T-(538|725)-MALV, Natural Language Processing
Adjacency pairs, speech acts and grounding in dialogs

Hrafn Loftsson[1] Hannes Högni Vilhjálmsson[1]

[1]School of Computer Science, Reykjavik University, Iceland

November 2010

Outline

Outline

Coordinating Dialog

Some Dialog Properties

Dialog is a kind of *discourse*...

- ...that is created by **more than one person**.
- ...which is constructed **on-the-fly**.
- ...affords **mutual understanding**.

Coordinating Dialog

Coordinated Understanding

If a computer system offers a natural language dialog as an interface, it could utilize dialog skills to ensure mutual understanding and to reduce ambiguitiy. In fact, one can view dialog as a mechanism for incremental mutual understanding.

For example, the system could:

- Confirm correct understanding.
- Ask user for missing information.
- Ask user for furher explanation.

Coordinating Dialog

Coordinated Understanding

If a computer system offers a natural language dialog as an interface, it could utilize dialog skills to ensure mutual understanding and to reduce ambiguitiy. In fact, one can view dialog as a mechanism for incremental mutual understanding.

For example, the system could:

- Confirm correct understanding.
- Ask user for missing information.
- Ask user for furher explanation.

Simple Dialog Systems

Speech Systems Without Language Processing

- Various phone services use speech recognition to retrieve information from users.
- Such systems often deal with difficult conditions (poor sound quality and ill formed sentences).
- Therefore they rely on simple but robust recognizers that only listen for pre-determined key words that fill information schemas.
- Further language processing is not performed.

Simple Dialog Systems

Automata for Dialog Systems

- When dialog systems are simple, they can be implemented as finite-state automatas (see Fig 15.1).
- A question is asked in each state, and a finite set of responses that a speech recognizer listens for, moves the dialog to a new state.
- An invalid input may require the user to repeat the response.

Simple Dialog Systems

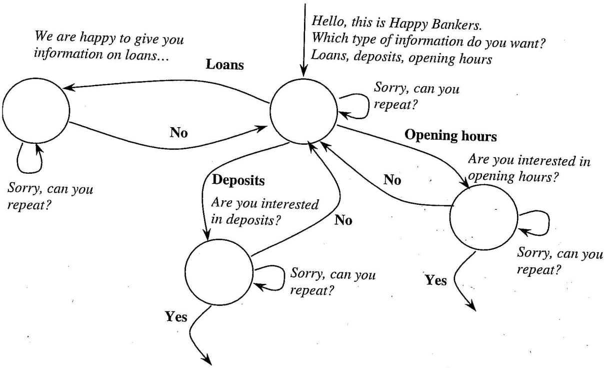

Fig. 15.1. An automaton for word-triggered dialogues.

Simple Dialog Systems

Implicit or Explicit Confirmation

The system can get confirmation about whether it understood the user correctly, either:

- **Implicitly**
  "Since you're interested in X, I recommend..."
- **Explicitly**
  "Are you interested in X?".

Simple Dialog Systems

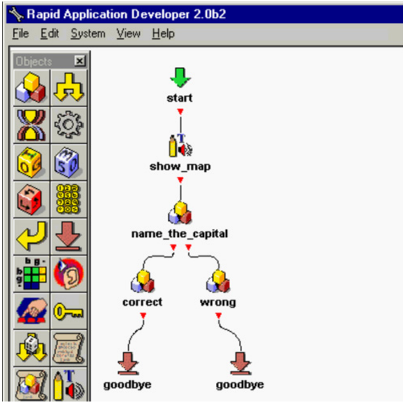

Figure: Visual tools, like the **CSLU Toolkit**, are available for quickly creating simple dialog systems

Simple Dialog Systems

Adjacency Pairs (í. grannpör)

- The dialog automata doesn't necessarily describe a very natural conversation.
- However, it serves as a good example of the pairing of sentences that happens in natural language and is an imporant aspect of more complex systems.
- Pairs of sentences, from two separate producers, where the first requires the presence of the second sentence, is called an **Adjacency Pair**

Examples

Questions/Answers, Greetings, requests/affirmations, ...

Simple Dialog Systems

Adjacency Pairs (í. grannpör)

- The dialog automata doesn't necessarily describe a very natural conversation.
- However, it serves as a good example of the pairing of sentences that happens in natural language and is an imporant aspect of more complex systems.
- Pairs of sentences, from two separate producers, where the first requires the presence of the second sentence, is called an **Adjacency Pair**

Examples

Questions/Answers, Greetings, requests/affirmations, ...

Preferred and Dispreferred Second Member

The response to a first member of an adjacency pair could be considered either preferred or dispreferred.

Table: 15.1 (in book)

First Member	Preferred Second	Dispreferred Second
Offer, Invitation	Acceptance	Refusal
Request	Compliance	Refusal
Assessment	Agreement	Disagreement
Question	Expected answer	Unexpected answer
Blame	Denial	Admission

Simple Dialog Systems

Three Sentence Intervention

Often a single pair of adjacent sentences is not enough to complete a simple exchange. Some kind of a conclusion or evaluation may be needed in a third sentence. This has been described by a **Three Sentence Intervention** model (see Table 15.2).

- Initiative Intervention ("Please pick X or Y!")
- Reaction Intervention ("X please.")
- Evaluation Intervention ("Thanks for choosing X.")

