# 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 for Course

- Monday: Introduction, Architecture of Dialogue Systems, Example Systems
- Tuesday: Simple structures: S-R, IR, finite State
- Wednesday:  Frame-based and Information State
- Yesterday : Plan-based and Logic Based
- Today: Advanced Topics: Grounding, Culture

# Outline for Today

- Grounding
  - Definitions
  - Models:
    - Clark & Schaeffer
    - Traum 94
    - Paek & Horvitz
    - Roque & Traum 08
- Example Information State Grounding System
  - EDIS
- Culture-specific dialogue agents
  - Culture
  - Culture-specific dialogue differences
  - Computational models of culture for agents
  - Examples

# What is Grounding?

- Not electrical grounding
- Not postponing space shuttle flights
- Not crashing a ship onto land
- Not symbol-grounding
- Establishing common ground (Clark & Wilkes-Gibbs ‘86)

# Grounding

- Common Ground
  - How do we model it?
  - How do we achieve it?
- Grounding Models
  - Clark & Schaefer
  - Traum 94
- Grounding & Media

# Models of Common Ground (MK, MB,…)

- Iterated (Schiffer 72)
  - $K_s p \wedge K_A p \wedge K_s K_A p \wedge K_A K_s p \wedge K_s K_A K_s p \wedge \ldots$
- Fixed Point (Harman 77): "A group of people have mutual knowledge of *p* if each knows *p and we know this*, where *this* refers to the whole fact known''
- Shared Situation (Lewis 69): Let us say that it is *common knowledge* in a population P that X if and only if some state of affairs **A** holds such that:
  1. Everyone in P has reason to believe that **A** holds.
  2. **A** indicates to everyone in P that everyone in P has reason to believe that **A** holds.
  3. **A** indicates to everyone in P that X.
- Primitive Attitude
- One-sided (e.g., Cohen '78 BMB)

# How is Common Ground Achieved/Assumed?

- Iterated: proof of individual attitudes
  - Truncation heuristics
  - Circular pointer in deepest beliefs (Cohen 78)
- Shared Situation
  - Observation of situation
  - Assumptions of sharedness (Clark & Marshall)
- Grounding
  - Feedback process

# Types of Feedback (Allwood et al 92)

- Levels:
  - Contact
  - Perception
  - Understanding
  - Attitudinal Reaction
- Signals types
  - Request feedback
  - Prepare other
  - Provide
    - Positive
    - negative

# Clark & Schaefer's contribution model

- Contributions to dialogue are collaborative achievements composed of two phases:
  - Presentation Phase: A presents utterance **u** for B to consider. He does so on the assumption that, if B gives evidence **e** or stronger, he can believe that B understands what A means by **u**
  - Acceptance Phase: B accepts utterance **u** by giving evidence **e'** that he believes he understands what A means by **u**. He does so on the assumption that, once A registers evidence **e'**, he will also believe that B understands.

# Contribution Model

- Each signal is also a presentation to be grounded
  - Recursive model
- Grounding Criterion: ``The contributor and the partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for the current purpose''
- Graded Evidence:
  - Display
  - Demonstration
  - Acknowledgement
  - Initiation of next relevant contribution
  - Continued attention

# Deficiencies of Contribution Model

- Off-line model
  - No way to tell recursion has finished until after the fact
  - No clear specification of moves (for interpretation & generation)
  - Not predictive of next utterances
- Issues with types of evidence

# Computational Model (Traum 94)

- Contribution recast as “DU”
  - (later “CGU”)
- Finite state network for CGU, tracking state of groundedness
- Set of Grounding acts to affect contents and state
- Interpretation and generation rules

# 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

| State | Entering Act | Preferred Exiting Act |
|---|---|---|
| S | ——— | $\text{Initiate}^{I}$ |
| 1 | $\text{Initiate}^{I}$ | $\text{Ack}^{R}$ |
| 2 | $\text{ReqRepair}^{R}$ | $\text{Repair}^{I}$ |
| 3 | $\text{Repair}^{R}$ | $\text{Ack}^{I}$ |
| 4 | $\text{ReqRepair}^{I}$ | $\text{Repair}^{R}$ |
| F | $\text{Ack}^{\{I,R\}}$ | $\text{Initiate}^{\{I,R\}}$ (next DU) |
| D | $\text{Cancel}^{\{I,R\}}$ | $\text{Initiate}^{\{I,R\}}$ (next DU) |

| 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 |

# Grounding Example: Trains

| UU# | Speaker: Utterance | | Act(s) | DU States 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|
| 3.3 | M: let's see | : | $init_1$ | 1 | | | |
| 3.4 | : where are there oranges | : | $cont_1$ | 1 | | | |
| 4.1 | S: the oranges are in the warehouse | : | $ack_1$,$init_2$ | F | 1 | | |
| 4.2 | : at Corning | : | $cont_2$ | F | 1 | | |
| 5.1 | M: oh okay | : | $ack_2$ | F | F | | |
| 5.2 | : and I see that there's a tanker car there | : | $init_3$ | F | F | 1 | |
| 5.3 | : oh we don't want a tanker car do we | : | $cancel_3$ | F | F | D | |
| 5.4 | : um | : | | F | F | D | |
| 5.5 | : I have to get a boxcar | : | $init_4$ | F | F | D | 1 |
| 5.6 | : to Corning | : | $cont_4$ | F | F | D | 1 |
| 5.7 | : and then I have to load it with oranges and eventually I have to get that to Bath | : | $cont_4$ | F | F | D | 1 |
| 5.8 | : by 8 o'clock | : | $cont_4$ | F | F | D | 1 |
| 6.1 | S: right | : | $ack_4$ | F | F | D | F |

# EDIS SYSTEM

- Uses PTT theory
- Trindikit implementation
- Autoroute domain

# PTT Informational Components

- Separate Views for System and User
  (System assumptions about User)
- Private, Public, and Semi-public components of View captures *grounding* process (Clark& Schaefer '87)
  - GND represents common ground
  - set of DUs represent partitioned semi-public information introduced but not (yet) grounded
  - UDUs structure accessible ungrounded DUs
- (Semi-)Public Information includes:
  - public events
  - social commitments of participants
- Private Information includes
  - Intentions
  - Beliefs

# EDIS Formalization of Information Components

- Record (AVM) for Views, with fields for each dialogue participant:
  - GND: PT-Rec
    Public Information
  - UDUS: list of accessible DU IDs
  - CDU (DU-ID,PT-Rec)
    current Discourse Unit
  - PDU (DU-ID,PT-Rec)
    penultimate Discourse Unit
  - INT: list of intended actions

- PT-REC contains:
  - DH: list of dialogue acts
    Dialogue History of performed dialogue acts
  - OBL: list of action types
    Obligations of participants to perform actions
  - SCP: list of states
    Social Commitments of agents to Propositions
  - COND: list of implications
    relevant conditional anticipated effects

# PTT Information State

$$
\begin{bmatrix}
\text{G} & : & \textit{PT-R} & & \\
\text{CDU} & : & \begin{bmatrix} \text{C} & : & \textit{PT-R} \\ \text{ID} & : & \text{DU-ID} \end{bmatrix} \\
\text{PDU} & : & \begin{bmatrix} \text{C} & : & \textit{PT-R} \\ \text{ID} & : & \text{DU-ID} \end{bmatrix} \\
\text{UDUs} & : & \text{List(DU-ID)} & & \\
\text{INT} & : & \text{List(Action)} & &
\end{bmatrix}
$$

$$
\textit{PT-R} : \begin{bmatrix}
\text{DH} & : & \text{List(Action)} \\
\text{OBL} & : & \text{List(Action)} \\
\text{SCP} & : & \text{List(Prop)} \\
\text{COND} & : & \text{List(Action)}
\end{bmatrix}
$$

# EDIS Dialogue Moves

- Forward-looking
  - assert(dp,Prop)
  - check(dp, Prop)
  - direct (dp,act-type)
  - info-request(dp,Q)
- Backward Looking
  - Address(dp,act)
    - accept
    - agree
    - answer
  - Understanding Act
    - Acknowledge(dp,DU-ID)

# Update Strategy

- Deliberation (produce new intentions)
- Acting on intentions (produce output dialogue moves)
- Update based on an observed utterance
  1. Create a new DU and push it on top of UDUs.
  2. Perform updates for backwards grounding acts.
  3. For other types, record in cdu.dh and apply the update rules for act class
  4. Apply inference update rules to all parts of the IS which contain newly added acts.

