

Hauptseminar Dialogsysteme und Voice-XML SS 2011

Christoph Draxler, Klaus U. Schulz

Sitzung 2: Dialogbegriff

Inhalte und Ziele

- Definition - was ist ein Dialog?
- Natur und Struktur von menschlichen Dialogen besser verstehen.
- Voraussetzungen zur Modellierung von Dialogen auf der Maschine beleuchten.

Warum sind menschliche Dialoge sehr komplex (Syntax, Semantik, Pragmatik, Prosodie,..), warum wirken Computerdialoge künstlich?

Hauptliteratur

Mc. Tear, Spoken Dialogue Technology, Ch. 3

“Definition”

“Definition”

Menschlicher Dialog

Kommunikation zweier oder mehrerer Menschen (Akteure) mittels gesprochener Sprache. Summe der Äußerungen bildet Situations-Einheit und/oder Sinneinheit.

Beispiele für "Sinneinheiten" (balanciert, gerichtet):

- Austausch zu Ansichten oder Informationen zu einem Thema
 - Gemeinsames Finden einer Lösung/eines guten Vorgehens
Problemlösung, Konfliktlösung
 - Austeilen und Entgegennahme (inkl. Nachfragen) einer Bitte, eines Vorschlags, Auftrags oder einer Instruktion
-

“Definition”

Manchmal Unterscheidung

Dialog versus „Konversation“

- Konversation informeller, dient der Pflege sozialer Beziehungen.
- Dialog im eingeschränkten Sinn eher an konkreten Zielen oder Aufgaben orientiert.

Dialog mit Rechner (nachfolgend)

Jede Art der menschlichen Kommunikation mit dem Rechner auf der Basis gesprochener Sprache.

“Definition”

Eigener Dialogbegriff in weiteren Disziplinen
(Psychologie, Soziologie,
Kulturwissenschaft, Religion,)

2. Charakteristika von Dialogen

Charakteristika von Dialogen

Überblick

- Dialog als Diskurs
(ling. Eigenschaften von Äußerungsfolgen, Syntax, Semantik)
 - Dialog als zweckgebundene Aktivität (Pragmatik)
 - Dialog als kollaborative Aktivität
("Spielregeln" für harmonischen Dialog)
 - Form von Äußerungen in Dialogen
(Unterschiede gesprochene, geschriebene Sprache)
-

2.1 Dialog als Diskurs

2.1 Dialog als Diskurs

Fokus syntaktische und semantische Merkmale von Dialogen

Notwendigkeit, zurückliegende Sätze, Äußerungen und Dialogteile in die Interpretation eines neuen Satzes miteinzubeziehen.

Welche Entitäten (und komplexere Konzepte) wurden eingeführt?

Wie kann man mit Pronomen und anderen anaphorischen Bezügen auf diese zugreifen? Wie Bezüge korrekt auflösen?

Dialog als Diskurs- Anaphernaauflösung

A: John won some money

B: What is **he** going to do with **it**?

Morphologie: he= John, it=money

A: John won some money in the lottery

B: What is he going to do with **it**?

?:it=money (kein Bezug auf letztes morphologisch mögliches Glied!)

?:Money ist "Fokus": it=money

(oder: "etwas mit Geld machen" versus „etwas mit einer Lotterie machen“)

A: Jim caught up with Bill outside the pub

B: Did **he** give **him** the tickets?

Most recent: he= Bill

Subject preference over object: he=Jim

Beide Strategien: he=him (aber Weltwissen zu „give“: he nicht him)

„Parallelreihenfolgen“: he=Jim, him=Bill

Dialog als Diskurs - Anaphernauflösung

Verwendung von Weltwissen statt Rückgriff auf eingeführte Entitäten

A: Did you see Bill in the pub last night?

B: No, **the barman** said he left early

Definite Beschreibungen (the barman) eigentlich in der Regel vorher eingeführt. Im Normalfall würde man in der Dialoghistorie nach einem Barman suchen.

Dialog als Diskurs

Komplexere anaphorische Bezüge: Rückgriff auf was??

A: Sollen wir am Do die Seminararbeit vorbereiten!

B: Ok, wir können ja **danach** noch ein Bier trinken!

A: Auf der Demo waren nur ein paar Fans von Guttenberg.

B: Meiner Ansicht nach spinnen **die**!

A: Autsch, ich habe mir den Fuß verdreht!