Nested Pairs and Interventions

It is possible that the second part in a pair or intervention contains a new pair or a series of interventions, i.e. they can be nested (see Table 15.3)

Table: 15.3 (in book)

Turn	Utterance	Intervention
S1	Pick one: loans or deposits	I1
U1	Deposits	R1
S2	Are you interested in deposits?	E1.I2
U2	Yes	E1.R2

Outline

Speech Act

Speech Act Theory

- It can be easy to create a dialog system where simple words or phrases move it forward.
- It is more complex for the system to truly understand the function of each full sentence.
- Language can be viewed as some kind of an act upon the receiver.
- From this act, called a **Speech Act**, the receiver can understand the producer's intent.

Speech Act

Speech Act Theory (cont...)

Lexical semantics (meaning of the words) don't fully describe the speech act. That is one reason the philosopher Austin (1962) divided an act into three distinct parts:

- **Locutionary**: The act of speaking, syntax and lexical semantics. Discourse devices.
- **Illocutionary**: The act that was intended, such as to ask or answer. Discourse function.
- **Perlocutionary**: The effects of these acts on the receiver, such as to frighten or worry.

Speech Act

From Locutionary To Illocutionary

Some types of syntactical forms are frequently associated with speech acts.

Table: 15.4 (in textbook)

Speech Act (Illocutionary)	**Syntatic Form (Locutionary)**
Assertions, statements	Affirmatives or declaratives
Orders, commands	Imperatives
Questions	Interrogatives

Speech Act

Classification of Illucutionary Acts

One can not always derive speech acts from the form or formal structure of sentences.

Example

- "Close the window!"
- "Can you close the window?"
- "It's geeting warm here.."

Speech Act

Classification of Illucutionary Acts

Therefore people tend to focus on the classification of illucutionary acts irrespective of syntactic or semantic categorization at the locutionary level. The starting point tends to be the original illocutionary classification by Searle (1969):

- Assertives: "This is a window"
- Directives: "Close the window!/Can you close the window?"
- Commissives: "I promise to close the window"
- Declaratives: "I charge you with treason"
- Expressives: "Pardon me"

Speech Act

Speech Acts as Dialog Acts

- The original illocutionary classification provides only a few high-level categories and is more concerned about the possible purpose of a sentence in isolation, rather than covering the myriad of distinct functions of the language device within a highly dynamic and evolving dialog.
- Therefore, when applied to analyzing natural dialogs, many researchers extended the original notion of speech acts to **Dialog Acts**, which include for example conversation coordination functions and allow for the possibility of more than one function being served at the same time.
- Some well known classification schemes include SUNDIAL, VERBMOBIL, DAMSL and SWBD-DAMSL.

Speech Act

Speech Act Tagging

The classification of speech acts suggests a **Speech Act Tagset** which can be used to tag sentences in a discourse in a similar fashion to tagging words within sentences. Both rule based and statistical approaches to automatic speech act tagging are being used.

Statistical Tagger

The tagger of Alexanderson (1996) is based on HMM and n-grams. One can train these taggers using one of the many existing corpora that have been annotated with speech act tags.

Speech Act

Speech Act Tagging

The classification of speech acts suggests a **Speech Act Tagset** which can be used to tag sentences in a discourse in a similar fashion to tagging words within sentences. Both rule based and statistical approaches to automatic speech act tagging are being used.

Statistical Tagger

The tagger of Alexanderson (1996) is based on HMM and n-grams. One can train these taggers using one of the many existing corpora that have been annotated with speech act tags.

Speech Act

Features Used to Identify Speech Acts

The features that are used by rules or statistical models to identify speech acts include:

- Cue words and phrases.
- Syntactic and semantic form (Locutionary)
- Adjacency or intervention pairing.
- Discourse structure, task model and plan progress.

Speech Act

Dialog Grammar

- For relatively simple kinds of conversations, it is possible to create a **Dialog Grammar** that describes all valid sequences of speech acts.
- This grammar can for example be used to parse a dialog that has been tagged with speech acts.
- This approach to full dialog interpretation is brittle since most natural dialogs don't seem to follow strict grammar rules.

Speech Act

Task Modeling

- Alternatively, a **Task Modeling** approach breaks the dialog into phases that each has its own expected outcome, usually described as a set of conditions that have to be fulfilled (e.g. certain supplied knowledge or a certain action performed).
- The phases fit together to form a sequence of steps towards completing some overall task.
- Depending on the task, the steps may be completed in any order, but often the system tries to guide the user through the "correct" order.
- While simple task modeling can be accomplished with automata, robust systems use more sophisticated plan recognition methods.

Outline

Grounding

Agent Model of User

- A system interacting with a human can build and use a so-called **Agent Model** of the user.
- It could then descibe that agent's mental state using **Modal Logic**:
  - $want(A, X)$; $cando(A, X)$; $believe(A, X)$
- Modal logic describes *personal truths* (nonuniversal)

Grounding

Belief about Belief

- The beliefs of the system K about the user A can be described like this:
  - $believe(K, want(A, X))$;
  - $believe(K, cando(A, X))$;
  - $believe(K, believe(A, X))$
- From this, the system can infer new facts about the user.
- But how can the system be sure its knowledge about the user is correct?

Grounding

Grounding

- If the system provides new knowledge **Y** verbally, is it guaranteed that the user has understood and that **believe(A, Y)**?
- The system can test this through a process of **grounding**, for example by asking:
  - "Did you get it?" eða "Do you know **Y**?"
- This is something that appears all the time in people's conversations. Many of the classic dialog patterns (e.g. some adjacency pairs) describe the grounding process.
- Whole categories of speech act tags (dialog act tags) are dedicated to grounding.