

Building Natural Language Generation Systems

Robert Dale and Ehud Reiter

1

What This Tutorial is About

- The design and construction of systems which
 - produce understandable texts in English or other human languages ...
 - ... from some underlying non-linguistic representation of information ...
 - ... using knowledge about language and the application domain.

Based on:

E Reiter and R Dale [1999] *Building Natural Language Generation Systems*. Cambridge University Press.

2

Goals of the Tutorial

- For managers:
 - to provide a broad overview of the field and what is possible today
- For implementers:
 - to provide a realistic assessment of available techniques
- For researchers:
 - to highlight the issues that are important in current applied NLG projects

3

Overview

- 1 An Introduction to NLG
- 2 Requirements Analysis and a Case Study
- 3 The Component Tasks in NLG
- 4 NLG in Multimedia and Multimodal Systems
- 5 Conclusions and Pointers

4

An Introduction to NLG

- What is Natural Language Generation?
- Some Example Systems
- Types of NLG Applications
- When are NLG Techniques Appropriate?
- NLG System Architecture

5

What is NLG?

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

[McDonald 1992]

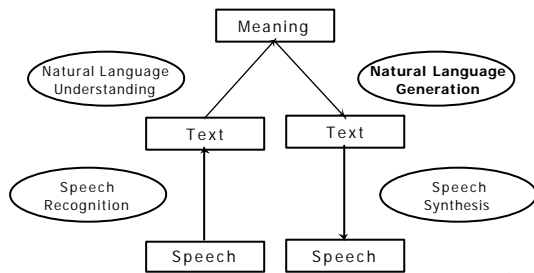
6

What is NLG?

- **Goal:**
 - computer software which produces understandable and appropriate texts in English or other human languages
- **Input:**
 - some underlying non-linguistic representation of information
- **Output:**
 - documents, reports, explanations, help messages, and other kinds of texts
- **Knowledge sources required:**
 - knowledge of language and of the domain

7

Language Technology



8

Example System #1: FoG

- **Function:**
 - Produces textual weather reports in English and French
- **Input:**
 - Graphical/numerical weather depiction
- **User:**
 - Environment Canada (Canadian Weather Service)
- **Developer:**
 - CoGenTex
- **Status:**
 - Fielded, in operational use since 1992

9

PlanDoc: Input

```
RUNID fiberall FIBER 6/19/93 act yes
FA 1301 2 1995
FA 1201 2 1995
FA 1401 2 1995
FA 1501 2 1995
ANF co 1103 2 1995 48
ANF 1201 1301 2 1995 24
ANF 1401 1501 2 1995 24
END. 856.0 670.2
```

13

PlanDoc: Output

This saved fiber refinement includes all DLC changes in Run-ID ALLDLC. RUN-ID FIBERALL demanded that PLAN activate fiber for CSAs 1201, 1301, 1401 and 1501 in 1995 Q2. It requested the placement of a 48-fiber cable from the C0 to section 1103 and the placement of 24-fiber cables from section 1201 to section 1301 and from section 1401 to section 1501 in the second quarter of 1995. For this refinement, the resulting 20 year route PWE was \$856.00K, a \$64.11K savings over the BASE plan and the resulting 5 year IFC was \$670.20K, a \$60.55K savings over the BASE plan.

14

Example System #3: STOP

- Function:
 - Produces a personalised smoking-cessation leaflet
- Input:
 - Questionnaire about smoking attitudes, beliefs, history
- User:
 - NHS (British Health Service)
- Developer:
 - University of Aberdeen
- Status:
 - Undergoing clinical evaluation to determine its effectiveness

15

STOP: Input

SMOKING QUESTIONNAIRE
Please answer by marking the most appropriate box for each question like this:

Q1 Have you smoked a cigarette in the last week, even a puff?
YES NO

Please complete the following questions. Please return the questionnaire unanswered in the envelope provided. Thank you.

Please read the questions carefully. If you are not sure how to answer, just give the best answer you can.

Q2 Home situation:
Live alone Live with husband/wife/partner Live with other adults Live with children

Q3 Number of children under 16 living at home boys girls

Q4 Does anyone else in your household smoke? (If so, please mark all boxes which apply)
husband/wife/partner other family member others

Q5 How long have you smoked for? ...10... years
Tick here if you have smoked for less than a year

16

STOP: Output

Dear Ms Cameron

Thank you for taking the trouble to return the smoking questionnaire that we sent you. It appears from your answers that although you're not planning to stop smoking in the near future, you would like to stop if it was easy. You think it would be difficult to stop because *smoking helps you cope with stress, it is something to do when you are bored, and smoking stops you putting on weight.* However, you have reasons to be confident of success if you did try to stop, and there are ways of coping with the difficulties.

