

Language Technology II: Language-Based Interaction

Dialogue Phenomena

Ivana Kruijff-Korbayová
korbay@coli.uni-sb.de
www.coli.uni-saarland.de/courses/late2/



6/7/05
Dialogue Phenomena

Language Technology II: Language-Based Interaction
Manfred Pinkal & Ivana Kruijff-Korbayová

1

Introduction

- A dialogue system engages as a participant/agent in interaction with a human
- So, it needs to have a model of what such interaction(s) looks like
 - What needs to be modeled?
- Easy and pleasant interaction is an essential design aspect
 - What characterizes easy and pleasant interaction?



6/7/05
Dialogue Phenomena

Language Technology II: Language-Based Interaction
Manfred Pinkal & Ivana Kruijff-Korbayová

2

Introduction

- How do we know what conversations look like?
 - Study human-human conversations
 - Ultimate benchmark for "naturalness"
 - But, dialogue systems have specific requirements
 - Study human-computer conversations (data collected with actual systems)
 - Realistic, but confined to implemented functionality
 - Study simulated human-computer conversations, (data collected in Wizard-of-Oz studies, where a human simulates the system, given an algorithm)



6/7/05
Dialogue Phenomena

Language Technology II: Language-Based Interaction
Manfred Pinkal & Ivana Kruijff-Korbayová

3

Characteristics of Conversation

- Human-human conversation shows specific characteristics
- Although humans simplify their language use in human-computer interaction, the essential characteristics remain



6/7/05
Dialogue Phenomena

Language Technology II: Language-Based Interaction
Manfred Pinkal & Ivana Kruijff-Korbayová

4

Characteristics of Conversation

- Turn taking: participants take turns in contributing
- Structure
 - Global: opening, body, closing
 - Local: sequences of moves/acts, e.g., question-answer
- Importance of context
 - Dialogue economy: anaphora (pronouns, etc.), ellipsis, fragments
 - Indirectness, e.g., asking a question to make a request
- Grounding: achieving mutual understanding, i.e., common ground
- Error recovery (identifying and resolving misunderstandings)
- Mixed initiative (either participant can be in control)
- Collaboration



Lecture Outline

- Turn-taking
- Initiative and Collaboration
- Global and local structure
- Dialogue economy
- Dialogue acts
- Grounding



Turn Taking

- Dialogue participants take turns (like in a game):
A, B, A, B
- Dialogue turn = a continuous "contribution" to the dialogue from one speaker
- Though it is generally not obvious when a turn in natural dialog is finished, turn-taking appears fluid in normal conversation:
 - Minimal pauses between speakers (few hundred ms)
 - Less than 5% speech overlap
- How does it work?



Turn Taking



Turn Taking Rules

- Conversational analysis (Sacks et al. 1974)
- When does turn-taking occur:
 - Transition-relevance places ---places where the dialog/utterance structure allows speaker shift to occur (typically at utterance boundaries, but also smaller units)
 - TRP signals include syntax (phrase boundaries), intonation, gaze, gesture; Also cultural conventions apply
- Who speaks next
 - At each TRP (current speaker A):
 1. If A selected B as next speaker, B should speak
 2. If A did not select the next speaker, then anyone may take a turn
 3. If no-one else takes a turn, then A may (continue)
 - To get a turn if not selected, a speaker must "jump in" at a TRP



Turn Taking Rules

- When do we get pauses or lapses?
- When do we get overlaps?



Turn Taking Models in Human-Computer Dialog

- Rigid:
 - System speaks until it completes its turn, then waits
 - How long to wait? (Is the user still there? Did s/he hear?)
 - System lets User to finish turn, then starts
 - How to determine end of user's turn? (Is the user finished?)
- Flexible, with barge-in:
 - User barge-in: system speaks until it completes its turn or until it detects an interruption
 - Open-mic: system listening all-the-time
 - Problems: talk vs. noise; system's own talk is "noise"
 - Push-to-talk: user pushes button to take a turn
 - System barge-in: When appropriate at all? When is a TRP?