# Update Rules

- effects of observed dialogue acts
  - formalized in terms of social commitments
- inference
  - Obligation Resolution
  - Conditional Resolution
  - Intention Resolution
- Deliberation
  - adopting new intentions

# Dialogue Act Effect Updates

| | |
|---|---|
| act | ID:2, **ack**(DP,DU1) |
| effect | peRec(w.Gnd,w.pdu.tognd) |
| effect | remove(DU1,UDUS) |
| act | ID:c, **forward-looking-act**(DP) |
| effect | push(obl,**u-act**(**o**(DP),CDU.id)) |
| act | ID:2, **accept**(DP,ID2) |
| effect | *accomplished via rule resolution* |
| act | ID:2, **agree**(DP,ID2) |
| effect | push(scp,**scp**(DP,**P**(ID2))) |
| act | ID:2, **answer**(DP,ID2,ID3) |
| effect | push(scp,**ans**(DP,**Q**(ID2),**P**(ID2))) |
| act | ID:2, **assert**(DP,PROP) |
| effect | push(scp,**scp**(DP,PROP)) |
| effect | push(cond,**accept**(**o**(DP),ID)→ **scp**(**o**(DP),PROP)) |
| act | ID:1, **assert**(DP,PROP) |
| effect | push(cond,**accept**(**o**(DP),ID)→ **scp**(**o**(DP),PROP)) |
| act | ID:2, **check**(DP,PROP) |
| effect | push(obl,**address**(**o**(DP),ID)) |
| effect | push(cond,**agree**(**o**(DP),ID) → **scp**(DP,PROP)) |
| act | ID:2, **direct**(DP,Act) |
| effect | push(obl,**address**(**o**(DP),ID)) |
| effect | push(cond,**accept**(**o**(DP),ID) → **obl**(**o**(DP),Act)) |
| act | ID:2, **info_request**(DP,Q) |
| effect | push(obl,**address**(**o**(DP),ID)) |

# Deliberation Factors

- obligations
  - to perform understanding acts
  - to address previous dialogue acts
  - to perform other actions
- potential obligations
  that would result if another act were performed,
  as represented in the cond field (or CDU.OBL)
- insufficiently understood dialogue acts
  with a 1 confidence level in cdu.dh
- intentions to perform complex acts

# Deliberation Rules

1. Grounding:
   OBL U-act, everything in CDU understood
   $\Rightarrow$ ack(W,CDU)

2. Address:
   OBL address act
   $\Rightarrow$ accept, agree, or answer

3. Anticipatory Planning:
   INT act1 $\wedge$ COND act1 $\rightarrow$ OBL act2
   $\Rightarrow$ act2 add an intention to perform an action

4. SubGoal: Int(act1) $\wedge$ NextSubact(Act1,Act2)
   $\Rightarrow$ Act2

   (a) check CDU.DH:1

   (b) info-request

# Sample Autoroute Dialogue

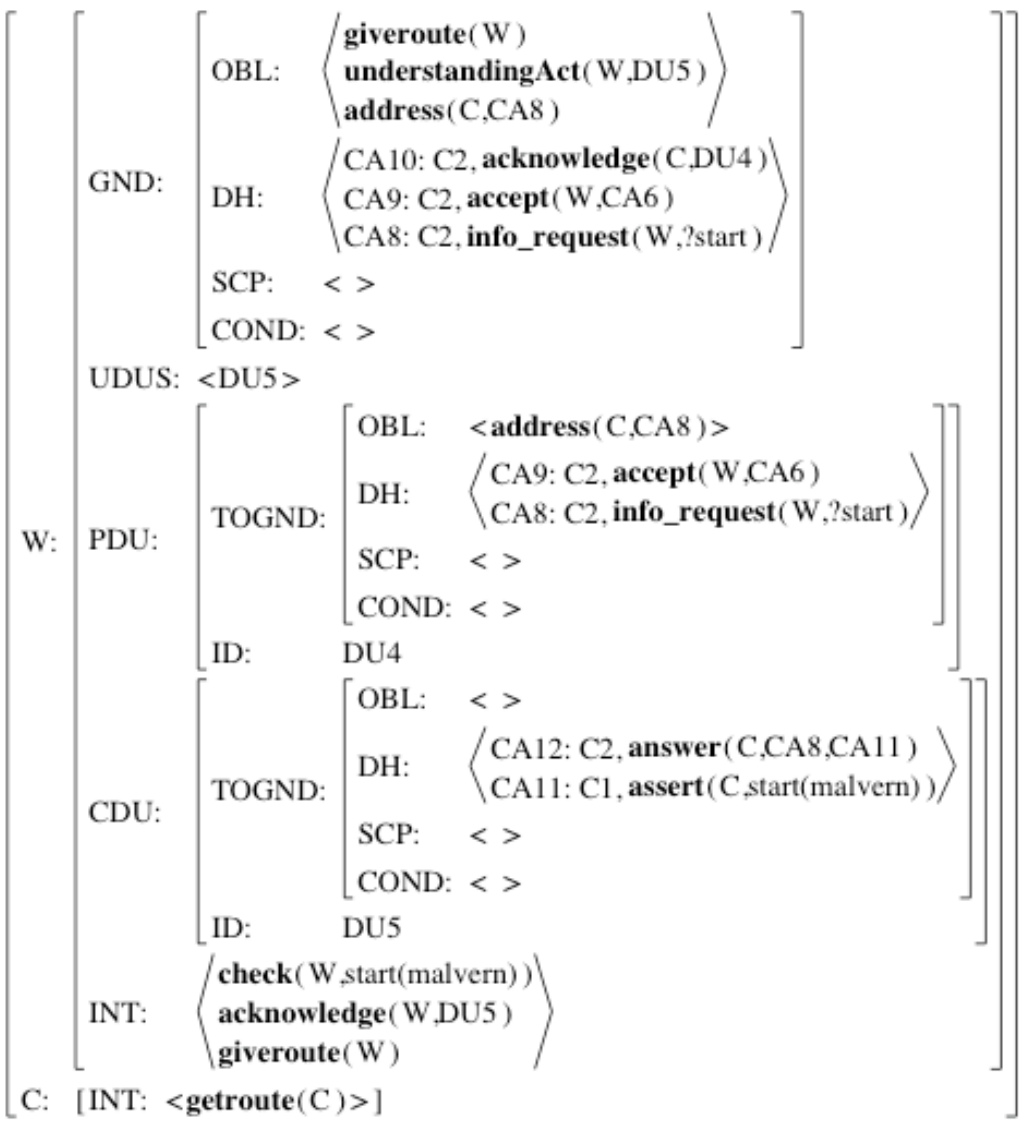
| W **WIZARD** | **CALLER** |
|---|---|
| [1]: How can I help you? | [2]: A route please |
| [3]: Where would you like to start? | [4]: Malvern |
| [5]: Great Malvern? | [6]: Yes |
| [7]: Where do you want to go? | [8]: Edwinstowe |
| [9]: Edwinstowe in Nottingham? | [10]: Yes |
| [11]: When do you want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

# InfoState after [2]: A route please

W:
- GND:
  - OBL: ⟨ **understandingAct**(W,DU3), **address**(C,CA2) ⟩
  - DH: ⟨ CA3: C2, **acknowledge**(C,DU2), CA2: C2, **info_request**(W,?helpform) ⟩
  - SCP: < >
  - COND: < >
- UDUS: <DU3>
- PDU:
  - TOGND:
    - OBL: <**address**(C,CA2)>
    - DH: <CA2: C2, **info_request**(W,?helpform)>
    - SCP: < >
    - COND: < >
  - ID: DU2
- CDU:
  - TOGND:
    - OBL: <**address**(W,CA6)>
    - DH: ⟨ CA6: C2, **direct**(C,giveroute(W)), CA5: C2, **answer**(C,CA2,CA4), CA4: C2, **assert**(C,want(C,route)) ⟩
    - SCP: <**scp**(C,want(C,route))>
    - COND: <**accept**(W,CA6) **->** **obl**(W,giveroute(W))>
  - ID: DU3
- INT: ⟨ **info_request**(W,?start), **giveroute**(W), **accept**(W,CA6), **acknowledge**(W,DU3) ⟩

C: [INT: <**getroute**(C)>]