B: Geht **es** wieder?

Dialog als Diskurs

Komplexere anaphorische Bezüge: Rückgriff auf was??

A: Wenn er sich nicht entschuldigt, zeige ich ihn an!

B: **Das** ist keine gute Idee!

Speichern von früher eingeführten "Topics,,

A: Am Sa gehe ich in das Konzert von ACDC

[B, A: ...Subunterhaltung über die Musik von ACDC...]

B: Wann beginnt **das Ganze**? Wer spielt denn **als Vorgruppe**?

Dialog als Diskurs

Einfache Dialogsysteme *steuern* den Benutzer bei der Auswahl von Topics und ggfs. beim Wechseln von Topics.

Für das System transparent, für den Benutzer teils umständlich.

Experimentelle Systeme können freie Topikänderung durch den Benutzer erlauben ("ach übrigens, wieviel Uhr ist es denn?").

Buchhaltung zu nicht "beendeten" Topics notwendig.

2.2 Dialog als zweckgebundene Aktivität

2.2 Dialog als zweckgebundene Aktivität

Dialog oft mit Zweck verbunden (Bitte, Auftrag, Transaktion,...).

- "Echte" Bedeutung der verwendeten Äußerungen hängt vom Situationskontext ab.
 - Auswahl einer adäquaten Sprache hängt vom Kontext ab.
 - Oft werden Bedeutungsanteile wahrgenommen, die in den reinen Wörtern nicht enthalten sind.
-

Dialog als zweckgebundene Aktivität

Bedeutung hängt vom Kontext ab

„Hier ist es warm“

Neutral -- Komm herein! -- Gehen wir raus! -- Unternimm etwas,
damit es hier kälter wird

Dialog als zweckgebundene Aktivität

Adäquatheit der Sprache (Situation, Teilnehmer)

Wunsch: mache die Türe zu

1. Mach die Türe zu!
 2. Würde es Ihnen etwas ausmachen, die Türe zuzumachen?
 3. Ich würde es begrüßen, wenn du die Türe schließt.
 4. Ist die Tür offen?
 5. Hier zieht es!
-

Dialog als zweckgebundene Aktivität

Sprechakte (Speech Acts - Austin und Searle, 60er)

- Dialoge nicht durch traditionelle Wahrheitswertbedingungen (wie in üblicher Semantik) erfassbar.
 - Dialog hat FUNKTION (Wittgenstein: "Worte sind Taten").
Es geht um die Aktionen, Absichten und Handlungen der Sprecher und Hörer (Information, Bitte, Kommando, Einschüchterung,...)
-

Dialog als zweckgebundene Aktivität

Charakterisierung von Sprechakten und ihrer Voraussetzung am Beispiel **Kommando** (Searle 1969):

- Äußerung betrifft zukünftige Aktion, die der Hörer ausführen soll.
- Er kann diese ausführen, oder zumindest glaubt der Sprecher, dass der Hörer dies kann.
- Es ist nicht offenkundig, dass bei einem normalen Verlauf der Hörer die Aktion ohnehin ausführt ("atme").
- Sprecher möchte, dass Hörer die Aktion ausführt.

Wenn Akteure Sprechakte ausführen sollen, müssen sie in der Lage sein, *Glaubenszustände, Wünsche und Absichten der Dialogpartner zu erfassen.*

Dialog als zweckgebundene Aktivität

Implikaturen

Bedeutungsanteil einer Äußerung, der allenfalls angedeutet ist, nicht explizit ausgedrückt ist.

Hörer braucht (soziale, Dialog-) Kompetenz, um zu verstehen.

Konversationelle Implikaturen basieren auf Kooperationsprinzip (Paul Grice), Kommunikation folgt rationalen Regeln,

Dialog als zweckgebundene Aktivität

1. Kooperationsprinzip

- * Gestalte deinen Gesprächsbeitrag so, dass er dem anerkannten Zweck dient, den du gerade zusammen mit deinen Kommunikationspartnern verfolgst.

2. Maxime der Quantität

- * Mache deinen Gesprächsbeitrag mindestens so informativ, wie es für den anerkannten Zweck des Gesprächs nötig ist.
- * Mache deinen Beitrag nicht informativer, als es für den anerkannten Zweck des Gesprächs nötig ist.