Initiative & Collaboration



Initiative

- Who is in control of the dialogue progression?
- Being the one who's talking does not necessarily mean being in control, e.g., just answering a question
- Dialogue initiative vs. task initiative
- Basically, two models:
 - Fixed initiative model (one participant in control)
 - System-initiative: can drive dialogue as wanted by prompting user, but may be unnatural and inconvenient for user
 - User initiative: can do what wants when wants (if knows what to do), but difficult for system, because it doesn't know what is coming
 - Mixed initiative model (either participant can assume initiative, depending on knowledge, skills, situation, etc.)
 - Typical in human-human conversation
 - How to decide whether to take initiative?



Collaboration

- Conversation (and communication in general) is a joint activity
 - It has a purpose (agreed on by the participants)
 - It involves collaboration/cooperation
- Collaborating (being cooperative): helping each other to accomplish goals by, e.g.,
 - Cooperative interpretation beyond literal meaning (inference)
 - Cooperative answering
 - Complying with requests or directives when possible
 - Providing more information than requested (when it is relevant or useful), also correcting false presuppositions or misconceptions
 - Intensional answers and generalizations
 - Taking initiative when this helps to accomplish the joint activity



Collaboration & Inference

- Discourse participants "read between lines"
 - Conversational implicatures due to Gricean maxims
 - Informativity principle: try strongest interpretation
- and cooperatively accommodate implicit assumptions, if possible
 - Accommodation of presuppositions: adjusting common ground unless conflicting information
 - Resolving reference to entities in common ground
- If not possible, repair problems, if relevant



Maxims of Conversation

- Cooperative Principle
 - Maxim of quality
 - Maxim of quantity
 - Maxim of relevance
 - Maxim of manner
- Conversational implicatures arise based on
 - Assumed adherence to maxims
 - Blatant violation of maxims



Conversation Structure



Conversation Structure

- Global: the overall structure of an entire conversation
- Local: relations between pairs or sequences of turns

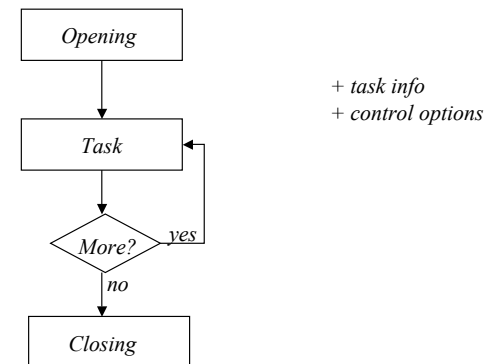


Global Structure

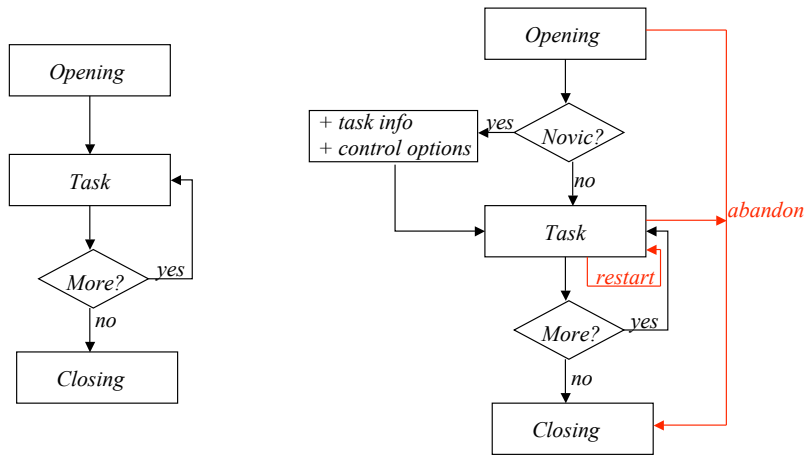
- Generic structure of a conversation:
 - Opening: "initialization" (establish contact, greetings, pleasantries)
 - Body: exchange about the subject matter(s) (accomplishing task(s), discussing topic(s)), sometimes a task is ended by a summary
 - Closing: winding down, farewell, breaking contact
- Conventions apply in all sections



Global Structure



Global Structure



Local Structure

- Adjacency pairs (Schegloff 1968):
 - Adjacent turns
 - Produced by different speakers
 - Ordered: First[^]Second
 - Typed: particular First requires a particular Second
 - Greet-greet, ask-answer, request-grant, offer-accept, compliment-downplay, etc. ⇒ preferences, expectations
- Insertion sequences: APs can be embedded
 - E.g., "sub-dialogue", misapprehension-correction, clarification