# InfoState after [4]: Malvern, prompting check



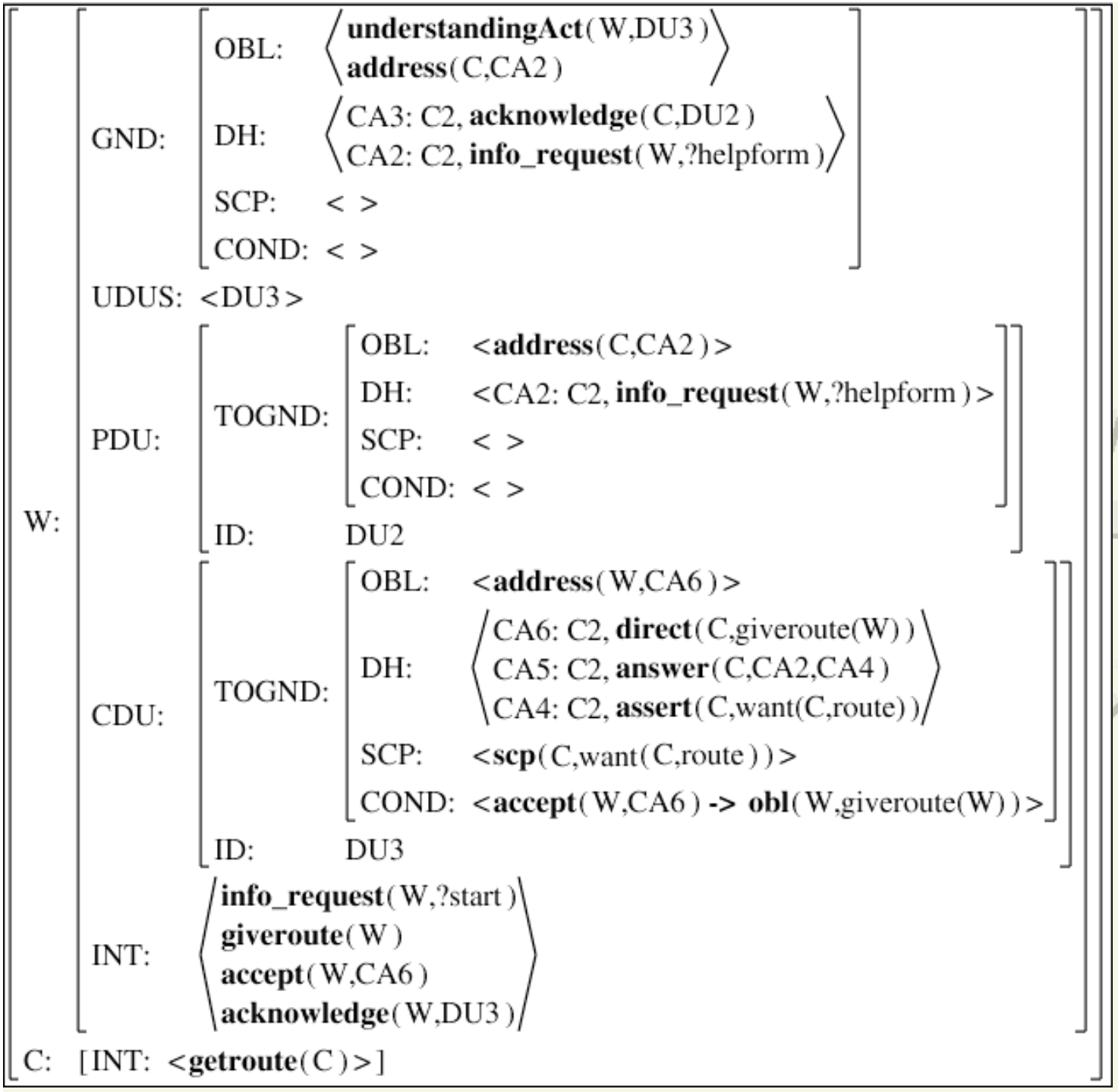
# InfoState after [5]: Great Malvern?

W:
- GND:
  - OBL: ⟨ **understandingAct**(C,DU6), **giveroute**(W) ⟩
  - DH: ⟨ CA13: C2, **acknowledge**(W,DU5), CA12: C2, **answer**(C,CA8), CA11: C1, **assert**(C,start(malvern)) ⟩
  - SCP: < >
  - COND: < >
- UDUS: <DU6>
- PDU:
  - TOGND:
    - OBL: < >
    - DH: ⟨ CA12: C2, **answer**(C,CA8,CA11), CA11: C1, **assert**(C,start(malvern)) ⟩
    - SCP: < >
    - COND: < >
  - ID: DU5
- CDU:
  - TOGND:
    - OBL: <**address**(C,CA14)>
    - DH: <CA14: C2, **check**(W,start(malvern))>
    - SCP: < >
    - COND: <**agree**(C,CA14) **->** **scp**(W,start(malvern))>
  - ID: DU6
- INT: <**giveroute**(W)>

C: [INT: <**getroute**(C)>]

## InfoState after [7]: Where do you want to go?

$$
\begin{bmatrix}
W: & \begin{bmatrix}
\text{GND:} & \begin{bmatrix}
\text{OBL:} & \left\langle \begin{array}{l} \textbf{understandingAct}(\text{C,DU8}) \\ \textbf{giveroute}(\text{W}) \end{array} \right\rangle \\
\text{DH:} & \left\langle \begin{array}{l} \text{CA17: C2,} \textbf{acknowledge}(\text{W,DU7}) \\ \text{CA16: C2,} \textbf{agree}(\text{C,CA14}) \end{array} \right\rangle \\
\text{SCP:} & \left\langle \begin{array}{l} \textbf{scp}(\text{C,}\textbf{start}(\text{malvern})) \\ \textbf{scp}(\text{W,}\textbf{start}(\text{malvern})) \end{array} \right\rangle \\
\text{COND:} & \langle\ \rangle
\end{bmatrix} \\
\text{UDUS:} & \langle \text{DU8} \rangle \\
\text{PDU:} & \begin{bmatrix}
\text{TOGND:} & \begin{bmatrix}
\text{OBL:} & \langle\ \rangle \\
\text{DH:} & \langle \text{CA16: C2,} \textbf{agree}(\text{C,CA14}) \rangle \\
\text{SCP:} & \langle \textbf{scp}(\text{C,}\textbf{start}(\text{malvern})) \rangle \\
\text{COND:} & \langle\ \rangle
\end{bmatrix} \\
\text{ID:} & \text{DU7}
\end{bmatrix} \\
\text{CDU:} & \begin{bmatrix}
\text{TOGND:} & \begin{bmatrix}
\text{OBL:} & \langle \textbf{address}(\text{C,CA18}) \rangle \\
\text{DH:} & \langle \text{CA18: C2,} \textbf{info\_request}(\text{W,?dest}) \rangle \\
\text{SCP:} & \langle\ \rangle \\
\text{COND:} & \langle\ \rangle
\end{bmatrix} \\
\text{ID:} & \text{DU8}
\end{bmatrix} \\
\text{INT:} & \langle \textbf{giveroute}(\text{W}) \rangle
\end{bmatrix} \\
\text{C:} & [\text{INT:} \langle \textbf{getroute}(\text{C}) \rangle]
\end{bmatrix}
$$


W:
GND:
OBL: understandingAct(W,DU3)
address(C,CA2)
DH: CA3: C2, acknowledge(C,DU2)
CA2: C2, info_request(W,?helpform)
SCP: < >
COND: < >
UDUS: <DU3>
PDU:
TOGND:
OBL: <address(C,CA2)>
DH: <CA2: C2, info_request(W,?helpform)>
SCP: < >
COND: < >
ID: DU2
CDU:
TOGND:
OBL: <address(W,CA6)>
DH: CA6: C2, direct(C,giveroute(W))
CA5: C2, answer(C,CA2,CA4)
CA4: C2, assert(C,want(C,route))
SCP: <scp(C,want(C,route))>
COND: <accept(W,CA6) -> obl(W,giveroute(W))>
ID: DU3
INT: info_request(W,?start)
giveroute(W)
accept(W,CA6)
acknowledge(W,DU3)
C: [INT: <getroute(C)>]

# Recognizing Grounding Acts

- Initiate:  core acts, no ungrounded CGU
- acknowledge: evidence of understanding (backward act, explicit, follow-up)
- Request-repair:clarify-parameter, or repetition request
- Repair: providing changing or solicited info

# Grounding Act Updates

- initiate:
  - New CGU, state -> 1, obligation to ground
- continue:
  - New content added to CGU
- Request-repair
  - State -> 2, obligation to repair
- Repair
  - State-> 1, change content
- Acknowledge
  - State -> F, content effects
- Cancel
  - State -> D, remove CGU from ^grounding, recent-cgus, remove grounding obligations for CGU

