CS 599: Computational Models of Dialogue Modelling Fall 2005 Lecture 1: Overview of Course

David Traum

traum@ict.usc.edu

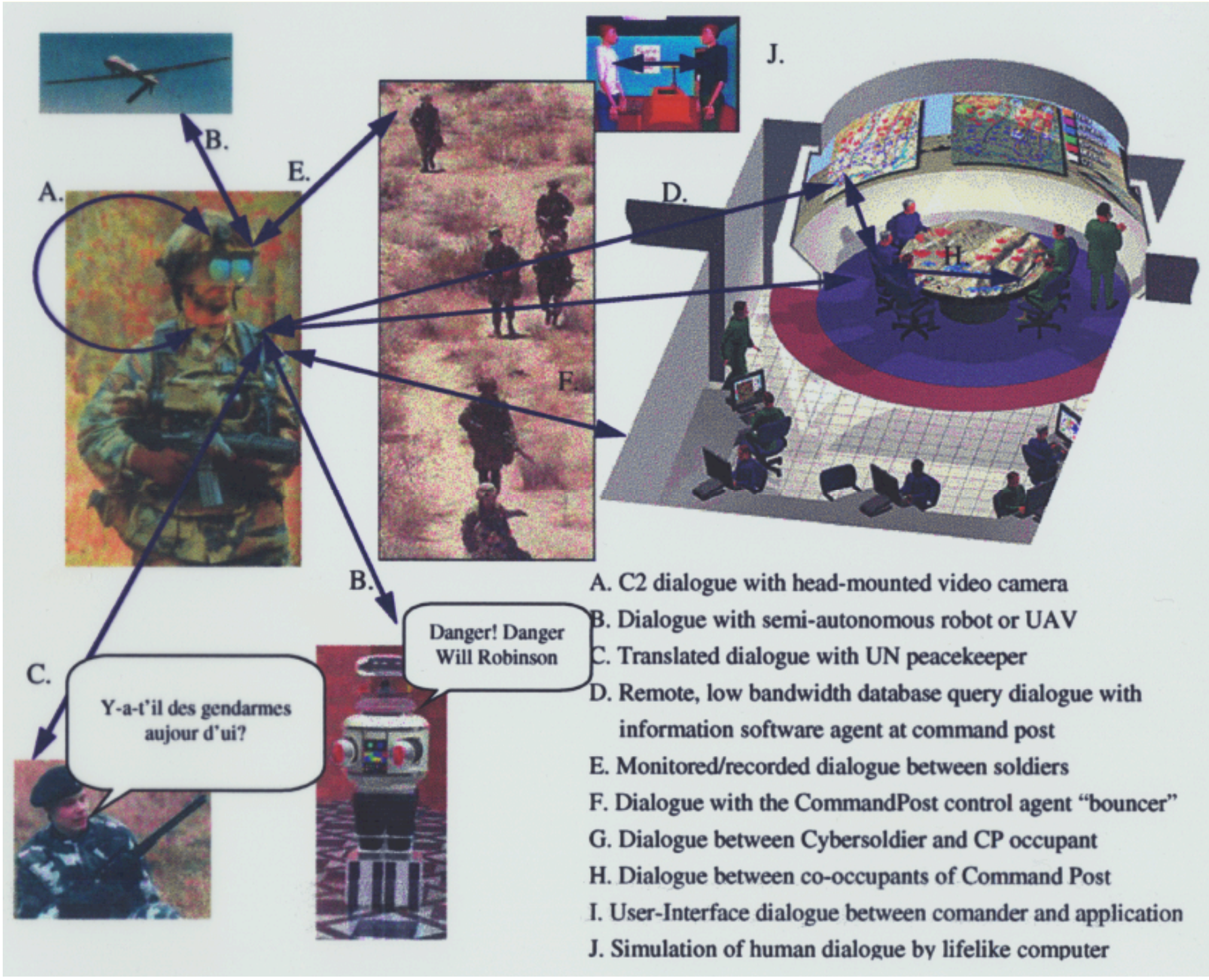
http://www.ict.usc.edu/~traum

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
- Types of Dialogue Agents
  - Information provider
  - Service provider
  - Instruction-giver
  - Advisor/Critic
  - Tutor
  - Collaborative partner
  - Conversational partner

Types of Dialogue Agents

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


A.
B.
C.
D.
E.
F.
G.
H.
J.
Danger! Danger
Will Robinson
Y-a-t'il des gendarmes
aujour d'ui?
A. C2 dialogue with head-mounted video camera
B. Dialogue with semi-autonomous robot or UAV
C. Translated dialogue with UN peacekeeper
D. Remote, low bandwidth database query dialogue with
information software agent at command post
E. Monitored/recorded dialogue between soldiers
F. Dialogue with the CommandPost control agent "bouncer"
G. Dialogue between Cybersoldier and CP occupant
H. Dialogue between co-occupants of Command Post
I. User-Interface dialogue between comander and application
J. Simulation of human dialogue by lifelike computer

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,…

Dialogue Management Tasks

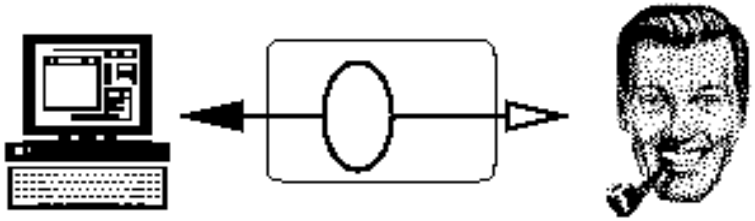
- Maintaining & Updating Context
- Deciding what to do next
- Interface with back-end/task model
- Provide expectations for interpretation

Dialogue Systems: State of the Art

- Deployed Commercial Systems
  - Call routing/call center first contact
  - Simple information tasks
  - Voice menus
- Useful systems
  - Medium-sized tasks (communicator, in car navigation)
  - Command & control
  - Language tutoring
- Advanced Research Prototypes
  - Collaborative systems
  - Adaptive systems
  - Multi-modal systems
  - Immersive Training

Two Approaches to Dialogue Systems:

front-end



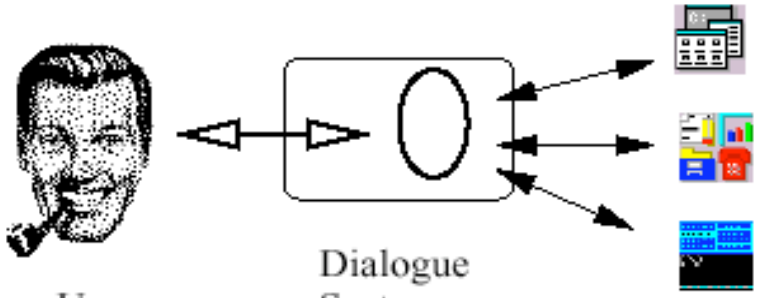
System as "translator" between user and backend system

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?