Local Structure: Preferences

- Significant silence (option 1 at a TRP)
 - If A selects B to speak next, but B doesn't
 - Then (assuming B has heard & understood) B's silence can be interpreted as a hesitation to give a dispreferred Second, e.g., B does not know the answer to A's question, B's response to A's offer or request is negative, etc.
- Other cases (silence at options 2 or 3 at a TRP) are just insignificant delays (pauses or lapses)

Local Structure: Insertions

- "Sub-dialogue":
 - A: Where are you going?
 - B: Why do you want to know?
 - A: I thought I'd come with you.
 - B: I'm going to the supermarket.
 - Clarification:
 - A: I'd like three sausages.
 - B: Which ones? Merquez or Lyoner?
 - A: Merquez.
 - B: Here you go.
 - Misapprehension-Correction:
 - A: I'd like three sausages.
 - B: Three pairs.
 - A: No, three single pieces.
 - B: OK.
- A: When is the next train from SB to Hamburg?
 B: The next train to Homburg Hauptbahnhof is at 1 p.m.
 A: Hamburg, not Homburg.
 B: Ah, Hamburg?
 A: Yes.
 B: OK, the next connection to Hamburg is at 3 p.m.

Dialogue Structure and Coherence

- Grosz and Sidner (1985)
 - Linguistic structure: discourse segments signaled by cues, e.g., discourse markers, prosody, etc.
 - Intentional structure: discourse segment purposes and relations between them (dominance, satisfaction-precedence); subdialogues vs. true interruptions
 - Attentional structure: entities in focus spaces corresponding to discourse segments; antecedents for anaphoric links; stack-model of focusing



Dialogue Economy



Dialogue Economy

- For reasons of economy, parts of structure are often "abbreviated" or omitted
⇒ anaphoric reference, ellipsis and fragments
- The missing structure can normally be recovered from the previous utterances and from the context
- Keeping track of the context is essential to coherent dialogue
- Without modeling these phenomena, dialogue can appear unnatural or even go wrong



Dialogue Economy

U: Do any samples contain bismuth and ruthenium?

S: Yes.

U: Give me their overall analyses.

U: Do any samples contain bismuth and ruthenium?

S: No.

U: Then what do they contain?

A: What time is Twelfth Night playing tonight?

B: It starts at 8:10 p.m.

A: And Hamlet?

G: where are you in relation to the top of the page just now?

F: about four inches

G: four inches?

F: yeah



Dialogue Acts (Conversation Acts) (Dialogue Moves)



Speech Acts

- Dialogue acts evolved from speech acts
- Speech act theory: do things with words (Austin, Searle)
- Utterances are acts that change context
 - Locutionary act: the act of uttering the words with their semantic content
 - Illocutionary act: the communicative act the speaker performs in saying the words --> speech acts
 - Perlocutionary act: the act that occurs as a result of the utterance (e.g., making someone laugh, scared...)



Speech Acts

Assertive	S commits to sth being the case	Comment, suggest, swear, boast, conclude
Directive	S attempts to get H do sth	Ask, order, request, beg, invite, advise
Commissive	S commits to future course of action	Promise, plan, vow, bet, oppose
Expressive	S expresses psychological state	Thank, apologize, welcome, deplore
Declarations	S changes world	Resign, name, fire



Dialogue Acts/Moves

- Generalization of speech acts to conversational functions of utterances at various levels
- Various taxonomies, typically tuned for a specific task or domain
- Attempts at reusable schemes:
 - Conversation acts (Traum and Hinkelman 1992, Traum 1994)
 - DAMSL (1997)
 - DATE (2001)



DAMSL

- DAMSL: hierarchical general DA classification scheme for task-oriented dialogue
 - Forward looking function (like speech act)
 - Backward-looking function: relationship to previous utterance(s) by other speaker (including grounding)
 - Information level
 - Task: doing the task
 - Task management: talking about the task
 - Communication management: managing communication
 - Other
 - Communicative status: intelligibility, interpretability, self-talk...