# Open Problems with this Model

- Binary grounded/ungrounded decision
  - No levels of "groundedness"
- Leaves the unit size unspecified
- Confusability of grounding acts
  - e.g. repetition = acknowledgment, repair, or request for repair?
- Only well-suited for spoken language grounding

## Levels of Analysis: **Quartet**: Paek & Horvitz 2000

lowest

highest

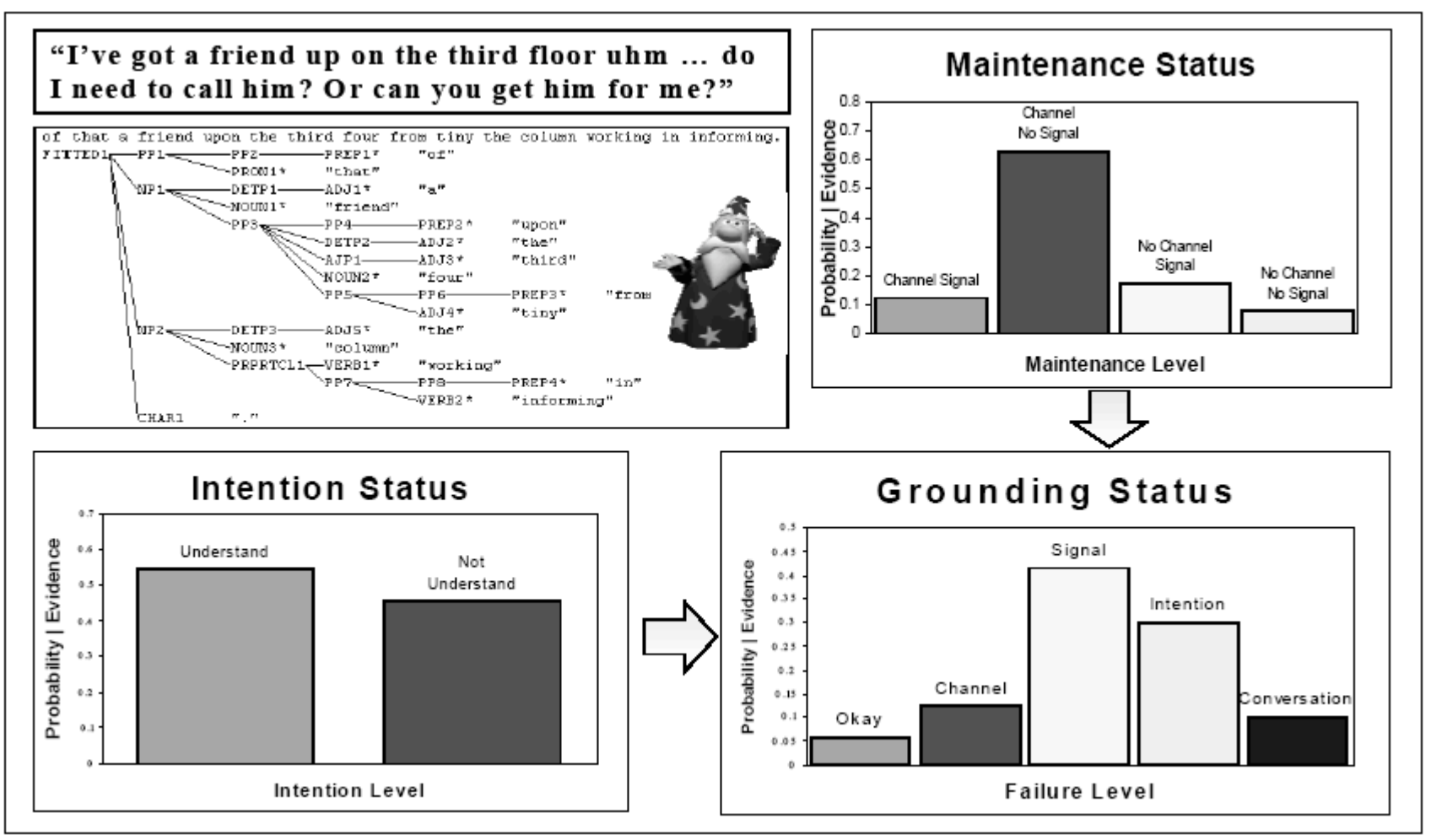
- **Channel Level**: attempt to open communication channel with some behavior
- **Signal Level**: behavior is intended as a signal
- **Intention Level**: understanding of semantic content occurs
- **Conversation Level**: a joint activity is proposed and responded to

*All levels require coordination between speaker and listener

# System Design

- Two modules:
  - **maintenance** — **Signal & Channel** level
  - **intention** — **Intention** level
- **Conversation Control** — **Conversation** level
  - exchanges info between the modules
  - determines grounding state
  - weighs costs and benefits
  - evaluates module performance & reliability

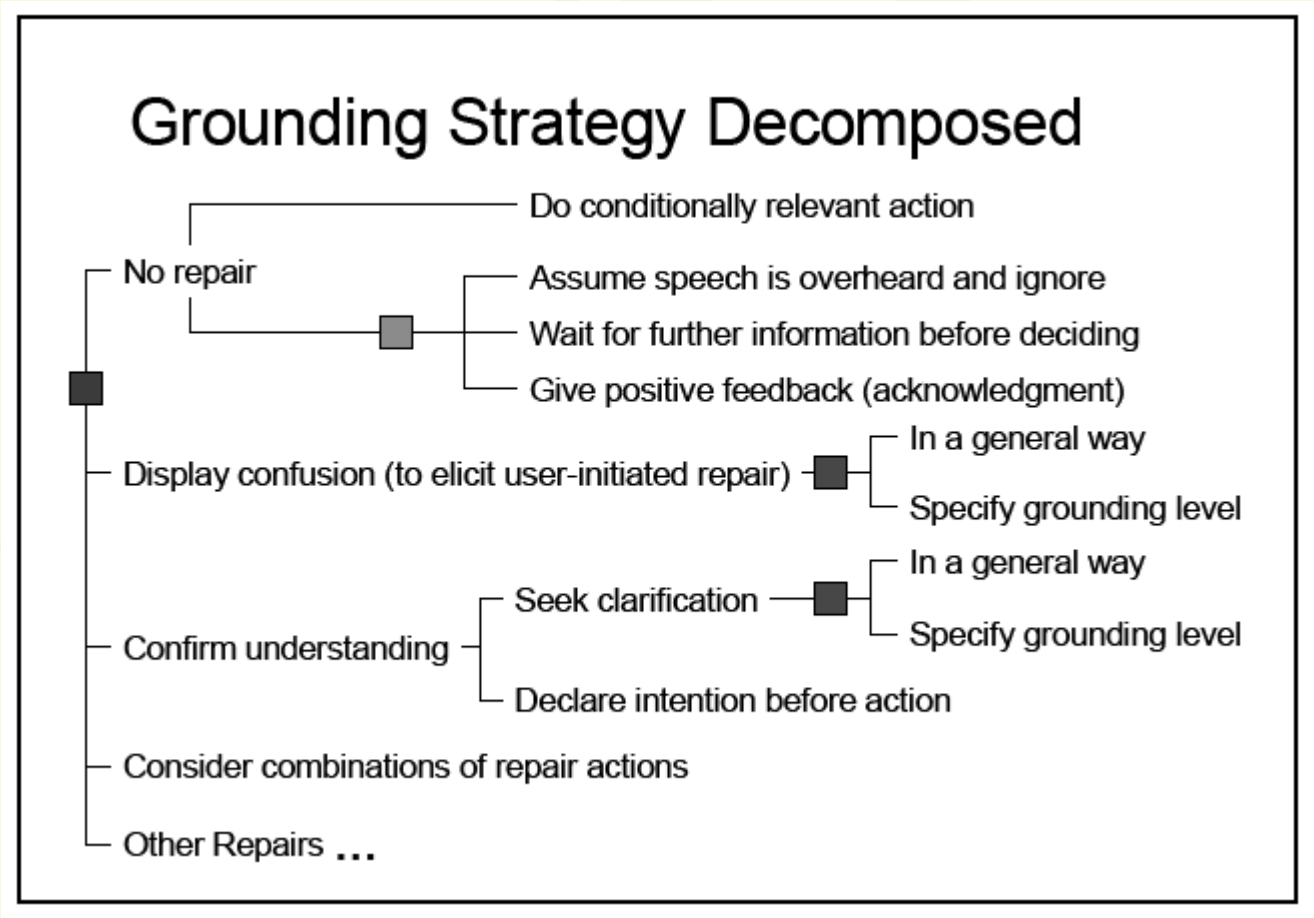
# Signal Failure



# Benefits of this Design

- ASR can model probabilistic dependencies among levels
- Easier to pinpoint and fix problems in system understanding
- Models psychological strategies for grounding on lower levels first
- Flexibility in multiple domains: simply changing the **intention** module

# Grounding Strategies



# Grounding Issues

- How is a particular grounding act realized?
- How important is the grounding?
  - How useful will it be to the system?
- What criteria are needed?
- How well will a particular act ground its intended content?
- And what is the opportunity cost of performing this act?
  - Is it worth it?