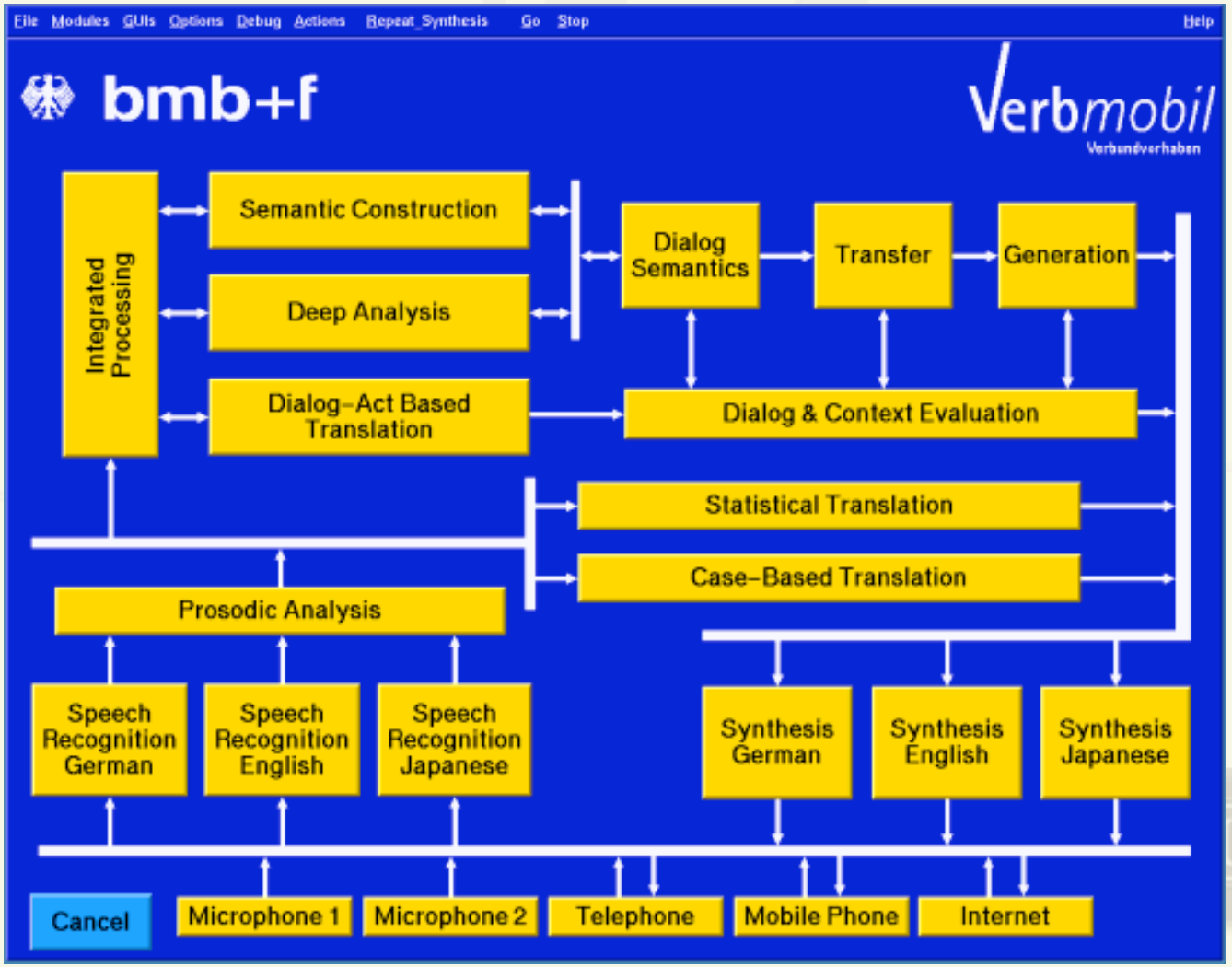
Key run-time question: what should be done now (given context, inputs, goals)?

Example Systems

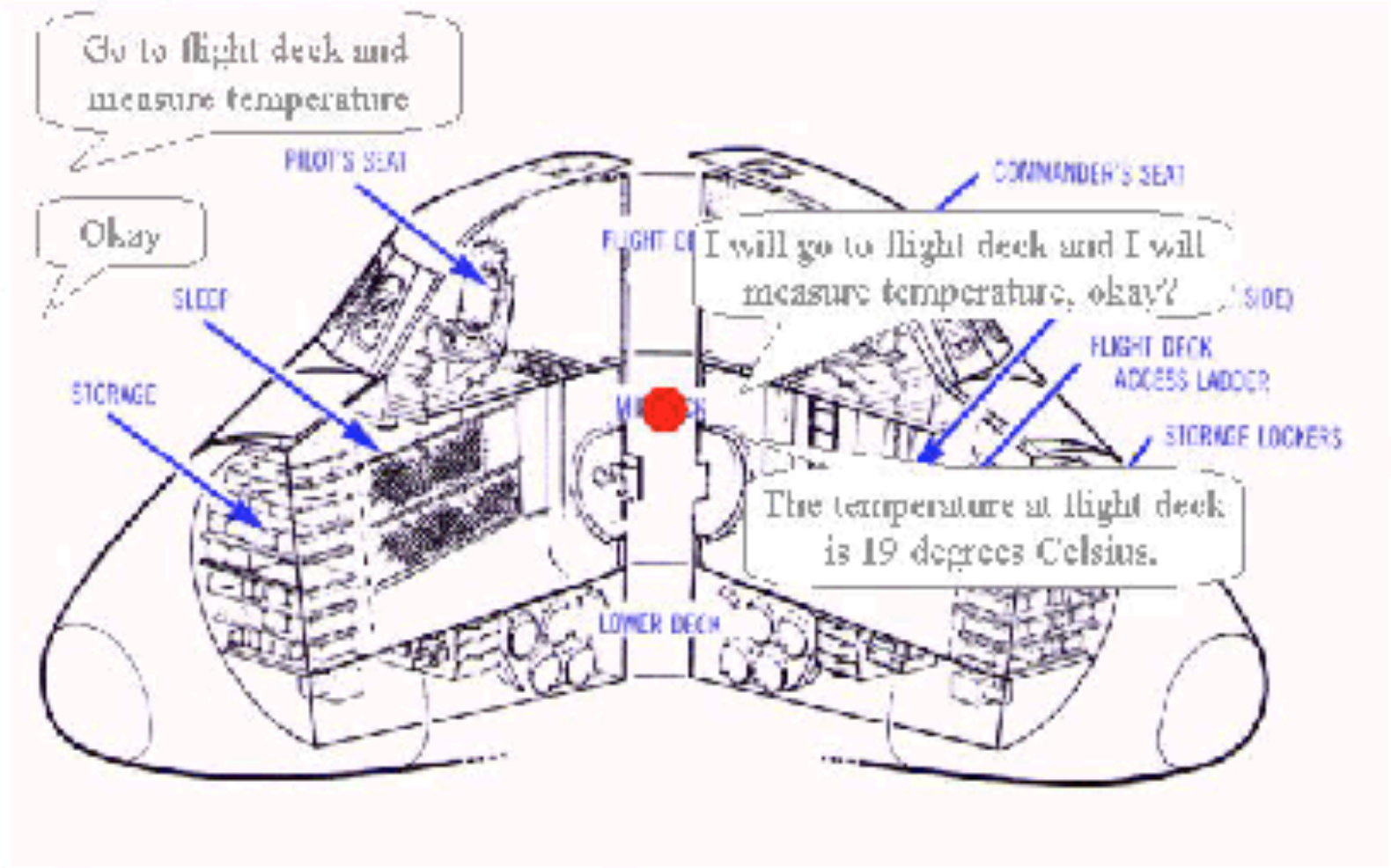
- United Airlines
- RAD
- Trains/TRIPS
- MRE & SASO

