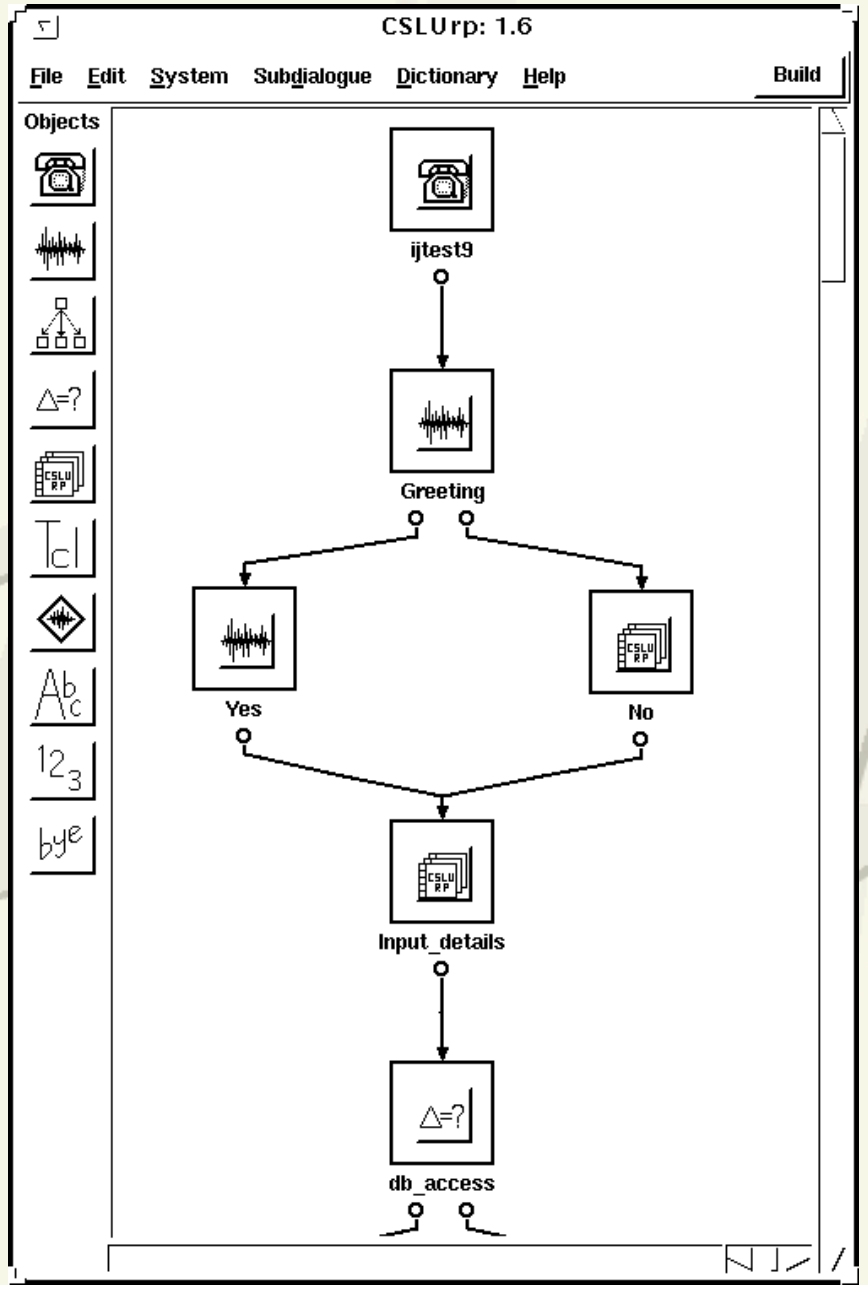
Bridgekeeper: Local structure

- (left as an exercise)

Bridgekeeper: FSM

- Draw on board

CSLUrp Interface



Voice XML

```
<?xml version="1.0"?>
<vxml version="2.0"><menu>
    <prompt>    Say one of:
    <enumerate/>  </prompt>
    <choice
    next="http://www.sports.exampl
    e/start.vxml">     Sports
    </choice>
<choice
    next="http://www.weather.exam
    ple/intro.vxml">     Weather
    </choice>
<choice
    next="http://www.news.example
    /news.vxml">     News
    </choice>
<noinput>Please say one of
    <enumerate/></noinput>
</menu></vxml>
```

Sample Dialogue:

- Computer: Say one of: Sports; Weather; News.
- Human: Astrology
- Computer: Please say one of: Sports; Weather; News.
- Human: Sports
- Computer: (proceeds to http://www.sports.example/start.vxml)

Outline for Course

- Yesterday: Introduction, Architecture of Dialogue Systems, Example Systems
- Today: Simple structures: S-R, IR, finite State
- Tomorrow: Frame-based and Information State
- Thursday: Plan-based and Logic Based
- Friday: Advanced Topics