# Factors Affecting Grounding Behavior

- Amount of grounding, type of act, content & realization of act, and model for groundedness depends on a number of factors including
  - Purposes& prior groundedness (Grounding Criterion)
  - Available communication channels and resources
    - Costs and affordances: Clark and Brennan '90
    - Traum & Heeman '96: only 3-5% of utterances in spoken trains corpus had no grounding
    - Dillenbourg & Traum '96, 05: over 50% of utterances in typed MOO mystery solving dialogues had no grounding
  - Content
    - Dillenbourg & Traum '96, 05
    - Sometimes shared situation model is better than explicit grounding model (for facts on shared whiteboard)

# Media and Activity factors in Grounding

- Clark and Brennan 90
  - Media influences amount and type of grounding
- E.g., Traum & Heeman '96: Trains Domain, spoken language, no visual contact

| Category | % utterances |
|---|---|
| Explicit Ack | 52% |
| Related | 29% |
| Unrelated after Explicit | 15% |
| Other Unrelated | 3% |
| Uncertain | 2% |

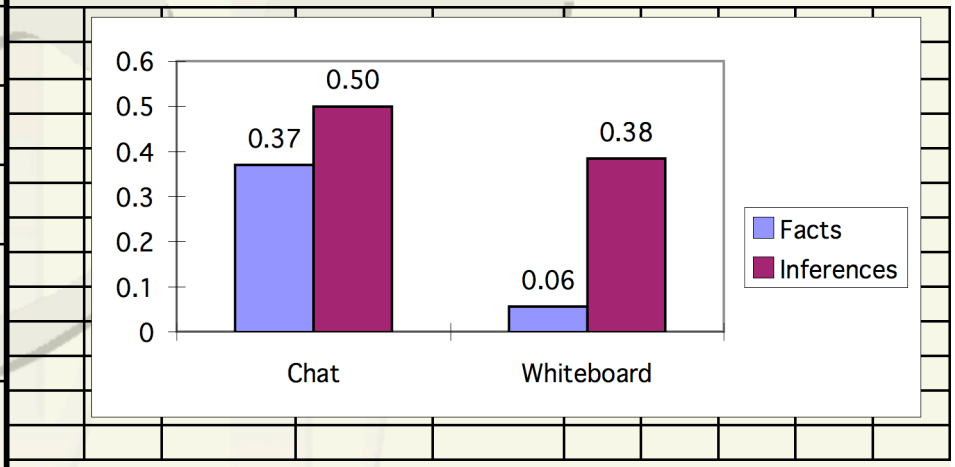
# Dillenbourg & Traum 96, 05

## Multi-modal computer-mediated grounding

- Grounding by category

| Content of interactions | Acknowledgment Rate |
|---|---|
| Task knowledge | 38% |
| Facts | 26% |
| Inferences | 46% |
| Task management | 43% |
| Meta-Communication | 55% |
| Technical problems | 30% |
| *All categories* | *41%* |

- Grounding by Category & Medium



# Degrees of Grounding model (Roque & Traum Sigdial 2008)

- defines degrees of groundedness
  - implied in concept of grounding criteria
  - implied in concept of strength of evidence of understanding
- distinguishes degrees of groundedness from evidence of understanding

# Model Components

- **set of types of** Evidence of Understanding
- **set of** Degrees of Groundedness
- Grounding Criteria **for each information element**
- Algorithms for dialogue management
  - given: interpretation, grounding criteria, history of Evidence
  - identify Evidence and Degrees
  - determine action

# Types of Evidence of Understanding

- Identifying Set of Types of Evidence
  - Began with list of Evidence from (Clark & Schaefer, 89)
  - Modified during analysis
- Corpus: JFETS-UTM, Call for Fire training at Ft. Sill
  - human FOs and FDCs, no automation
  - 4 sessions, 17 missions, 456 utterances
  - 1222 markables
    - 886 dialogue move parameters
    - 336 periods of silence

# Types of Evidence of Understanding

- Example:
  - **Submit:** derived from Clark & Schaefer 'Presentation' phase
  - **Repeat Back:** related to Clark & Schaefer's "Display" evidence
  - **Acknowledge:** from Clark & Schaefer

```
G91    direction 6120 over        Submit
S19    direction 6120 out         Repeat Back
G91    roger out                  Acknowledge
```

# Types of Evidence of Understanding

- Example:
  - **Request Repair**: next turn repair initiator
  - **Resubmit**: third turn repair, in this example

```
G91    grid 5843948 over          Submit
S19    say again grid over        Request Repair
G91    grid 5843948 over          Resubmit
```

# Types of Evidence of Understanding

- Example: Move On:
  - derived from Clark & Schaefer's "Initiation of the relevant next contribution"
  - example below: G91 would not submit *grid* if they did not consider the *fire for effect* to be grounded.

```
G91   fire for effect over      Submit
S19     fire for effect out     Repeat Back
G91     grid 4542368 over       Submit, Move On
```

# Types of Evidence of Understanding

- Lack of Response:
  - in the first example below, a reply is expected; suggests lack of grounding
  - in the second example below, the response is optional; suggests that neither speaker has an objection to the submission

| | | |
|---|---|---|
| G91 | S19 this is G91 over | Submit |
| | **(12 seconds of silence)** | **Lack of Response** |
| G91 | S19 this is G91 over | Resubmit |
| - - - - - - - - - - | | |
| G91 | b m p in the open over | Submit |
| S19 | b m p in the open out | Repeat Back |
| | **(10 seconds of silence)** | **Lack of Response** |

# Types of Evidence of Understanding

- Standalone
  - initially **Submit** material
  - generally **Acknowledge**
  - speaker B **Repeats Back** material presented by speaker A
  - speaker B makes a **Repair Request** of speaker A
  - speaker A **Resubmits** material speaker A previously presented
- Additional
  - **Uses** material previously introduced by speaker A
  - **Moves On** in terms of steps to task completion
- Silence-Related
  - after an utterance, a **Lack of Response**

# Degrees of Groundedness

- Given evidence related to an information component,
  what can we say about how grounded it is?
- Define Degrees of Groundedness before/after Evidence

*What is the Degree before the Submit?*

**G91 fire for effect over** **Submit**

*What is the Degree after the Submit?*

**S19 fire for effect out** **Repeat Back**

*What is the Degree after the Repeat Back?*

**G91 grid 4542368 over** **Submit, Move On**

*What is the Degree after the Move On?*

# Degrees of Groundedness

- Unknown - **material has not yet been introduced**
- Misunderstood - **after a Request Repair**
- Unacknowledged - **after a Submit, Lack of Response**
- Accessible - **after a Submit or Resubmit**
- Agreed-Signal - **after an Submit, Acknowledgment**
- Agreed-Signal+ - **after a Submit, Acknowledgments, other**
- Agreed-Content - **after a Submit + Repeat Back**
- Agreed-Content+ - **after a Submit + Repeat Back + Acknowledgment(s) / other**
- Assumed - **both participants already know material**

# Algorithms for Dialogue Management

- Identify Evidence
  - given dialogue act interpretation and history of Evidence
  - rules based on definitions of Evidence
- Identify Degree of Groundedness
  - given current Degree and history of Evidence
  - rules based on observations of Evidence patterns
- Decide on grounding action
  - given current Degree and Grounding Criterion
  - decide on Evidence to provide (or not)

# *Notes on Enculturated Interfaces and culture-specific aspects of communication*

David Traum, USC Institute for Creative Technologies

USC

ICT

INSTITUTE FOR CREATIVE TECHNOLOGIES

# What is Culture?

- **Kroeber & Kluckholm (52): 164 different definitions**
- **Allwood (85): all the characteristics common to a particular group of people that are learned and not given by nature.**
  1) **Patterns of thought**: common ways of thinking, where thinking includes factual beliefs, values, norms, and emotional attitudes.
  2) **Patterns of behavior**: common ways of behaving, from ways of speaking to ways of conducting commerce and industry, where the behavior can be intentional/unintentional, aware/unaware or individual/interactive.
  3) **Patterns of artifacts** : common ways of manufacturing and using material things, from pens to houses (artifact = artifical object),
  4) **Imprints in nature**: the long-lasting imprints left by a group in the natural surroundings, where such imprints include agriculture, trash, roads or intact/ruined human habitations.