Verbmobil: Spoken Translation

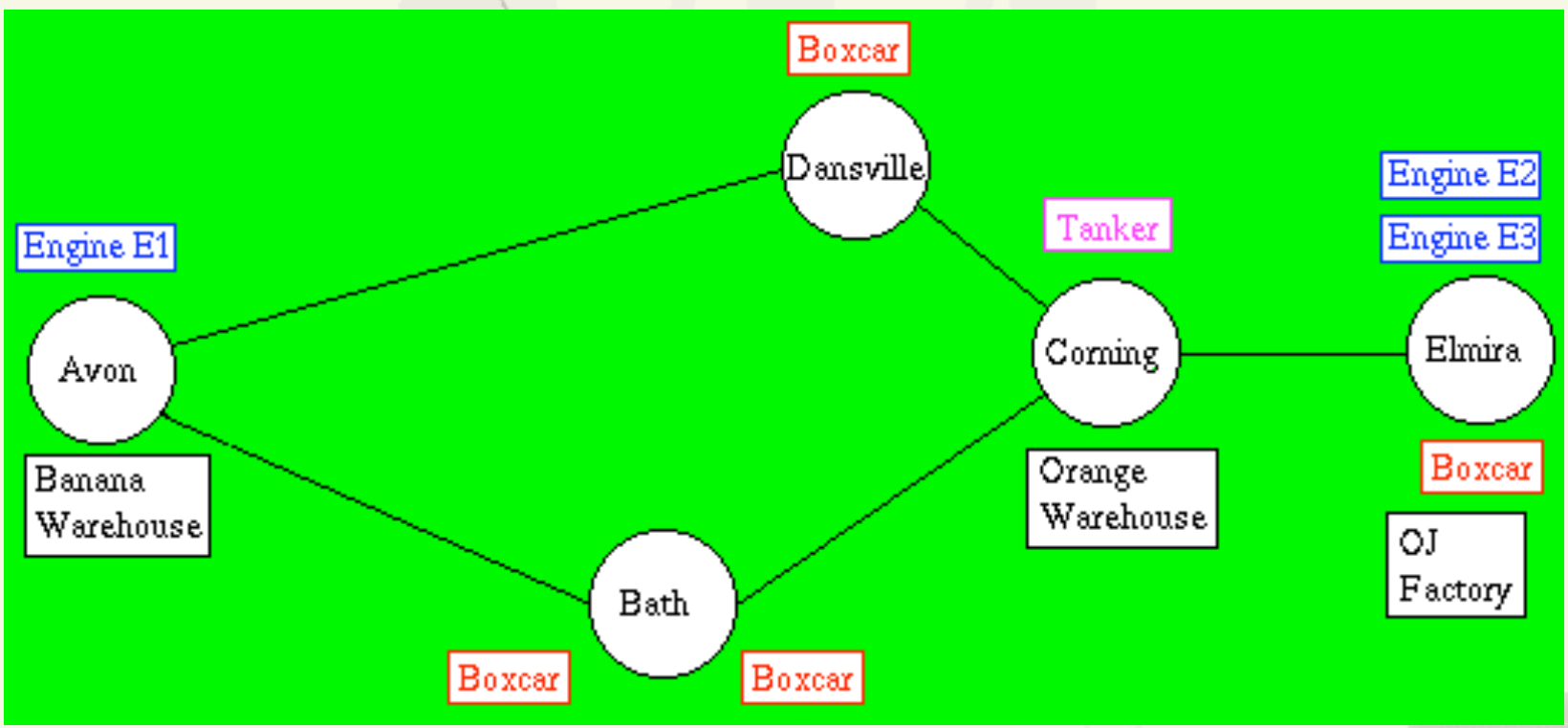
Verbmobil Architecture



NASA Rialist System



TRAINS-93 Domain



Trains-93 Dialogue

```
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)
```

