Dialogue Modeling for Spoken Language Systems

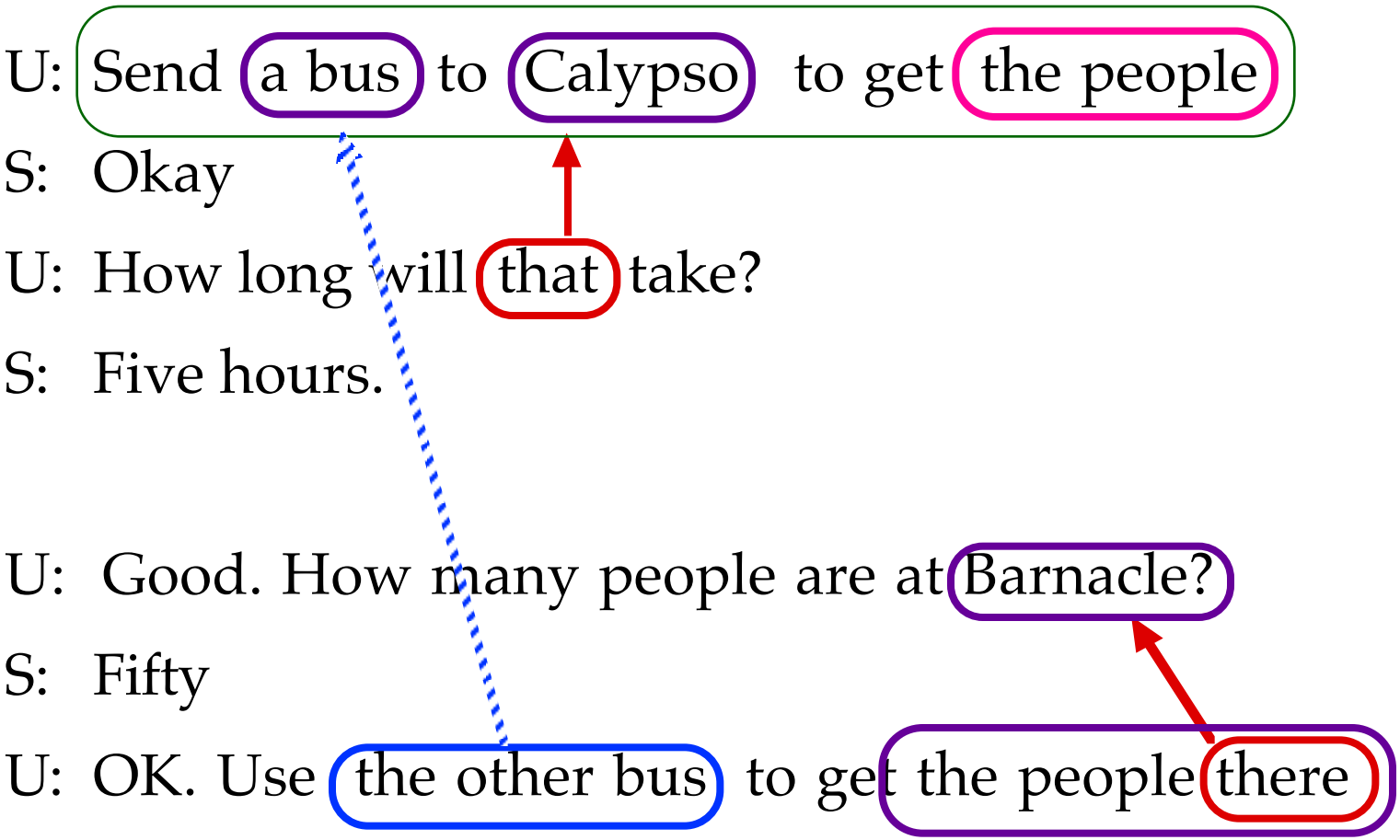
James F. Allen
University of Rochester

ACL Workshop * July 1997

What are Dialogue Models used for?

- Providing context for understanding
  - reference resolution
  - disambiguation/intention recognition
- Providing context for determining appropriate responses
  - determining what content needs to be realized
  - generating effective utterances
- Managing initiative and system activity
  - determining how the system responds to different situations
  - managing subdialogues and topic flow

Reference



Reference requiring knowledge of situation,
Reference requiring problem solving knowledge ,
Anaphoric reference from discourse context,
Complex reference using discourse context

Disambiguation/Intention Recognition

U: Is there a shorter route?

"Lexical Ambiguity" -- less time or less distance?

U: Let's plan to evacuate Abyss as quickly as possible

"Structural Ambiguity" -- plan quickly or evacuate quickly?

U: Where are the people?

"Temporal Context Ambiguity" -- the current actual physical locations, the planned locations, ...

"Intended Reference Ambiguity" -- all the people or the people just discussed?

What Content Needs to be Realized?