# Working definitions of culture for communication

- **A store of knowledge used to assign meaning to behaviors in context**
  - range of conceivable/expressible meanings
  - Ontology of relevant contextual elements
  - Ontology of meaningful behaviors
  - Behavior+context ---> meaning relationship
- **A group who assume that this knowledge is shared within the group**
  - National culture
  - Ethnic or religious group culture
  - Institutional culture
  - Professional culture
  - Family/clan culture

# Culturally-Specific Behavior

- **How universal are behaviors, and**

**Behavior+context --> meaning relationships?**

  - Universal
  - Culture-specific
  - Activity or role-specific
  - Individualistic

- **What are the differences?**
  - different behaviors
  - different meanings
  - different situations
  - Different mappings
  - different frequencies

# How do you enculturate an interface?

- **Incorporate behavior+context ---> meaning relationship for behaviors and interpretations of interface**
  - But this is true of all interfaces!
  - Just as for embodiment, emotion, cognition: Can't opt out of the meaning game
    - People will attribute meaning (whether intended or not)
- **All interfaces are enculturated - just a question of which culture**
  - Candidates
    - Assumed universal UI culture
    - Designers' culture
    - Badly-designed interface culture (e.g. software emanating from Redmond Washington)
  - Why does specific interface culture matter?
    - Communication may be inefficient/hard to learn
    - Crossed cultures (actor & interpreter) leads to misunderstanding

# Why adopt a specific (existing) culture?

- **Ease of understanding/efficiency**
- **Virtual Humans/ECAs**
  - Realism
  - Culture-training
  - Culture translation

# At what level to add culture?

- **Behavior**
  - Performance
  - Frequency
- **Cognition**
  - Meaning & Context-specific, goal-directed Behavior Generation
  - Meaning & context-specific interpretation
- **Behavior is easier to induce from corpus study, ultimately less useful/appropriate for micro-analysis/specific attribution**

# Computational Models: Parameterizable Culture models

- General (Universal?) behavior <--> meaning template
- Culture fills in parameter values to complete the relation
- Parameter values determinable from observation

# Corpus-based culture-studies

- **Method:**
  - Record lots of data of participants from different cultures
  - Try to normalize other factors (activities, relationships, status, setting,…)
  - Recover regularities (and ideally meanings) in behaviors
  - Fill in parameter values
- **Dangers:**
  - How big is the relevant culture (national? Ethnic? Age-group? clique?)
  - How universal are the findings?
    - Many differences in behaviors within cultures
      - Gender
      - Age
      - Status
      - relationship
  - How representative are the participants?
  - What else is going on/confusion/meaning modifiers
  - How natural is the experimental setting?

# Aspects of Communication

- **Verbal Language: phonemes, morphemes, words, sentences**
- **Non-verbal behavoirs:**
  - Proxemics
  - Gaze
  - Facial Expressions
  - Body Posture
  - Hand Gestures
  - Prosody & Intonation
- **Social interaction**
  - Turn-taking
  - Greetings & closings
  - Sequential interaction
  - Grounding
  - Boosting & downplaying
  - Ritualized behavior

# Speech act distribution (Traum 2000 Journal of Semantics)

- **Frequency of understanding and answer acts, within tasks, across subject populations**

| Damsl TRAINS | Damsl Monroe | SWBD-Damsl Switchboard | HCRC HCRC Maptask | HCRC DCIEM Maptask | Verbmobil II Verbmobil English | Verbmobil II Verbmobil German | Verbmobil I Verbmobil I German |
|---|---|---|---|---|---|---|---|
| statement 45.9 | 51.4 | 49 | explain 7.9 | 7.9 | Inform,... 22.8 | 21.2 | 12.2 |
| info-request 15.2 | 9.9 | questions 4.9 | query,check,align 23.5 | 20.3 | | | |
| action-dir,oo 12.2 | 12.9 | 0.7 | instruct 15.6 | 15.2 | request,suggest 26.0 | 27.0 | 32 |
| commit,offer 23.8 | 16.8 | 0.1 | | | commit 0.5 | 0.8 | |
| conventional 2.5 | 0.6 | 1.4 | | | 13.4 | 15.6 | 16.5 |
| answer 14.7 | 8.4 | 3 | reply,clarify 22.8 | 20 | feedback 15.2 | 9.8 | 0.6 |
| accept 30.0 | 23.0 | 5 | | | accept,confirm 10.3 | 12.3 | 13.5 |
| reject 2.2 | 0.5 | 0.2 | | | reject,explained 3.3 | 4.4 | 8.2 |
| other agree 3.6 | 1.8 | 0.3 | | | clarify 2.3 | 1.9 | 8.9 |
| Understanding 30.2 | 28.5 | 23 | acknowledge 20.5 | 28.1 | backchannel 3.6 | 3.3 | |
| non-understand 1.2 | 0.5 | 0.1 | | | | | |

Table 1: Percentage Distributions of Dialogue Acts in Corpus Coding



USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

# Ex: Proxemics

- **How physically close to stand to someone of a particular relationship in conversation?**
  - Close enough to hear speech?
  - Far enough to be able to bow?
  - Close enough to shake hands?
  - Close enough to smell?
  - Close enough to feel breath?
  - Far enough to not?
- **How close is uncomfortably "too" close**
- **How far is too far?**

# Ex 2: Cultural Variability in Turn-taking Anglo vs Native Americans from Warm Springs Reservation (Philips '76)

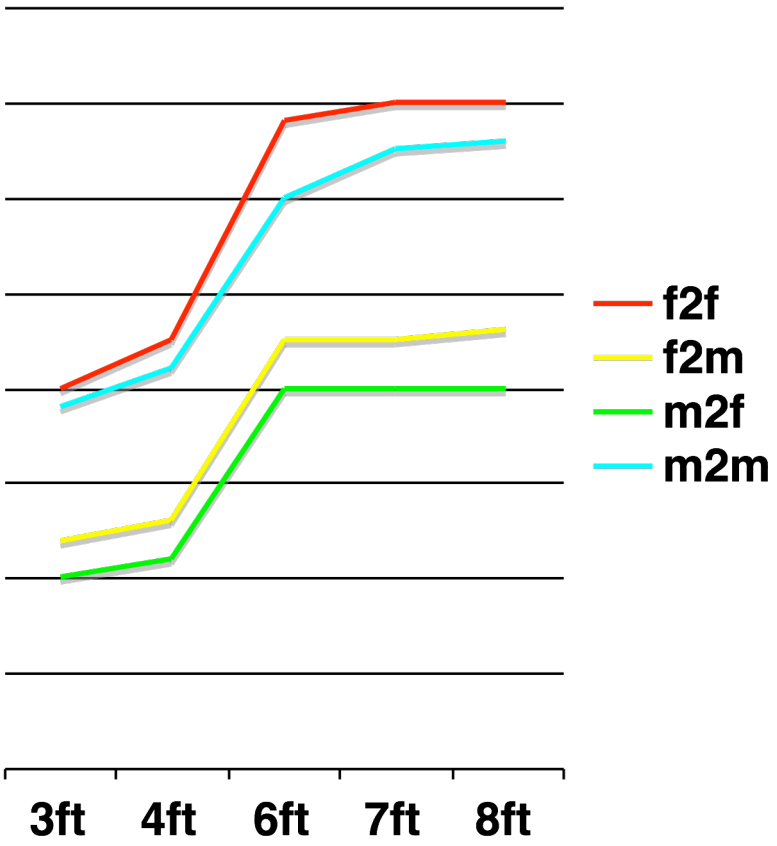
- **Anglo pattern**
  - Analysis from Striegnitz yesterday
    - (e.g., Sacks & Schegloff, Duncan, Kendon, Goffman)
  - Speaker and addressee signals & regulation
  - Sequential relationship (adjacency pairs)
  - Obligation to respond (quickly to questions)
- **Warm Springs pattern**
  - Less regulation of speakers
  - Less selection of addressees
  - No pressure to respond immediately to questions
    - Questions can remain on floor longer