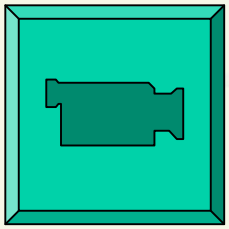
TRAINS-96 Domain

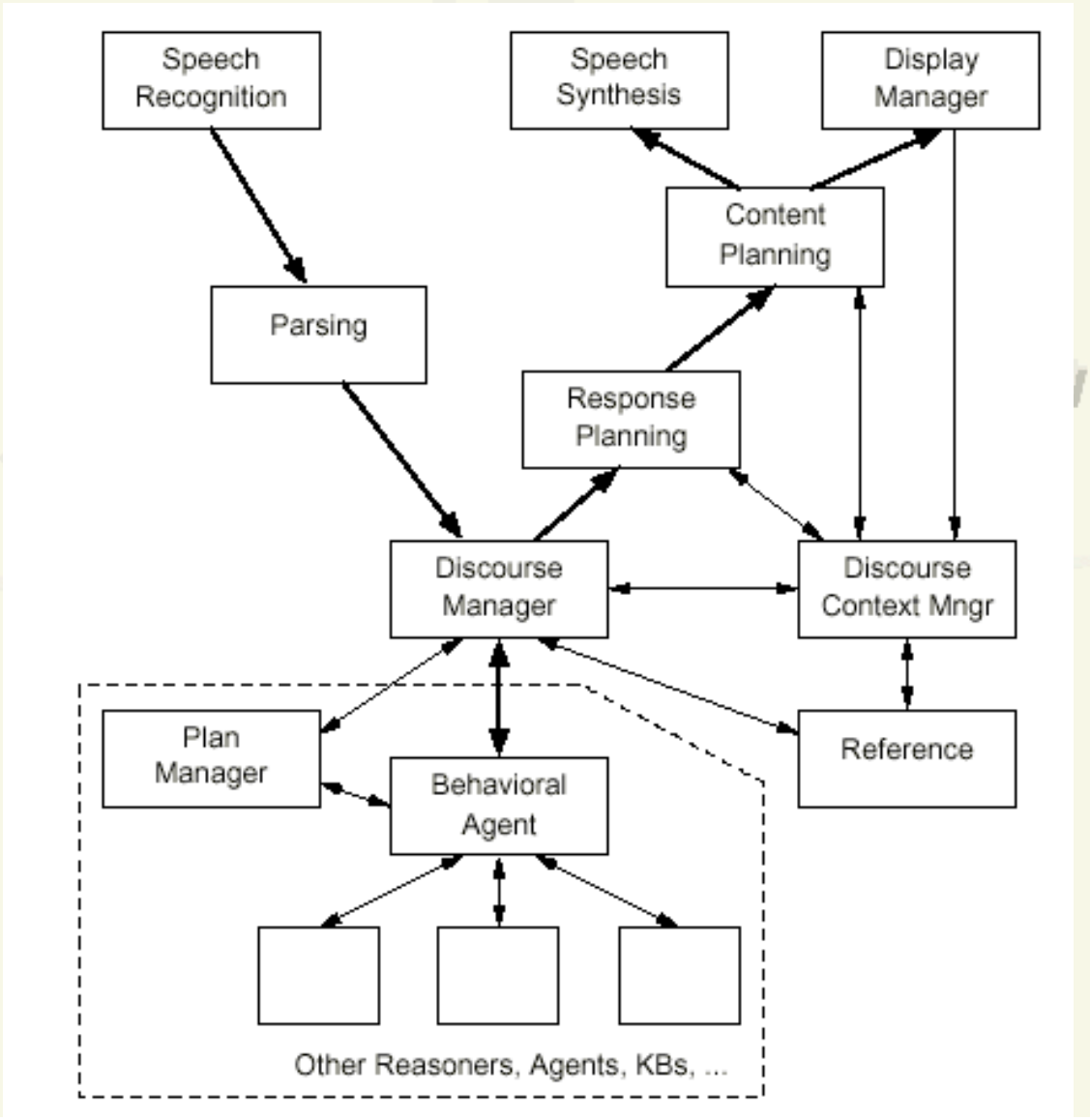
The TRAINS Project

An Interactive Natural Language-Based Planning Assistant

Department of Computer Science
University of Rochester

TRIPS Architecture





Trips Module Description

Module	Function
Speech Recognition (SR)	Transforming speech input into a word stream or word lattice
Parser	Transforming the SR output into interpretations, each a set of conventional speech acts, using full and robust parsing techniques
Reference Manager (REF)	Identifying the most salient referents for referring expressions such as noun phrases
Discourse Context Manager	Maintaining the global (topic flow) and local (salience with a topic) discourse context
Discourse Manager (DM)	Identifying the intended speech act, current task, current step in the current task, and system obligations arising from the dialogue
Behavioral Agent (BA)	Determines system actions (*e.g.*, answer a question, notify of a problem, request clarification); Manages the interface to the back-end systems.
Plan Manager	Constructing, modifying, evaluating, and executing plans (whether they are the subject of the conversation or the task being executed)
World KB	Maintains a description of the current state of the world under differing assumptions (*e.g.*, based on different plans or hypotheses)
Response Planner	Determining the best communicative act(s) (and their content) to accomplish the system's current goals and discourse obligations

SDS Components

- Architecture
- Input Interface (Audio, Keyboard,etc)
- Interpretation (internal representation)
- Dialogue Management
- Generation
- Output Interface

Dialogue Manager Architectures

- Integrated (tree-based)
- Finite-state
- Frame-based
- Plan-based
- Agent-based (BDI)

Interpretation: Speech Recognition

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

Interpretation: Parsing

- Styles
  - Key-word
  - Grammar-based
  - Concept-based (semantic parser)
  - Expectation-driven
- Spoken Dialogue vs. Written text
  - Utterance length, grammaticality, interactivity, repairability, transience,

Dialogue Management Tasks

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

Generation & Synthesis

- Template-based or Fixed
- Prosodic cues, multimodal generation
- Voice Clip, or TTS
- TTS or Concept to Speech

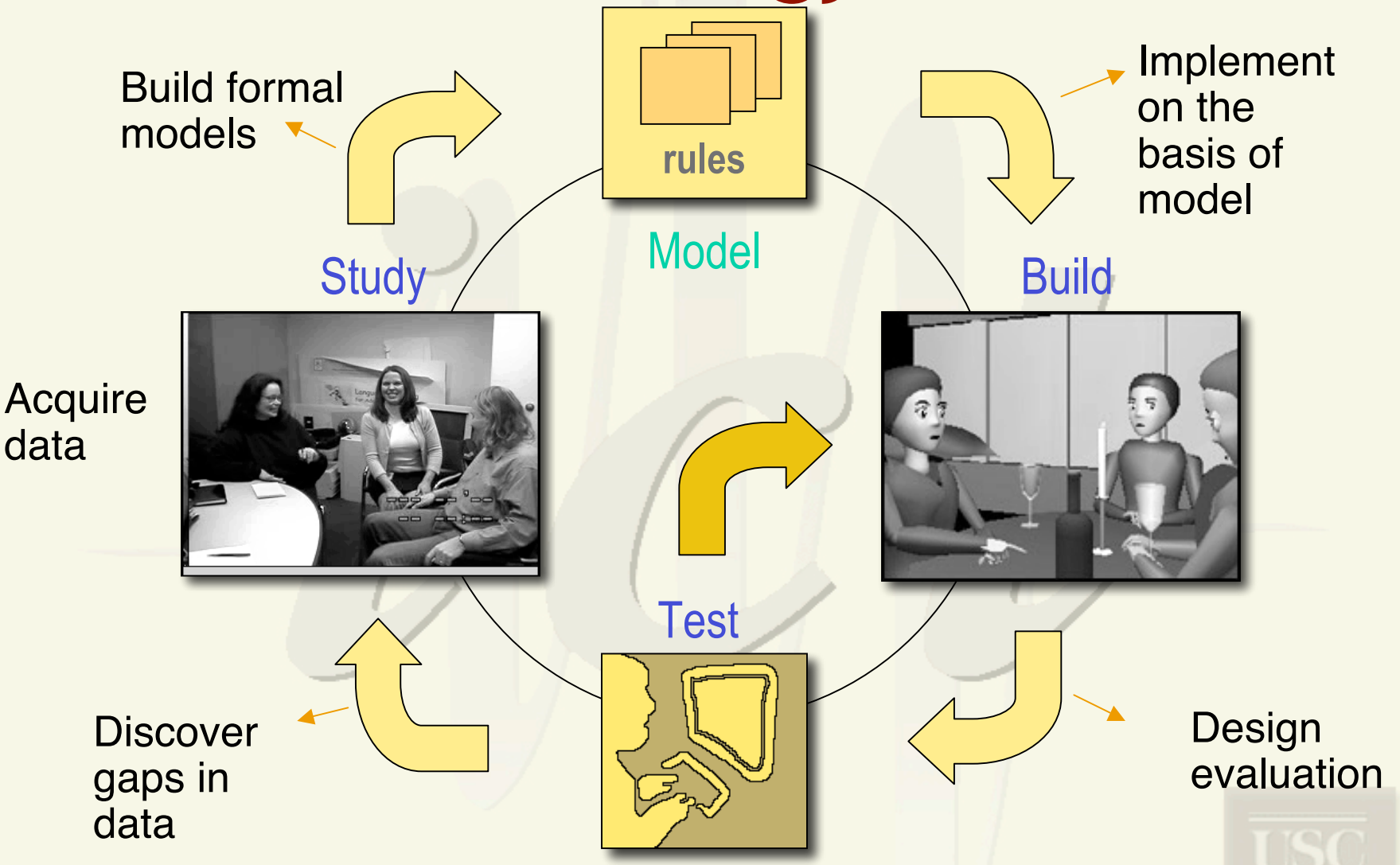
Using Data

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

Evaluation

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

Methodology



Each Step is subject to Evaluation

- Was the data appropriate?
- Is the model of human behavior correct?
- Does the system implement the model correctly?
- → Evaluated by micro-analysis.