3. Maxime der Qualität

- * Versuche einen Gesprächsbeitrag zu liefern, der wahr ist.
 - * Sage nichts, wovon du glaubst, dass es falsch ist.
 - * Sage nichts, wofür du keine hinreichenden Anhaltspunkte hast.
-

Dialog als zweckgebundene Aktivität

4. Maxime der Relevanz/Relation

- * Sage nichts, was nicht zum Thema gehört, wechsle das Thema nicht einfach so plötzlich.

5. Maxime der Modalität

- * Vermeide Unklarheit.
 - * Vermeide Mehrdeutigkeit.
 - * Vermeide unnötige Weitschweifigkeit.
 - * Vermeide Ungeordnetheit.
-

Dialog als zweckgebundene Aktivität

Eine konversationelle Implikatur kann z.B. dadurch entstehen, dass das Gesagte nur dann einen Sinn ergibt, wenn wir eine Implikatur hinzufügen. Zum *Erfassen* von Implikaturen dienen teils o.g. Maximen.

Autofahrer: „Mein Benzin ist alle.“

Fußgänger: „Gleich um die Ecke ist eine Tankstelle.“

➤(impliziert) An der Tankstelle gibt es Benzin zu kaufen.

A: Leihst du mir etwas Geld für einen Kaffee?

B: Es ist nicht Weihnachten!

System: Wann wollen Sie von London abfahren?

User: Ich muss das Meeting um 10 erreichen.

Dialog als zweckgebundene Aktivität

Reasoning-Beispiel (Variante eines Beispiels von Searle):
„Können Sie mir eine Liste der Flüge von München nach Hamburg
geben?“

Dialog als zweckgebundene Aktivität

Reasoning-Beispiel (Variante eines Beispiels von Searle):
„Können Sie mir eine Liste der Flüge von München nach Hamburg geben?“

X hat mich gefragt, ob ich in der Lage bin, eine Flugliste anzugeben. Ich nehme an, dass X „kooperativ“ (im Grice-schen Sinn) ist und seine Äußerung ein Ziel hat. X weiß, dass ich in der Lage bin, eine solche Flugliste anzugeben, es gibt keinen Anhaltspunkt, dass X ein rein theoretisches Interesse an meinem Fähigkeiten zum Geben von Fluglisten habe. Daher steckt wohl hinter der Anfrage etwas anderes. Was kann es sein? Eine Voraussetzung für „Direktiven“ ist es, dass der Hörer die geforderte Aktion ausführen kann. Aus diesem Blickwinkel macht es Sinn, wenn X fragt, ob ich eine Flugliste angeben kann. Dazu sind wir in einer konversationellen Situation, wo es eine normale Aktivität ist, eine Flugliste anzugeben. Daher, weil es keine plausible Interpretation eines anderen Sprechakts gibt, ist die Frage wohl als Aufforderung gemeint.

Dialog als zweckgebundene Aktivität

Oberflächenform und Sprechaktform

Kannst du mir das Salz geben (?/!) (Frage - Aufforderung)

Es wäre schön, wenn du mitkämost (./!) (Feststellung - Bitte)

Der Euro steht bei 7 Dollar (./?) (Feststellung - Frage)

Und Sie sagten, Sie wollten nächste Woche fliegen (./?) (Feststellung - Nachfrage)

Wer kann das bezahlen (?/!) (Frage – Feststellung, Beschwerde)

Rolle der Prosodie!

2.3 Dialog als kollaborative Aktivität

2.3 Dialog als kollaborative Aktivität

Dialogpartner müssen zu einem geordneten Dialogablauf beitragen.

„Turn-Taking“: Übernahme der Rede-Aktivität.

Es gibt Regeln, nach denen im Gespräch abgewechselt wird.

Keine Pause zwischen beiden Sprechern im Normalverlauf.

"Staffellauf mit Übergabe des Staffelholzes".

Vielseitige "Signale", wie man Ende des eigenen Gesprächstückes bzw. auch den Wunsch, etwas zu sagen, signalisieren kann. („Wie sehen Sie das?“ „Was halten Sie davon?“)

Abwehr einer "gewaltsamen Übernahme": "Bitte lassen Sie mich kurz ausreden".

Dialog als kollaborative Aktivität

Bei Mensch-Computer Dialogen Turn-Taking meist streng markiert.

Andeuten des Endes des eigenen Beitrags des Rechners durch „Floskeln“ zu komplex und umständlich. Stattdessen „Beep!“

Bei vielen Systemen kann Mensch kein Signal ausgeben, dass er die Rede übernehmen möchte.