DAMSL: Forward Looking Functions

- Statement a claim
- Information request a question
 - Check a question confirming info
- Influence-on-addressee (= Searle's directives)
 - Open-option a weak suggestion or list of options
 - Action-directive command or instruction
- Influence-on-speaker (= Searle's commissives)
 - Offer offer to do something (subject to confirmation)
 - Commit commitment to do something
- Conventional
 - Opening greetings
 - Closing farewell
 - Thanking thanking and responding to thanks



Backward Looking Functions

- Agreement (speaker's attitude toward an action, plan, object, etc.)
 - Accept
 - Accept part
 - Maybe
 - Reject
 - Reject part
 - Hold
- Answer (answer to question)
- Understanding (whether speaker understood previous turn)
 - Signal-non-understanding
 - Signal-understanding
 - Acknowledgement (demonstrated by a continuer or assessment)
 - Repeat-paraphrase (demonstrated by a repetition or rephrase)
 - Completion (demonstrated by collaborative completion)



Dialogue Games/Sequences

- Some sequences of dialogue acts occur regularly, are even conventionalized; cf. adjacency pairs
 - Greeting-greeting
 - Question-answer
 - Compliment-downplayer
 - Accusation-denial
 - Offer-acceptance
 - Request-grant
 - ...
- Obligation to respond
- Preferred responses



Dialogue Act Recognition

- How do we decide what DA a user input is, e.g., statement vs. info-request
- At first glance, this looks simple
 - Different syntax:
 - Yes-no-questions have subj-verb inversion
 - Statements have declarative syntax
 - Commands have imperative syntax
- However, the mapping between surface form and illocutionary act is not one-to-one



Dialogue Act Recognition

- For example, what looks like a yes/no question
Can you give me a list of the flights from A to B
Can be a polite form of directive or request
Please give me a list of flights from A to B
- What looks like a statement
And you said you wanted to travel next week
Can actually be a question, used to verify sth.
(but, intonation!)



Dialogue Act Recognition

- Another example of "indirectness":
 - A: That's the telephone.
 - B: I'm in the bath.
 - A: OK.
- Can be paraphrased as follows:
 - A requests B to perform action (answer phone)
 - B states reason why he cannot comply (in bath)
 - A undertakes to perform action (answer phone)



Dialogue Act Recognition

- Idiom-based model:
 - Literal meaning (direct speech act)
 - Idiomatic meaning (indirect speech act)the grammar would list idiomatic meanings for each construction, e.g., Can you X? would have request as one possible meaning
- Inferential model: indirect speech acts arrived at by inference



Automatic DA Recognition

- This is one of the tasks of the dialogue management module (see next lecture)
- Plan-based interpretation
 - Essentially the inference model, differences lie in amount and depth of actual reasoning
 - Symbolic
 - Requires hand-coding and domain-knowledge
- Cue-based recognition
 - Essentially derived from the idiom model
 - Using a combination of utterance features and context features (supervised machine learning methods)
 - Requires hand-annotated data



Grounding



Grounding in Joint Activities

- Conversation (and communication in general) is a joint activity
 - Involves collaboration/cooperation of agents
 - Requires coordination of actions among agents
 - Requires common ground
- CG = sum of mutual knowledge, beliefs, suppositions
- Grounding = process of augmenting CG (by engaging in conversation)



Common Ground

- CG basis:
 - Communal CG: based on membership in cultural communities; includes facts, beliefs, and assumptions about objects, norms of behavior, conventions, procedures, skills, and even ineffable experiences.
 - Personal CG: based on joint personal experiences
- CG status is not all-or-nothing:
 - Graded evidence
 - Feeling of knowing, false consensus



Common Ground

- The Byzantine generals problem:
 - actual mutual knowledge cannot be achieved in situations in which communication is fallible.
- Individual agents act on their individual beliefs or assumptions about what their CG is
- Stalnaker (speaker presupposition)
 - Discrepancies may lead to failures in communication
 - A context is close enough to being nondefective if the divergences do not affect the conversation issues



Grounding

- (Clark 1996)
- Principle of (joint) closure: Agents performing a (joint) action require (CG) evidence, sufficient for current purposes, that they have succeeded in performing it
- The optimal evidence isn't usually the strongest, most economical and most timely evidence possible, for that may be too costly.



Grounding Levels

- Grounding means establishing CG well enough for current purposes, at all levels of interpretation:

Intention	S proposes project w	H considers project w
Proposition	S signals that p	H recognizes that p
Signal	S presents signal s	H identifies signal s
Channel	S executes behavior t	H attends to behavior t

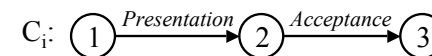


- Upward completion, downward evidence



Grounding

- Contribution (to conversation):
 - Presentation phase: A presents a signal for B to understand. He assumes that, if B gives evidence e or stronger, he can believe that B understands what he means by it.
 - Acceptance phase: B accepts A's signal by giving evidence e' that she believes she understands what A means by it. She assumes that, once A registers e' , he too will believe she understands.