Each Step is subject to Evaluation

- Is the system well implemented?
- Does the interface succeed better than X:
  - Do (which) people prefer/trust/enjoy the interface?
  - Does it make work easier/more efficient/better?
  - What uses of embodiment are most powerful
- → Evaluated by macro-analysis

Hot Topics In Dialogue Research

- Mixed Initiative
- Grounding
- Discourse Structure
- User/Agent Modeling
  - Affective dialogue
- Adaptive dialogue management
- Social context
  - Social roles
  - Obligations & commitments
  - politeness
- Multi-party (more than two) dialogue
  - Turn-Taking
  - Speaker and addressee id
  - Multiple conversations

Views on initiative (control)

- Any Contribution
  - MI Planning
  - Turn (Donaldson, Hagen)
- Type of Dialogue move
  - Initiative/Response (Dahlback et al, Carletta et al, Ishizaki)
  - Patterns: command, question, assertion, prompt
    - (Whittaker, Stenton &Walker, Smith and Hipp)
  - Amount/type of information
- Goal Interactions
  - Whose goals are being addressed
  - Game Playing: Sente or Tempo - forcing moves of other
  - Obligations vs. Goal (Traum & Allen)
- Multi-level concepts:
  - Choice of speaker, task, outcome (Novick & Sutton)
  - Discourse vs Task (Chu-Carroll & Brown), Local vs. Global (Rich and Sidner)
  - Hierarchical (Whittaker&Walker)

Example: Chu-Carroll & Brown

1. Customer:
   - I need some money. How Much do I have in my 6-month CD?
2. T alternatives:
   A. T: no initiative
      - You Have $5000 in that CD.
   B. T: Dialogue initiative
      - You Have $5000 in that CD, but that CD will not mature for another 3 months.
   C. T: both dialogue and task initiative
      - You Have $5000 in that CD, but that CD will not mature for another 3 months. However you have $3000 in another CD that will mature next week.

Views on Mixed-initiative

- Contributions by multiple parties
- Changing initiative-holder mid-interaction
  - Fixed phases, or variable shift
- User providing more input than asked for
  - Middle level between system and user
- Ability to handle set of complex behaviors
  - Answer, ignore, over-answer, barge-in (Hagen)

Example: Narayanan et al

- System Initiative (SI)
  - System: “VPQ. Please say the name of the person.”
  - Acceptable Response from User: “Larry Rabiner.”
- Mixed Initiative (MI)
  - System: “VPQ. Please say the name of the person.”
  - Acceptable Response from User: “Larry Rabiner’s fax number, please.”
- User Initiative (UI)
  - System: “VPQ. What can I do for you?”
  - Acceptable Response from User: “I’d like the fax number for Larry Rabiner.”

Styles of Response

1	**Sys**:	Where do you want to go?
2	**User**:	Boston.
3a	**Sys**:	When would you like to go?
3b		Tell me more about your travel plans.
3c		When would you like to go to Boston?
3d		Do you want to go to Boston?
3e		Did you say Boston?
3f		Boston?
3g		Boston or Austin?
3h		Where?
3i		Please Repeat.

Grounding Acts

Label	Description
initiate	Begin new DU, content separate from previous uncompleted DUs
continue	same agent adds related content to open DU
acknowledge	Demonstrate or claim understanding of previous material by other agent
repair	Correct (potential) misunderstanding of DU content
Request Repair	Signal lack of understanding
Request Ack	Signal for other to acknowledge
cancel	Stop work on DU, leaving it ungrounded and ungroundable

Grounding Automaton

Next Act	In State						
	S	1	2	3	4	F	D
initiate^I	1						
continue^I		1			4		
continue^R			2	3			
repair^I		1	1	1	4	1	
repair^R		3	2	3	3	3	
ReqRepair^I			4	4	4	4	
ReqRepair^R		2	2	2	2	2	
ack^I				F	1	F	
ack^R		F	F			F	
ReqAck^I		1				1	
ReqAck^R				3		3	
cancel^I		D	D	D	D	D	
cancel^R			1	1		D	

Grounding Example

(1)

1	I: Move the boxcar to Corning
2	I: and load it with oranges
3	R: ok

(2)

1	I: Move the boxcar to Corning
2	R: ok
3	I: and load it with oranges
4	R: ok

(3)

utt:	Grounding Act	DU1
1:	$\text{init}^{\text{I}}(1)$	1
2:	$\text{cont}^{\text{I}}(1)$	1
3:	$\text{ack}^{\text{R}}(1)$	F

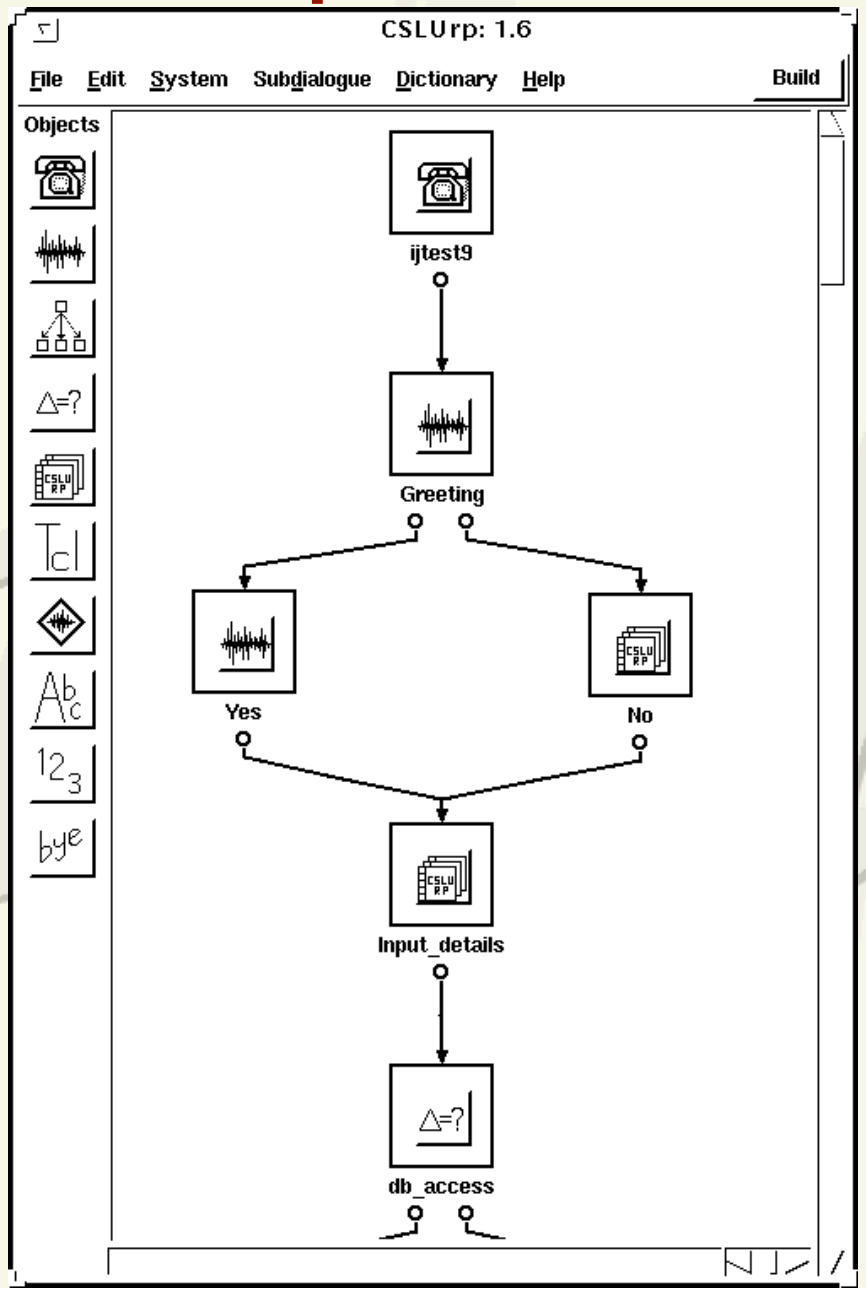
(4)

utt:	Grounding Act	DU1	DU2
1:	$\text{init}^{\text{I}}(1)$	1	
2:	$\text{ack}^{\text{R}}(1)$	F	
3:	$\text{init}^{\text{I}}(2)$	F	1
4:	$\text{ack}^{\text{R}}(2)$	F	F