„Barge-in“ als Möglichkeit einiger Systeme, mit Problemen verbunden (Spracherkennung)

Mensch signalisiert durch Pause Ende seines Beitrags.

Dialog als kollaborative Aktivität

Grounding

Sicherstellen einer gemeinsamen Basis für das Verständnis.

Signalisieren des "akustischen" und/oder "inhaltlichen" Verstehens, ggfs. Nachfrage.

Wo notwendig, Voraussetzungen schaffen für gemeinsames inhaltliches Verständnis.

Beispiel: Situation, wo neues Mitglied zu Dialogpartnern hinzukommt. "Wir haben uns über xy unterhalten und waren dabei...".

2.4 Äußerungen in Dialogen

2.4 Äußerungen in Dialogen

Äußerung in gesprochener Sprache versus Sätze in Texten.

Grammatikalische Komplexität echter Sätze in Äußerungen meist gering

Äußerungen haben oft Form partieller Sätzen, „elliptische“ Antworten.

A: Ist die Milch im Kühlschrank?

B: Ja.

A: Gib mal bitte!

Äußerungen in Dialogen

Zusatzphänomene ungrammatikalischer Art

Falsche Starts: "Did - did you see - have you ever seen a lion in the wild?"

Zögern, Räuspern, Füllfloskeln, Pausen, ..

Prosodie

Tonhöhenkontur ("Ok" aufsteigend: checking act, "Ok" absteigend: Akzeptanz)

Betonung (der wichtigsten Einheit)

Phrasieren von Gruppen "Call the (ticket office in Belfast)" versus

"Call the ticket office (- in Belfast),"

Nichtverbales Verhalten

Rolle von Stille

Stille als unausgesprochenes "Nein,":

"Könntest du mir vielleicht bis Mai 20,000 Euro leihen?"

3. Modellierung von Dialogen

3. Modellierung von Dialogen

Dialogstruktur

Hierarchische Struktur, Teile, Muster für Abfolgen.

Dialogprozess

Was „treibt“ Dialoge.

Wie läuft Interpretation im Kontext, Grounding etc. ab?

Was sagt man wann wie warum?

3.1 Dialogstruktur

Dialogstruktur

"Dialogue-Act"

Kleinste Einheiten (z.B. Bitte, Aufforderung)
Charakteristische Abfolgen (Bitte, Antwort)

- Können Sie mir den Zucker reichen?
- Bitteschön!
- Danke!

DAMSL (Dialogue Act Markup in Several Layers)

System zur Annotation von Dialogen

Sinnvolle Kategorien für Acts hängen stark von Funktion des Dialogs ab.

Dialogstruktur

DAMSL (Dialogue Act Markup in Several Layers)

Was wird in DAML annotiert?

- „Kommunikativer Status“: Äußerung verständlich und vollständig?
 - „Information-Level“: Charakterisierung des semantischen/ pragmatischen Inhalts: bringt er die Aufgabe vorwärts, bezieht er sich auf die Kommunikation,..
 - Forward-Looking Function: wie beeinflusst Äußerung den folgenden Dialog? Glaubt der Hörer nun bestimmte Dinge, oder hat er zugesagt, bestimmte Aktionen zu machen?
 - Backward Looking Function: Wie bezieht sich Äußerung auf zurückliegenden Dialog?
-

3.1 Dialogstruktur

DAMSL (Dialogue Act Markup in Several Layers) Annotation von Acts in Dialogstrukturen

Table 3.1. DAMSL: Forward-looking dialogue acts

Statement	A claim made by the speaker or an answer to a question.
Assert	The speaker is trying to change the beliefs of the addressee.
Reassert	The speaker thinks the claim has already been made.
Other-statement	Other categories of statement.
Influencing-addressee- future-action	Request other than to perform an action.
Open-option	Suggests an action without obligating the hearer.
Action-directive	Obligates hearer to perform the action, or to communicate a refusal or an inability to perform the action.
Info-request	Questions and requests for information.
Committing-speaker- future-action	Commits the speaker to some future action.
Offer	Conditional on the hearer's agreement.
Commit	Not conditional on the hearer's agreement, e.g., promise.
Conventional	Conventional conversational acts.
Opening	Greetings, phrases used to start an interaction.
Closing	Saying goodbye, phrases used to close a dialogue.
Explicit-performative	Performing an action by virtue of saying the utterance, e.g., "I apologise" as way of doing an apology.
Exclamation	An exclamation, e.g., "Ouch".
Other-forward-function	Any other forward-looking function.