- By the principle of joint closure, contributors require positive evidence that their partners understood. (vs. the assumption of understanding unless negative evidence to the contrary is presented, e.g. (Grosz&Sidner 1986))



Grounding Feedback (Acceptance)

1. Continued attention
2. Relevant next contribution (presupposing understanding, e.g., answer to question, doing action)
3. Acknowledgement: asserting understanding with
 1. A backgrounded acknowledgement token (= continuer / backchannel), e.g., "Yeah", "mmm-mm", "I see", "uh-huh", or nodding;
 2. An assessment, e.g., that's great
 3. Unison completion
4. Demonstration (by paraphrasing, reformulating or cooperatively completing)
5. Display (verbatim repetition)



Expanded Contributions

- Expanded acceptance phase:
 - often when B has trouble understanding A's presentation \Rightarrow grounding subdialog, e.g., request for verification, clarification, repair
- Expanded presentation phase ("installments"):
 - often when A anticipates B may have trouble understanding (or when A unsure) \Rightarrow dividing up and (possibly) requesting feedback through grounding subdialog, e.g., request for confirmation



Grounding Problems

A: Can I speak to Jim Johnston please?
B: Jim who?
A: Jim Johnston.
B: And what about him?
A: Can I speak to him please?
B: Jim Johnston Senior?
A: Yes.
A: No. Junior.
A: What, are there two?
B: Junior and senior.
A: Eh, I'm not sure.
B: ~~Oh, yes, I am connecting you to Jim Johnston Junior.~~
B: ~~Oh, yes, I am connecting you to Jim Johnston Junior.~~



Grounding Problems

B: I am connecting you.
B: **Jim Johnston.** **Yeah.** **Sure.** I am connecting you.
B: Oh, yes. I am connecting you.
B: OK, I am connecting you to Jim Johnston Junior.



Grounding Problems

- Grounding problems are due to
 - Lack of perception or understanding
 - Ambiguity
 - Conflicts (cannot link to CG)
 - Misunderstanding (discrepancies in CG)
- Clarification and repair strategies, e.g., ask for repetition, rephrase, clarify



Expanded Contributions

- Packaging of Installments:
What is the optimal size of a contribution?
 - The smaller the chunks with grounding feedback, the more certainty, but the communication takes longer
 - The larger the chunks, the more danger of snowball effect of a misunderstanding at some point
 - Working memory constraints
- ⇒ Variable size, depending on skills and purposes.



Verification Strategies in Systems

- Immediate explicit feedback (and verification request)
 - S: Where do you want to go?
 - U: Hamburg.
 - S: Traveling to Hamburg. (OK?)
 - U: Yes.
 - S: When do you want to go?
- Immediate "implicit" feedback by incorporating material to be grounded in the next system turn (see if user accepts or protests)
 - S: Where do you want to go?
 - U: Hamburg.
 - S: And when do you want to go to Hamburg?
- Delayed explicit feedback by summarizing at task end
 - ...
 - S: So. Traveling from Saarbrücken to Hamburg on Monday June 6
 - ...



Conclusions

- Characteristics of human-human dialogue that also (should) hold for human-computer dialogue:
 - Turn-taking
 - Initiative and Collaboration
 - Global and local structure
 - Dialogue economy
 - Dialogue acts and indirectness
 - Grounding
- but they present challenges for modeling



Reading

- D. Jurafsky and J. H. Martin. Speech and Language Processing. Chapter 19. Prentice Hall. 2000.
- H. Clark. Using Language. Chapters 4 and 8. Cambridge University Press. 1996.
- DAMSL annotation manual
<http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/>



Exercise

- Purpose of dial-a-dialogue assignment:
 - See for yourselves how and to what extent the presented dialogue phenomena are handled in a sample system
 - What would you want the/a system to be able to do
 - Think what it takes to achieve that behavior
- See course website:
<http://www.coli.uni-saarland.de/courses/late2/>
--> Schedule --> Exercise June 9 2005