Dialogue Toolkits

- Software Integration (OAA,Trains/Trips,Verbmobil)
- FSM Dialogue Kits (Nuance, OGI, …)
- Slot-Filling (Phillips)
- Current Development Kits:
  - Utterance-based (DARPA Communicator)
  - Information-based (TrindiKit)

CSLUrp Interface



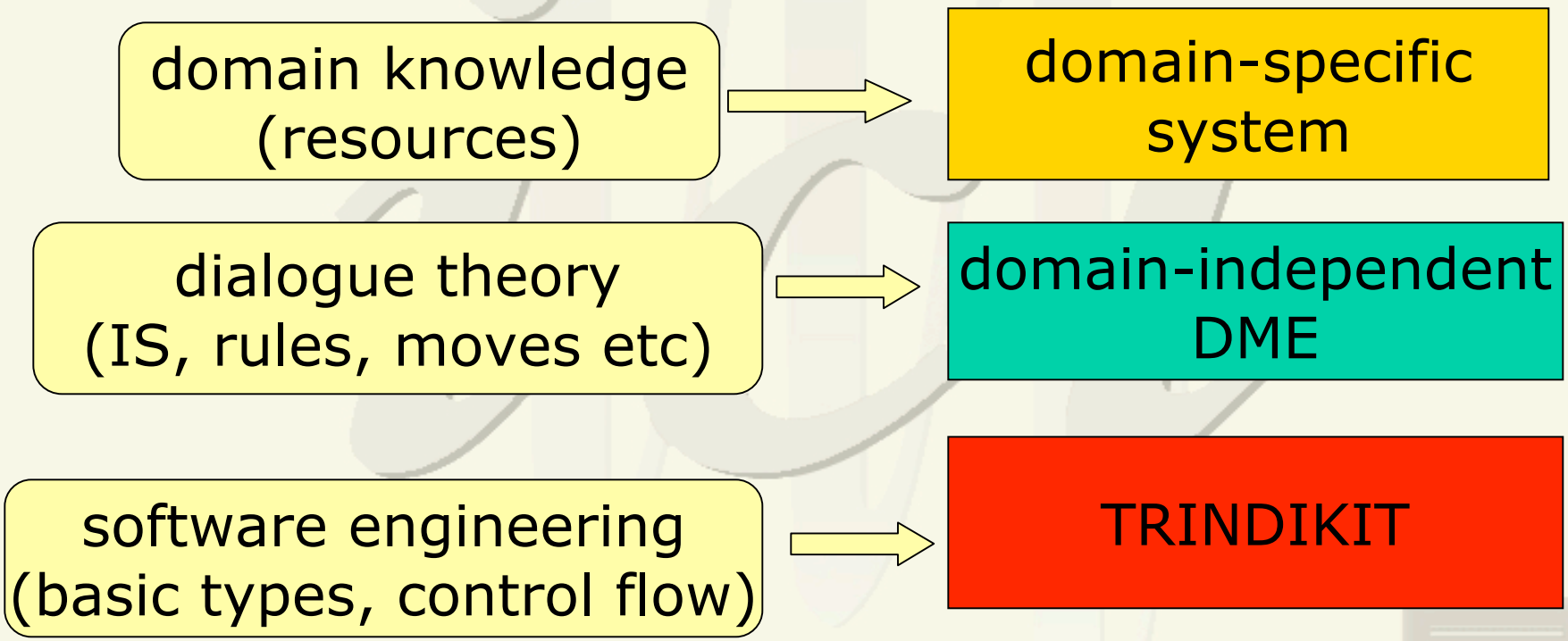
Trindi: Information State Theories of Dialogue

- 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

Sample Autoroute Dialogue

	W **WIZARD**		C **CALLER**
W	[1]: How can I help you?	C	[2]: A route please
W	[3]: Where would you like to start?	C	[4]: Malvern
W	[5]: Great Malvern?	C	[6]: Yes
W	[7]: Where do you want to go?	C	[8]: Edwinstowe
W	[9]: Edwinstowe in Nottingham?	C	[10]: Yes
W	[11]: When do you want to leave?	C	[12]: Six pm
W	[13]: Leaving at 6 p.m.?	C	[14]: Yes
W	[15]: Do you want the quickest or the shortest route?	C	[16]: Quickest
W	[17]: Please wait while your route is calculated.	C	…

Building a system



TrindiKit Systems

- GoDiS (Larsson et al) – information state: Questions Under Discussion
- MIDAS – DRS information state, first-order reasoning (Bos &Gabsdil, 2000)
- EDIS – PTT Information State, (Matheson et al 2000)
- SRI Autoroute – information state based on Conversational Game Theory (Lewin 2000)

  Robust Interpretation (Milward 1999)