U: Where are the people

No prior context, the goal is to evacuate island --->

Show Map, Highlight Locations, Identify how many people at each location

Map on display, goal to identify all cities with people --->

Highlight Locations, list locations

Generating Effective Utterances

Realize (AT-LOC TRUCK-1 SEGMENT-445)

in response to "Where is the truck?"

1. "Between Abyss and Barnacle"
2. "Here" + highlight truck on the map
3. "It is on the road from Abyss to Barnacle" + highlight on map

in response to "Are there any trucks?"

4. "There is one on the road between Abyss and Barnacle"
5. "A truck carrying the people from Calypso is between Abyss and Barnacle"

in response to "What is that on the road from Abyss to Barnacle"

6. "The truck carrying the people from Calypso"
7. "Truck one"

Managing the System Responses

U: Use truck one to get the people from Calypso to Delta

case 1: where plan can be constructed

system passive

1. "Which route would you like to take?"

system fills in details

2. "OK" + system picks a route

system helps with the planning process

3. "It would be faster to use a helicopter"

4. "OK. What should we do when we get them to Delta?"

case 2: where there is no "truck one"

5. "Which truck do you mean?"

6. "Do you mean the truck at Delta?

Managing Subdialogues

U: Use a truck to get the people from Calypso to Delta

S: OK. Which one do you want to use?

U: Which is the fastest truck?

S: The truck at Barnacle.

U: Can that truck get to Calypso faster than the others?

S: No.

U: OK. let's use the truck at Delta then.

S: OK. Which route do you want to use

U: Go along the coast

S: Fine.<<Builds new plan and displays it>>

Topic Flow

U: Use a truck to get the people from Calypso to Delta

S: OK <<Displays route>>

U: Would it be faster to go along the coast?

S: That option would take X hours <<Show alternate plan>>

U: Why so long?

S: The coastal road is likely to be in bad condition.

U: OK. Let's forget the coastal route.

U: So when will the truck arrive?

Topic/Option 1: Travel inland

Topic/Option 2: Travel along the coast

Classifying Dialogue Systems

Reference Capabilities

How can the context be used to allow efficient referring expressions?

Task Complexity

Can task change as dialogue develops, how compelx can a task be?

Dialogue Management Capabilities

Can the dialogue itself be used to manage the dialogue?

Initiative

Who controls the dialogue?

Speech Acts Supported

What different things can a person do in a dialogue?

Modality

What modes of communication are available?

Incrementality

Can topics and utterances be developed incrementally?

Classifying Dialogue Systems II

Reference	Static - proper names, unique descriptions	Descriptions resolved by status of task	Personal Pronouns	Arbitrary Anaphoric Expressions
Task Complexity	Static	Adaptive Single	Sequential	Hierarchical / Complex
Explicit Dialogue Management	None	minimal pair Clarification / Correction	Clarification / Correction Subdialogues	"Meta" conversation
Initiative	Fixed Single	Fixed Mixed	Dynamic Mixed Initiative	

Speech Acts Supported

Questions only, requests, assertions, promises, ...

Modality

keyboard NL, speech, graphics, mouse, pen, handwriting, ...

Incrementality

“installment” queries, back channel responses, ...

The Need for Focus-based Strategies

Referring Expressions

*Send **it** on to Abyss* ----> Most relevant movable object

*How long will **it** take?* ----> Most relevant action/plan

*Pick up the people **there**?* ----> Most relevant location

*And take **them** to Barnacle?* ----> Most relevant set of movable objects

Implicit Arguments

*Send the truck to Abyss and pick up **the people***

----> The people at Abyss, not last people mentioned

----> Pick up the people with the truck

Implementing Reference

Fixed Strategy

Hard-wired interpretation strategy

Works only for names, and descriptions in limited contexts

Topic/Task Based

Fixed interpretation strategy parametrized by Task Context

e.g., in ATIS, *there* -> destination, *it* --> the plane or trip

Recency over last N sentences (history list)

Maintain list of all objects mentioned over last N sentences

Pick most recent that matches type restrictions

Hierarchical

Relevant focus objects determined by recency strategy organized by topic structure

Task Complexity

Static

System performs a single task not affected by prior utterances

Every utterance is interpreted using identical strategy

Possible applications

answer single queries on a static database with no follow-up

Adaptive Single Task

System performs single task, affected by prior utterances

Interpretation strategy parametrized by task requirements **at current time in the dialogue**

Example application

multiple interaction information retrieval, e.g., ATIS, where missing parts of current query "filled in" from prior queries

Static & Adpative Single Task Models

Advantages/Disadvantages

+ Simplicity

  - Interpretation can be built into parser
  - Simple robustness techniques can be effective
  - Adaptive models allow for incremental specification of task

– Generality

  - limited follow-up and elaboration
  - no clarification
  - can't support complex tasks

Topic Flow Models

Sequence of Adaptive Topics

Utterances are interpreted with respect to one of a fixed set of tasks (e.g., find directions, look up weather, etc)

