NLP
Introduction to NLP

Text Generation
Basic NLP Pipeline

- (U)nderstanding and (G)eneration
One Definition of NLG

- Natural language generation is the process of deliberately constructing a natural language text in order to meet specified communicative goals.

  [McDonald 1992]
What is NLG?

- Mapping meaning to text
- Stages:
  - Content selection
  - Discourse structure
  - Sentence structure
  - Lexical choice
  - Linguistic realization
Examples

• Weather forecasts
• Directions
• Train schedules
• Descriptions of objects
Content Selection

- What is important?
- What content should be presented to the user
- Depends on the domain, the audience, communicative goal
- Semantic representation
Discourse Structure

• Rhetorical structure
• Order in which to present the information
Lexical Choice

- What words to use
- Paraphrases
- Different syntactic structures
Sentence Structure

• Linking words
  – e.g., therefore, and, but

• Aggregation
  – You have two messages from Jason and one from Claire
Linguistic Realization

• Grammaticality
  – Agreement
  – Morphology
  – Adjective order
  – Articles (a/an)

• Relative clauses

• Referring expressions

• Presentation constraints
  – Media
  – Length
  – User background
  – Genre constraints

• Extralinguistic issues
  – Layout
Choices

• NLG is about choices
  – Content
  – Coherence
  – Style
  – Media
  – Syntax
  – Aggregation
  – Referring expressions
  – Lexical choice
  – Syntax

• Evaluation