Case Studies: Virtual Human Dialogue @ ICT

- MRE

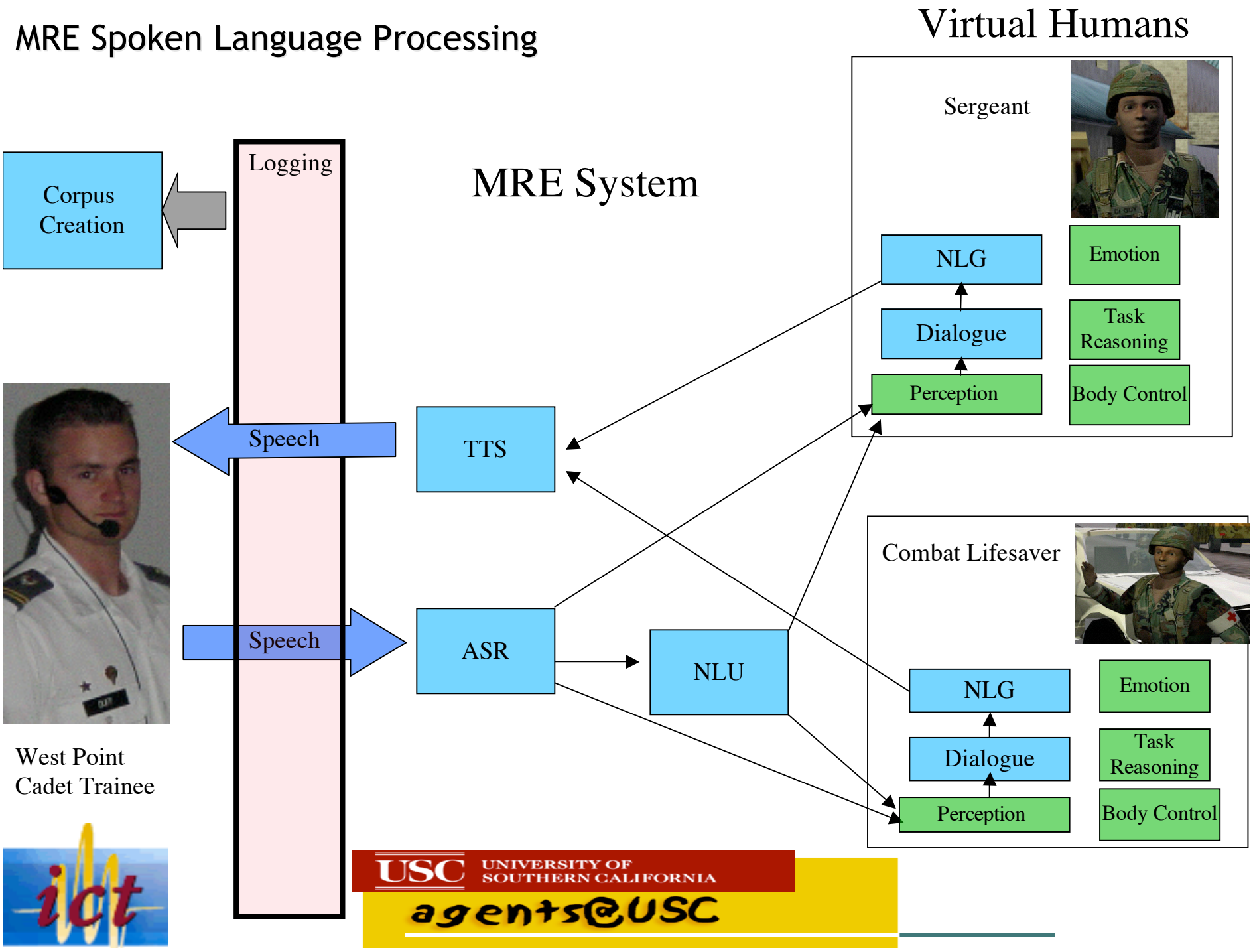
- SASO-ST

Immersive Training Environment

- **VR Theatre**
  - 8’ 150˚ Curved Screen, Multiple Projectors
  - 10-2 3-d spatialized sound

- **Mission Rehearsal Exercise** (Swartout et al ‘01)
  - **Human lieutenant (student) faces peacekeeping dilemmas**
    - Appears in video offsceen
  - **Artificial agents interact with user**
    - Mentor (e.g., sergeant, front left)
    - Teammates (e.g., medic, front right)
    - Locals (e.g., mother, front center)

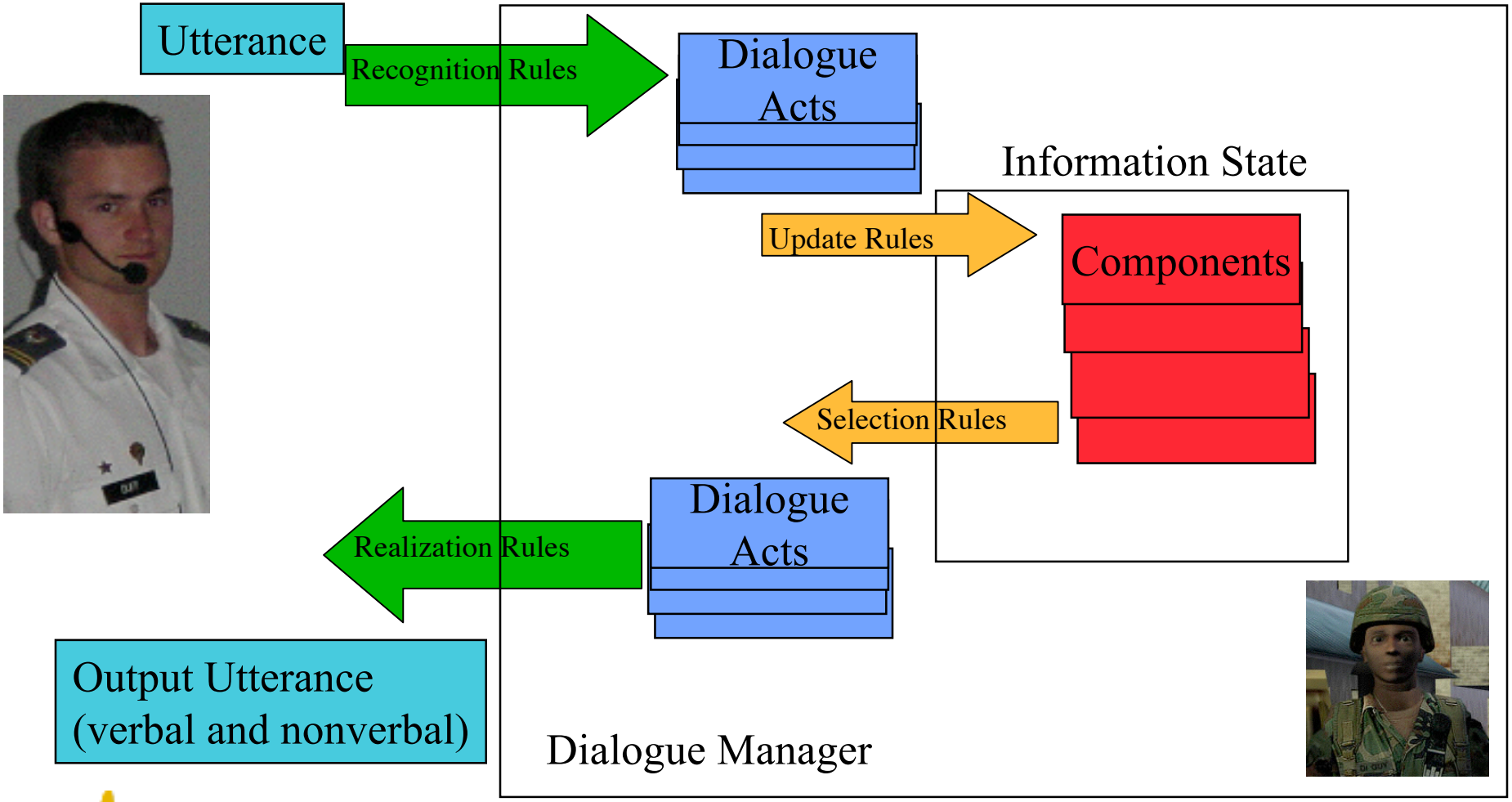
MRE Spoken Language Processing



Dialogue Layers

- **Information State components**
  - Capture coherent aspect of communicative interaction (e.g., turn, grounding, obligations)
- **Dialogue Acts**
  - Recognition Rules
    - Observables + current context
  - Updates: ISC X DA -> ISC
  - Selection rules
  - Realization rules
    - Verbal (NLG)
    - Non-verbal (gesture, other behavior)