Dialogstruktur

DAMSL (Dialogue Act Markup in Several Layers) Annotation von Acts in Dialogstrukturen

Table 3.2. DAMSL: Backward-looking acts

Agreement	The second participant's response to a previous proposal, request or claim.
Accept	Accepts all of the proposal, request or claim.
Accept-part	Accepts the proposal, request or claim in part.
Maybe	Defers a definite answer.
Reject-part	Rejects part of the proposal, request or claim.
Reject	Reject all of the proposal, request or claim.
Hold	Leaves the proposal, request or claim open, e.g., counterproposal, request for further information.
Understanding	Actions taken by the speakers to ensure they understand each other as the dialogue proceeds.
Signal-non-understanding	Explicitly indicate a problem in understanding the utterance.
Signal-understanding	Explicitly signal understanding.
Acknowledge	Indicate that the previous utterance was understood without necessarily signalling acceptance.
Repeat-rephrase	Repeat or rephrase the previous utterance to signal understanding.
Completion	Signal understanding by completing the speaker's utterance.
Correct-misspeaking	Correction of the previous speaker's utterance.
Answer	Compliance with an Info-Request.
Information-relations	How the content of the current utterance relates to the content of the previous utterance.

Dialogstruktur

Verbindungen von Acts

- Aufeinanderfolgende Paare

 - (Frage-Antwort, Gruß-Grußerwiderung, Einladung-Aannahme der Einladung)

- Aufeinanderfolgende Tripel

 - (Initiierung, Antwort, Feedback im Unterricht)

Größere Einheiten

- hängen stark von Funktion des Dialogs ab. (Topics, Transactions)

Dialogstruktur

Dialog-Grammatiken?

Kann man die mögliche Struktur von Dialogen mittels
„Dialoggrammatik“ beschreiben?#

- Satzgrammatik: Es gibt ungrammatikalische Sätze!
- ?Dialoggrammatik: Wann ist Dialog „ungrammatikalisch“??
Versuch, aus jedem Verlauf einen Sinn zu machen!

„Dialog-Grammatiken“ könnten allenfalls typische Verläufe festlegen.
Dialog ist dynamisch, wird getrieben durch Verlauf, Teilnehmer,..

Dialogstruktur

Warum Dialoge annotieren?

- Analyse der Einheiten, Vorkommen in Dialogarten, Analyse der Struktur
 - Design von Dialogsystemen, Vorlage für sinnvolle Verläufe
 - Machine Learning: Was ist natürlichste Weise, Dialog fortzusetzen?
-

3.2 Dialogprozess

3.2 Dialogprozess

Teile von Dialogen oder Abfolgen in Dialogen können strukturell gut analysiert werden, wenn man eine adäquate Betrachtungsweise der Funktion des Dialogs hat:

Dialog als Teil sozialen Verhaltens

- (Alltagsdialoge; „Guten Tag!“-“Guten Tag!“)

Dialog als Erzielen eines gemeinsamen Verständnisses (Grounding)

- („Hast du heute schon Nachrichten gehört?“)

Dialog als zielgerichtete Aktion (z.B. Aussuchen eines Weins)

- „Welche Rotweine können Sie empfehlen?“
 - Planung („Was sollen wir am Wochenende machen?“)
-

Literaturergänzung

Pragmatics and Computational Linguistics, Dan Jurafski (in Handbook of Pragmatics, im Netz unter cite-seer): Logik-basierte und probabilistische Interpretation von Sprechakten.

Was haben wir gesehen?

Sprache in Bezug auf Form, Inhalt und Rolle viel komplizierter als naive Vorstellung.

Wollte man mit Maschine „freie und natürliche“ Dialoge erreichen, müßte Rechner

- in der Lage sein, Wissens-, Glaubens-, Fähigkeitenzustände des Dialogpartners zu erfassen,
 - ein gutes Bild von typischen Dialogverläufen in Abhängigkeit vom Dialogtyp und -Zweck haben, momentane „Dialogzustände“ beurteilen können,
 - Welt- und sprachliches Wissen zur Auflösung von anaphorischen Referenzen zur Verfügung haben,
 - Sprache nach Situation passend wählen können,
 - ...
 - ...
-