# Warm Springs (Native American) pattern:

- **Fewer distinctions between addressees and other hearers**
- **Large distances between questions and answers (15 minutes, 30 minutes,…)**
- **Slower pace**
- **Longer pauses**
- **Interruptions rare**
- **More evenly distributed talk**
- **Less body motion**
  - Less head alignment
  - fewer posture shifts
  - No head-bobbing while talking
  - Fewer arm and head gestures
  - Arms kept closer to body, fingers rarely open
- **More facial and eye movement**
  - Widening of eyes
  - Eye movement
  - Brow movement
  - Shorter gazes (for speakers and hearers)

# Ex 3: Interactions: Gaze, distance, gender (Argyle & Dean '65)

- Women looked more than men at same and other gender
- Same sex looked much more
- Everyone looks more when further
- Distance and gaze in complementary relationship at establishing closeness





USC
ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

# ICT&UTEP Work on Culture-specific behaviors: Methodology

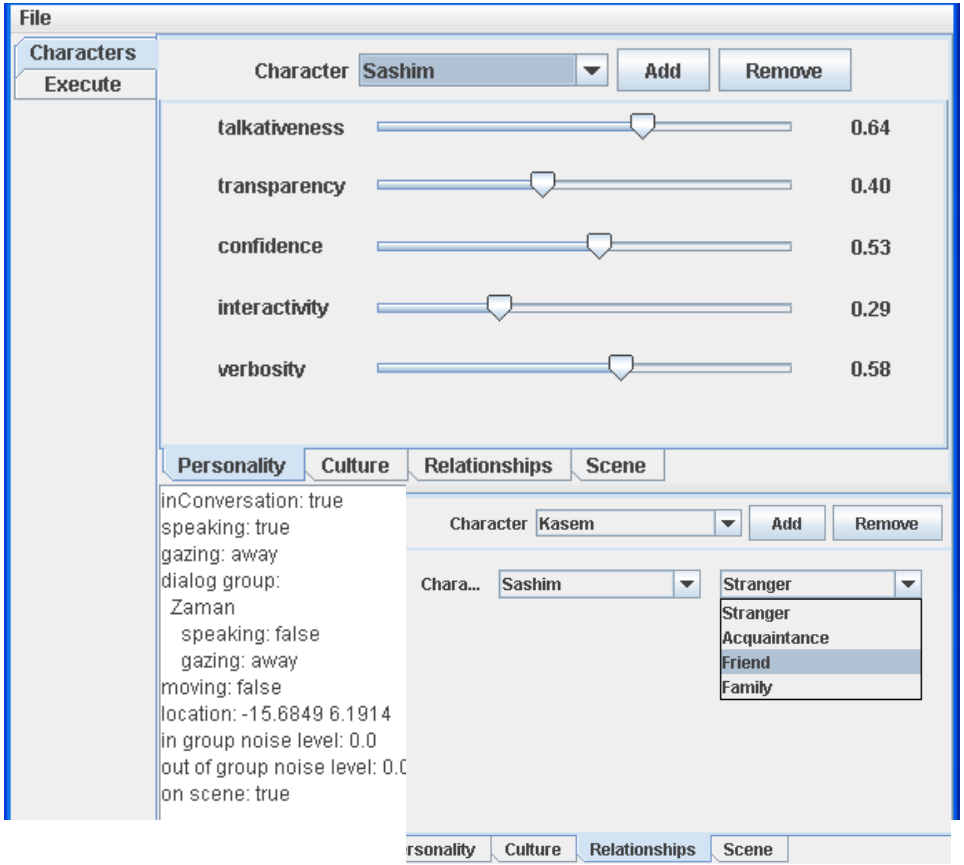
- Study literature for basis of parameterizable models of phenomena
- Study literature to find initial culture-specific settings
- Experiment/corpus study to validate/fine-tune model
- Use model to animate culture-specific agents
- Validate by perception studies from members of cultures

# ICT&UTEP Work on Culture-specific behaviors: Phase I: Group Conversational Behaviors (Jan et al IVA 2007)

- **Phenomena**
  - Proxemics
  - Silence/overlap at turn boundary
  - Gaze in conversational roles
- **Cultures**
  - Levantine Arab
  - Mexican
  - Anglo-American

# Believable group conversation simulation

## Character information state and personality model




## Group Conversation: A new character joins the conversation




USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

# Proxemics model

- **Social force model for positioning and movement**
  - $F_{speaker}$ : attractive force towards speaker
  - $F_{noise}$ : repelling force from outside noise
  - $F_{proximity}$ : repelling force from characters that are too close
  - $F_{convex}$ : force towards convex hull of all conversation participants
  - Sum forces to calculate positional goal
- **$F_{proximity}$ takes parameters for sizes of zones relevant for social relationship (Intimate, personal, social, public, Hall 68)**

| Culture | Social Zone |
|---|---|
| Anglo-American | 1.2m – 2.7m |
| Mexican | 1.0m – 2.0m |
| Arab | 0.7m – 1.5m |

# Gaze & Silence

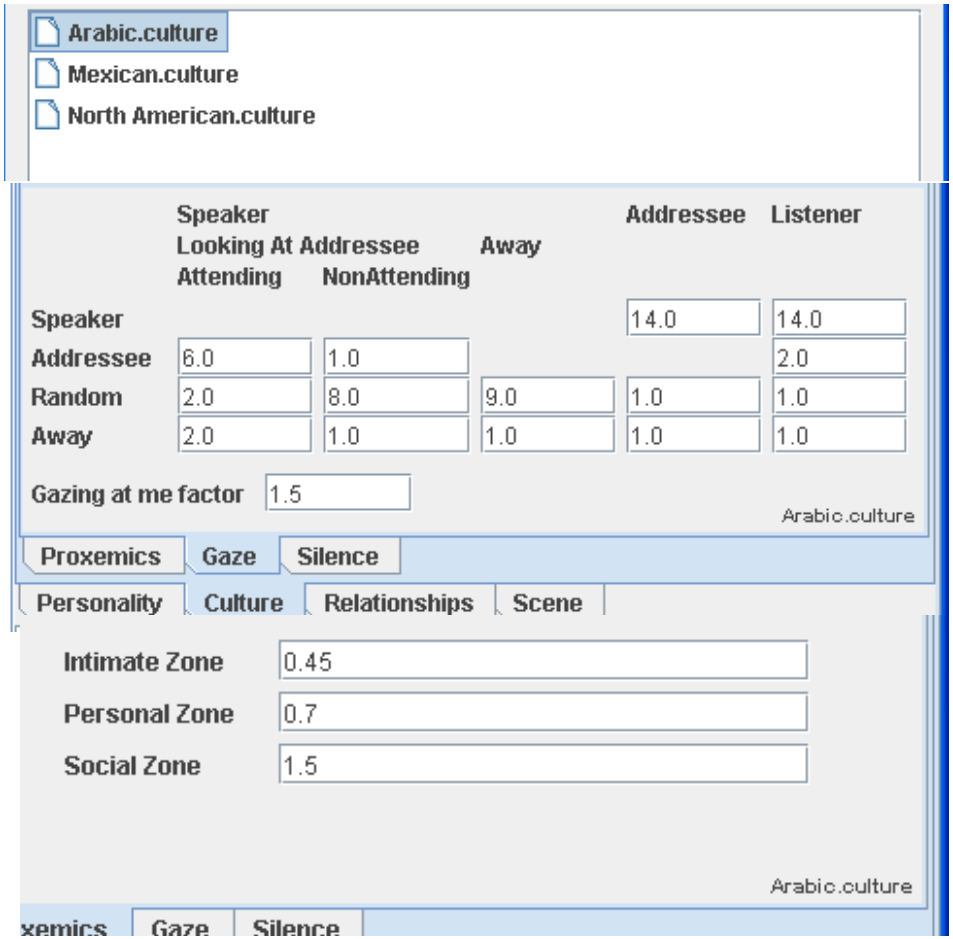
- **Gaze**
  - Likelyhood of looking at other participant depending on role
  - Gazer role: speaker, addressee, other listener
  - Gazee role: Speaker, addressee, other, none
  - Other factors: where is speaker looking previously, where is addressee looking
- **Turn-transitional silence and overlap**
  - At turn-transition: Time between speech end and speech start of new speaker (negative in case of overlap)
  - Gaussian distribution based on cultural parameters
    - Mean offset between speech end and speech start
    - Variation of offset