Dialogue Processing

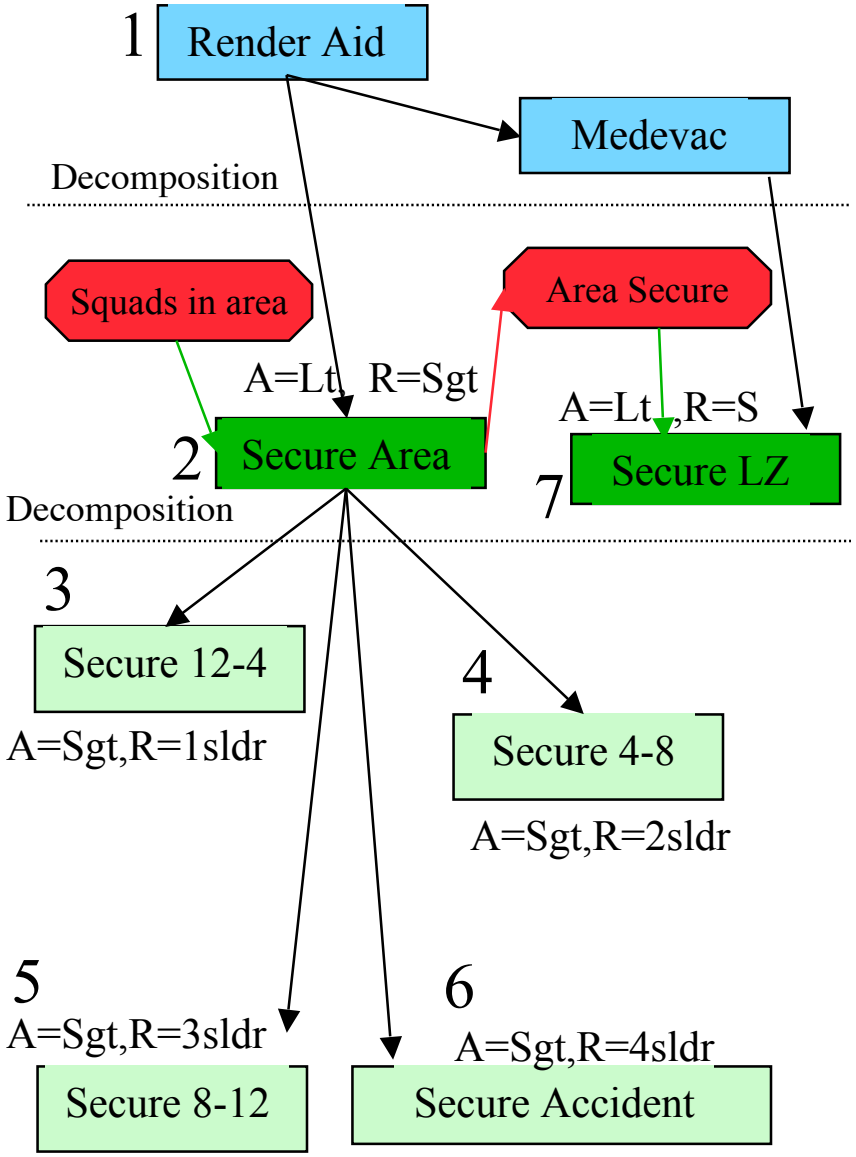


MRE Dialogue Layers (Traum & Rickel AAMAS 2002)

- **Contact**
- **Attention**
- **Conversation**
  - Participants
  - Turn
  - Initiative
  - Grounding
  - Purpose
  - Rhetorical
- **Social**
  - Obligations-Commitments
  - Negotiation-Collaboration
  - Social Roles
- **Individual**
  - Perception
  - Rational
    - belief,desire, intention,..
  - Emotional
    - Coping strategies

  (Marsella & Gratch, yesterday)

Sgt's Negotiation Behavior



Focus=1
Lt: U9 "secure a landing zone"
Committed(lt,7,sgt), 7 authorized, Obl(sgt,U9)
Sgt: U10 "first we should secure the assembly area"
Disparaged(sgt, 7,lt), endorsed(sgt,2.lt), grounded(U9)
Lt: U11"secure the area"
Committed(lt,2,sgt), 2 authorized, Obl(sgt,U11),grounded(U10)
Sgt: U12"yes sir"
Committed(sgt,2,lt), grounded(U11), Push(2,focus)
Goal7:Announce(2,{1sldr,2sldr,3sldr,4sldr})
Goal8: Start-conversation(sgt, ,{1sldr,2sldr,…},2)
    Goal8 -> Sgt: U21 "Squad leaders listen up!"
    Goal7 -> Sgt: U22 "I want 360 degree security"
    Committed(sgt,2,{1sldr,2sldr,3sldr,4sldr})
Push(3, focus)
    Goal9:authorize 3
    Goal9 -> Sgt:U23"1st squad take 12-4"
    Committed(sgt,3, {1sldr,2sldr,3sldr,4sldr}), 3 authorized
Pop(3), Push(4)
    Goal10: authorize 4
    Goal10 -> Sgt: U24"2nd squad take 4-8"
    Committed(sgt,4,{1sldr,2sldr,3sldr,4sldr}), 4 authorized
Pop(4)
    …
    A10: Squads move
    Grounded(U21-U26)
    ends conversation about 2, Happened(2)
Push(7,Focus)

SASO-ST

SASO Dialogue Example

1. Doc: What do you want?
2. Captain: I have orders to assist you in moving this clinic to a safer location

Understand 2:

- Captain takes turn
- acknowledges question
- answers question
- asserts he has an obligation
- but how does it relate to doctor?

SASO Dialogue Example

1. Doc: What do you want?
2. Captain: I have orders to assist you in moving this clinic to a safer location
3. Doc: you want to move the clinic?

Produce 3:

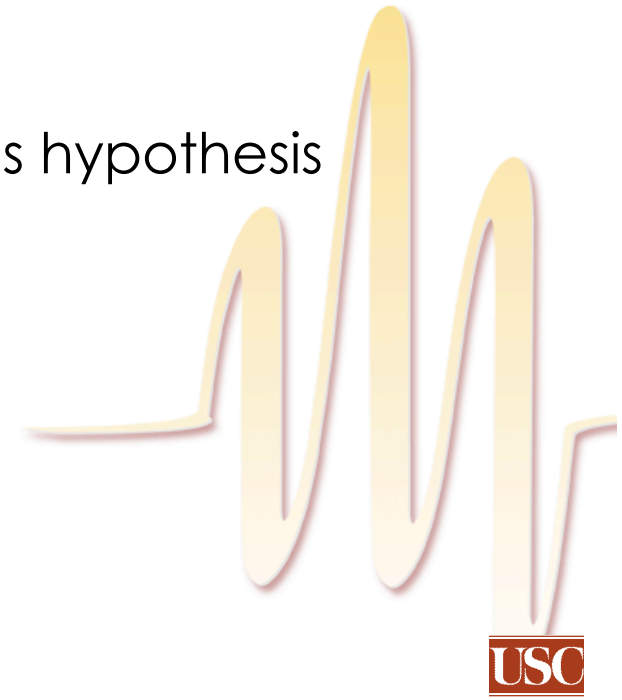
- Doctor attempts to verify current understanding
  - Don't assume most likely understanding
  - Don't ask open question
- Subsequent action depends on reply

SASO Dialogue Example

1. Doc: What do you want?
2. Captain: I have orders to assist you in moving this clinic to a safer location
3. Doc: you want to move the clinic?
4. Captain: Yes

Understand 4:

- Captain takes turn, answers question, verifies hypothesis
- Captain's goal at odds with Doctor
- Topic of Conversation: move clinic



SASO Dialogue Example

1. Doc: What do you want?
2. Captain: I have orders to assist you in moving this clinic to a safer location
3. Doc: you want to move the clinic?
4. Captain: Yes
5. ??

Produce 5:

- Use negotiation strategy to influence response type:
  - Avoid: talk about something else (e.g., casualties)
  - Attack: point out problems with move (e.g., no supplies)
  - Negotiate: evaluate relative merits:
    - bargain (e.g trade supplies)