What is a discourse?

• The linguistic term for a contextually related group of sentences or utterances

• Basic discourse types
  • Monologue
  • Dialogue
  • HCI turn taking / "dialogue"

Cohesion and Coherence

• Cohesion
  • The bond that ties sentences to one another on a textual level

• Coherence
  • The application of cohesion in order to form a discourse

Reference Phenomena 1

• Indefinite noun phrases
  • an apple, some lazy people
• Definite noun phrases
  • the fastest computer
• Demonstratives
  • this, that
• One-anaphora

Reference Phenomena 2

• Inferrables
  • car ⇒ engine, door
• Discontinuous sets
  • they, them
• Generics
  • they

Referential Constraints

• Agreement
  • Number
  • Person and case
  • Gender

• Syntactic constraints
• Selectional restrictions
Coreferential Expressions

- Coreference
  - Expressions denoting the same discourse entity corefer

- Anaphors
  - Refer backwards in the discourse
  - The referent is called the antecedent

- Cataphors
  - Refer forwards in the discourse
    Although he loved fishing, Paul went skating with Mary.

Pronouns

- Seldom refer more than two sentences back
- Requires a salient referent as antecedent

- Antecedent Indicators:
  - Recency
  - Grammatical role
  - Parallellism
  - Repeated mention
  - Verb semantics

Text Coherence

- Coherence relations
  - Result
  - Explanation
  - Parallel
  - Elaboration
  - Occasion

A Discourse Tree

Discourse Structure

- John went to the bank to deposit his paycheck (S1)
- He then took a train to Bill’s car dealership (S2)
- He needed to buy a car (S3)
- The company he works for now isn’t near any public tranportation (S4)
- He also wanted to talk to Bill about their softball league (S5)

Inference 1

Rule: If it rains the ground gets wet
Observation: It rains
Conclusion: The ground gets wet

Deduction: rule + observation → conclusion (modus ponens)
Induktion: observation +conclusion → rule (modus tollens)
(Abduktion: rule + conclusion - (?!) → obeservation)
Inference 2

- John hid Bill’s car keys. He was drunk.
  1. ⇐ John usually does stupid things when drunk
  2. ⇐ Bill often drives when drunk
- Bill was drunk. John hid his car keys.
  1. ⇐ Bill tends to “borrow” cars when drunk
  2. ⇐ Bill often drives his car when drunk

Background Knowledge

- The problem of encoding inference is usually said to AI-complete
- AI-completeness indicates that the problem requires all of the knowledge – and utilities to utilize it – that humans possess

Different Levels

- Syntax
  - Rules for constructing grammatical sentences
- Semantics
  - Rules for assigning meaning to statements
- Pragmatics
  - Rules (of thumb) for applying contextual constraints on the semantics of a statement

Pragmatics

- The study of meaning contained by utterances in situations (Leech, 1983)
- Relates the content of a clause (semantics) with the content of an utterance of that clause (pragmatics)
- Pragmatic rules often rules of thumb
- Dialogues – Cooperative Principles

Grice’ Cooperative Principle

- **Quantity**
  - Don’t say more that necessary
- **Quality**
  - Don’t say anything you do not believe in or have proof of
- **Relevance**
  - A response should be an answer to the question
- **Form**
  - Be clear / avoid ambiguity
  - Be concise
  - Be methodical

Discourse, what for?

- Information Retrieval
- Summarization
- Pronoun Resolution
  - Natural Language Generation
What Is Natural Language Generation?

A process of constructing a natural language output from non-linguistic inputs that maps meaning to text.

Related Simple Text Generation

- Canned text
  - Outputs predefined text

- Template filling
  - Outputs predefined text with predefined variable words/phrases

Areas of Use

NLG techniques can be used to:
- generate textual weather forecasts from representations of graphical weather maps
- summarize statistical data extracted from a database or spreadsheet
- explain medical info in a patient-friendly way
- describe a chain of reasoning carried out by an expert system
- paraphrase information in a diagram for inexperienced users

Goals of a NLG System

To supply text that is:
- correct and relevant information
  - non redundant
- suiting the needs of the user
- in an understandable form
- in a correct form

Choices for NLG

- Content selection
- Lexical selection
- Sentence structure
  - Aggregation
  - Referring expressions
  - Orthographic realisation
- Discourse structure

Example Architecture
Discourse Planner

1. Text shemata
   • Use consistent patterns of discourse structure
   • Used for manuals and descriptive texts

2. Rhetorical Relations
   • Uses the Rhetorical Structure Theory
   • Used for varied generation tasks

Discourse Planner – Rhetorical Relations

Rhetorical Structure Theory
(Mann & Thompson 1988)

• Nucleus
  • Multi-nuclear
  • Satellite

Surface Realisation

1. Systemic Grammar
   • Using functional categorization
   • Represents sentences as collections of functions
   • Directed, acyclic and/or graph

2. Functional Unification Grammar
   • Using functional categorization
   • Unifies generation grammar with a feature structure

Surface Realisation – Systemic Grammar

• Emphasises the functional organisation of language
• Surface forms are viewed as the consequences of selecting a set of abstract functional features
• Choices correspond to minimal grammatical alternatives
• The interpolation of an intermediate abstract representation allows the specification of the text to accumulate gradually

Surface Realisation – Systemic Grammar

Mood
Major
Minor
Mood
Declarative
Interrogative
Imperative
Present-Participle
Past-Participle
Infinitive
Polar
Bound Relative
Wh-
**Surface Realisation – Functional Unification Grammar**

**Basic idea:**
- Input specification in the form of a **FUNCTIONAL DESCRIPTION**, a recursive `<attribute,value>` matrix
- The grammar is a large functional description with alternations representing choice points
- Realisation is achieved by unifying the input FD with the grammar FD

```plaintext
((cat clause)
  (process ((type composite)
            (relation possessive)
            (lex 'hand')))
  (participants ((agent ((cat pers_pro)
                             (gender feminine)))
                  ((affected n ((cat np)
                                 (lex 'editor'))
                     ((possessor n ((cat np)
                                      (lex 'draft'))))))
  She hands the draft to the editor.
```

**Microplanning 1**

- Lexical selection
- Referring expression generation
- Morphological realization
- Syntactic realization
- Orthographic realization

**Microplanning 3**

**Aggregation**

Some possibilities:
- Simple conjunction
- Ellipsis
- Set introduction

**Aggregation**

Without aggregation:
- It has a snack bar.
- It has a restaurant car.

With set introduction:
- It has {a snack bar, a restaurant car}.
- It has a snack bar and a restaurant car.

Caution! Need to avoid changing the meaning:
- John bought a TV.
- Bill bought a TV.
- John and Bill bought a TV.

**Further Reading**

- Siggen
- Allen 1995: Natural Language Understanding
  - [http://www.uni-giessen.de/~g91062/Seminare/gk-cl/Allen95/al1995co.htm](http://www.uni-giessen.de/~g91062/Seminare/gk-cl/Allen95/al1995co.htm)