# Focus: Culture-specific group conversation

## Arab cultural parameters


## American cultural parameters


## Culture-specific parameters






USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

# Initial Results

- **Subjects evaluate realism in 6 movies, 2 minutes each**
  - 20 Anglo-American subjects
  - 20 Mexican subjects
  - 20 Arab subjects
- **Answer questions based on the way people talked with each other in the culture in which they grew up**
  - Seven-point Likert scale (not realistic – very realistic)

- **There are differences in evaluation of proxemics**
  - Arab subjects found Arab proxemics and animation realistic
  - Mexican and Anglo-American subjects found no significant cultural differences according to proxemics and overall animation
- **The t-tests show no significant difference in cross-cultural evaluation of gaze and turn-taking**

# Culturally-affected Behavior (Solomon et al 2008)

Figure 4 Farid and Fritz



USC

ICT
INSTITUTE FOR CREATIVE TECHNOLOGIES

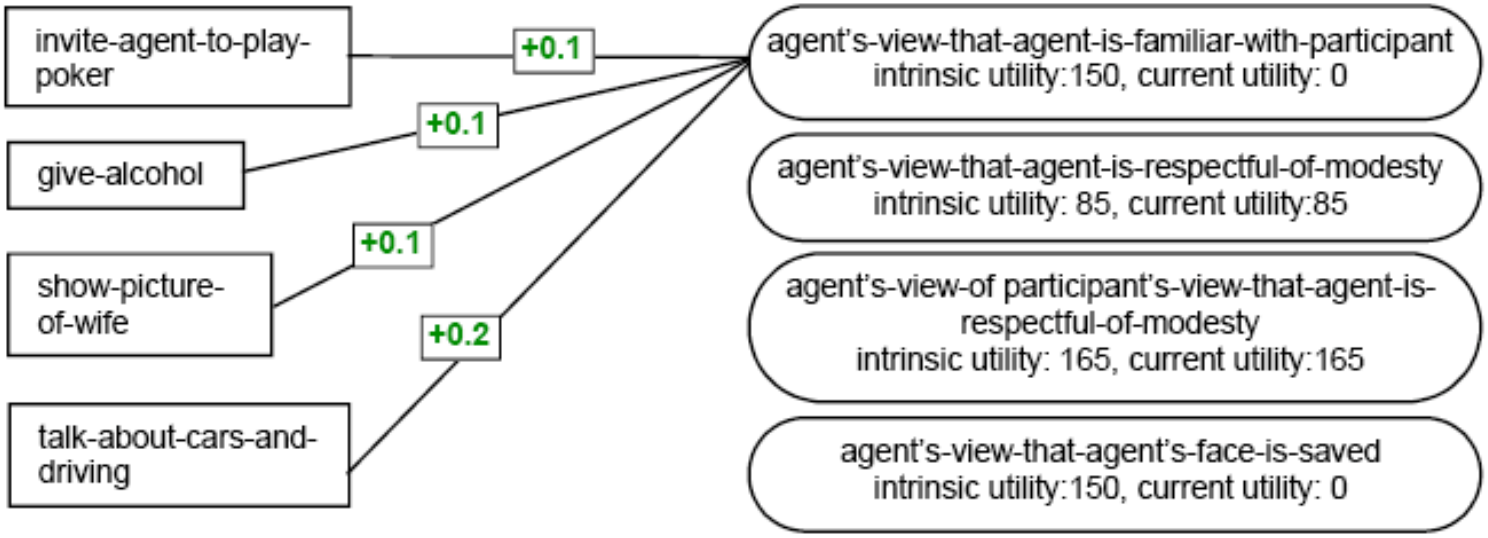



**Figure 3 Sample of Socio-Cultural Network for German Culture**

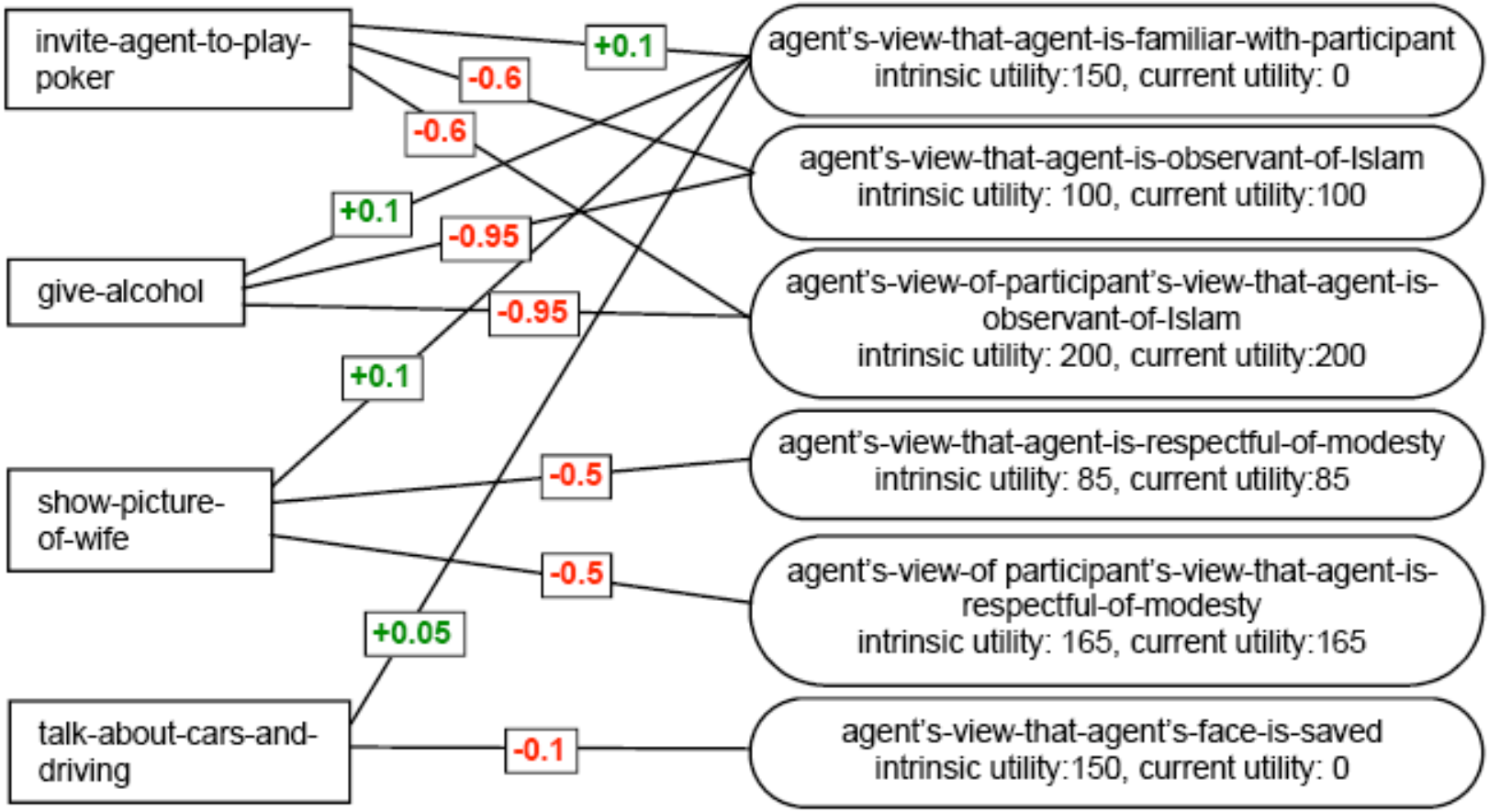



**Figure 2 Sample of Socio-Cultural Network for Iraqi Sunni Culture**

# Outline for Today

- Grounding
  - Definitions
  - Models:
    - Clark & Schaeffer
    - Traum 94
    - Paek & Horvitz
    - Roque & Traum 08
- Example Information State Grounding System
  - EDIS
- Culture-specific dialogue agents
  - Culture
  - Culture-specific dialogue differences
  - Computational models of culture for agents
  - Examples

# Outline for Course

- Monday: Introduction, Architecture of Dialogue Systems, Example Systems
- Tuesday: Simple structures: S-R, IR, finite State
- Wednesday:  Frame-based and Information State
- Yesterday : Plan-based and Logic Based
- Today: Advanced Topics: Grounding, Culture