Topic change signaled lexically by operation requested

Focus strategy allows "transfer" of focused object between tasks

Pluses/Minuses

+ Allows multiple tasks to be addressed in a single dialogue

– Generality
  - limited interaction between tasks

General Topic Flow Models

Topics can be related hierarchically as well as sequentially

Requires a strategy for managing **topic flow**, i.e., detecting

- new topics
- refinements of topics
- resumption of old topics

Complexity varies with complexity of back end reasoning required

Advantages/Disadvantages

+ Generality
  - a wide range of applications can be supported, including "mixed initiative" interactions

– Complexity
  - tracking topic flow can be complex

Single Initiative Models

User Initiative

System responds to user requests

Example

Question answering systems

Pluses/Minuses

+ User can direct the dialogue as desired

– System cannot guide the dialogue

System Initiative

System directs user by specifying the information it needs

Example

Form filling based applications

Pluses/Minuses

+ Better reliability since system has strong expectations to use in interpreting user 's utterances

– Often cumbersome and annoying for user, viz current touch-tone "dialogues" for handling inquiries

Mixed Initiative

The Idea

System and user may both take initiative at different times

Dialogue strategy varies throughout interaction

Varying Degrees of Mixed Initiative

1. Pro-active system that allows user clarification

2. Passive system except for clarification subdialogues and "alerts"

3. Predetermined initiative strategies

    e.g., user specifies the next task, system directs the performance of the task until completed

4. Initiative is "negotiated" as the dialogue progresses based on the needs of both parties.

Pluses/Minuses

+ user friendly, mirrors natural dialogue, effective interaction

– complicated control structure requires detailed model of task and problem solving methods

Dialogue Management

Vary by the degree to which the dialogue itself can be discussed, including topics such as

what was meant - e,g., clarification, correction, confirmation

what to discuss - e.g., changing the topic

how to discuss it - e.g., negotiating responsibilities & initiative

Typical Levels

None - utterances can only perform the task

Minimal Diversions- e.g., ability to handle clarification Q/A

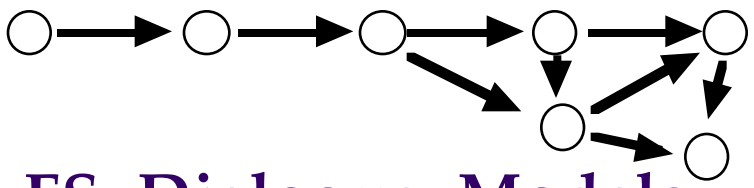
Limited Subdialogues - e.g., special purpose handling of clarification subdialogues

"Meta" Discussion - e.g., explicitly negotiating what to talk about first

Computational Models for Enabling Dialogue Systems

- Static
- Finite State
- Task-based control
- Agent-based control

Utterance-based FSM



Utterance-based FS Dialogue Models

State corresponds to a "minimal pair"

Examples

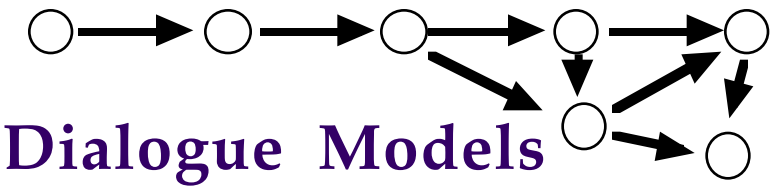
- System asks for social security number, User provides it
- System provides list of options, User selects one.

Pluses/Minuses

+ Practicality
  - high degree of system control, can prompt for "easy" utterances
  - can tailor language model to each state

– Limited Application, User Satisfaction
  - how to avoid the plague of "touch-tone" telephone interfaces

Topic-based FSM



Topic-based FS Dialogue Models

State corresponds to a topic /task that defines an interpretation context

Interaction within a topic might be mixed-initiative

Examples

- Book an airline ticket, book a hotel, rent a car, ...
- Introduce a goal, clarify it, develop a plan, evaluate plan, execute plan, ...

Pluses/Minuses

+ Practicality
  - high level system control, can tailor language model and interpretation strategies to each topic/state

– Complexity of determining topic shifts

Plan/Agent-based Dialogue Models

Dialogue results from the interaction of two or more agents

Models of agent's beliefs, including beliefs about other agent

Agents choose their actions based on their intentions and beliefs, according to "rules" of behavior

- cooperation
- helpfulness
- sincerity

Communication modeled in a Theory of Action

Speech acts change the mental states of the participants

Parsing maps language -> speech acts

Understanding --> speech act recognition (plan recognition)

Generation --> speech act planning

Evaluating Plan-based Models

+ Generality - the richest model of dialogue we have
  - unification of language and reasoning processes, for both interpretation and generation
  - unification of speech and GUI interfaces into a uniform framework
  - ability to support extended interactions, negotiation, clarification, ...

– Complexity
  - requires complex knowledge representations
  - computational complexities of planning and plan recognition