17

Example System #4: TEMSIS

- Function:
 - Summarises pollutant information for environmental officials
- Input:
 - Environmental data + a specific query
- User:
 - Regional environmental agencies in France and Germany
- Developer:
 - DFKI GmbH
- Status:
 - Prototype developed; requirements for fielded system being analysed

18

TEMSIS : Input Query

```
((LANGUAGE FRENCH)
(GRENZWERTLAND GERMANY)
(BESTAETIGE-MS T)
(BESTAETIGE-SS T)
(MESSSTATION \"Voelklingen City\")
(DB-ID \"#2083\")
(SCHADSTOFF \"#19\")
(ART MAXIMUM)
(ZEIT ((JAHR 1998)
(MONAT 7)
(TAG 21))))
```

19

TEMSIS : Output Summary

- Le 21/7/1998 à la station de mesure de Völklingen -City, la valeur moyenne maximale d'une demi-heure (Halbstundenmittelwert) pour l'ozone atteignait 104.0 µg/m³. Par conséquent, selon le décret MKK (MKK-Verordnung), la valeur limite autorisée de 120 µg/m³ n'a pas été dépassée.
- Der höchste Halbstundenmittelwert für Ozon an der Meßstation Völklingen -City erreichte am 21. 7. 1998 104.0 µg/m³, womit der gesetzlich zulässige Grenzwert nach MKK-Verordnung von 120 µg/m³ nicht überschritten wurde.

20

Types of NLG Applications

- Automated document production
 - weather forecasts, simulation reports, letters, ...
- Presentation of information to people in an understandable fashion
 - medical records, expert system reasoning, ...
- Teaching
 - information for students in CAL systems
- Entertainment
 - jokes (?), stories (??), poetry (???)

21

The Computer's Role

Two possibilities:

#1 The system produces a document without human help:

- weather forecasts, simulation reports, patient letters
- summaries of statistical data, explanations of expert system reasoning, context-sensitive help, ...

#2 The system helps a human author create a document:

- weather forecasts, simulation reports, patient letters
- customer-service letters, patent claims, technical documents, job descriptions, ...

22

When are NLG Techniques Appropriate?

Options to consider:

- Text vs Graphics
 - Which medium is better?
- Computer generation vs Human authoring
 - Is the necessary source data available?
 - Is automation economically justified?
- NLG vs simple string concatenation
 - How much variation occurs in output texts?
 - Are linguistic constraints and optimisations important?

23

Enforcing Constraints

- Linguistically well-formed text involves many constraints:
 - orthography, morphology, syntax
 - reference, word choice, pragmatics
- Constraints are automatically enforced in NLG systems
 - automatic, covers 100% of cases
- String-concatenation system developers must explicitly enforce constraints by careful design and testing
 - A lot of work
 - Hard to guarantee 100% satisfaction

24

Example: Syntax, aggregation

- Output of existing Medical AI system:
**The primary measure you have chosen, CXR shadowing, should be justified in comparison to TLC and walking distance as my data reveals they are better overall. Here are the specific comparisons:
TLC has a lower patient cost TLC is more tightly distributed TLC is more objective walking distance has a lower patient cost**

25

Example: Pragmatics

- Output of system which gives English versions of database queries:
 - **The number of households such that there is at least 1 order with dollar amount greater than or equal to \$100.**
 - Humans interpret this as number of "households which have placed an order \geq \$100"
 - Actual query returns count of all households inDB if there is any order in theDB (from any household) which is \geq \$100

26

NLG System Architecture

- The Component Tasks in NLG
- A Pipelined Architecture
- Alternative Architectures

27

Component Tasks in NLG

- 1 Content determination
- 2 Document structuring
- 3 Aggregation
- 4 Lexicalisation
- 5 Referring expression generation
- 6 Linguistic realisation
- 7 Structure realisation

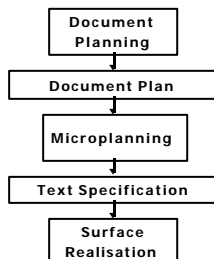
28

Tasks and Architecture in NLG

<ul style="list-style-type: none">• Content Determination• Document Structuring	Document Planning
<ul style="list-style-type: none">• Aggregation• Lexicalisation• Referring Expression Generation	Micro-planning
<ul style="list-style-type: none">• Linguistic Realisation• Structure Realisation	Surface Realisation

29

A Pipelined Architecture



30

Other Architectures

#1: Variations on the 'standard' architecture:

- shift tasks around
- allow feedback between stages

#2: An integrated reasoner which does everything:

- represent all tasks in the same way: eg as constraints, axioms, plan operators ...
- feed specifications into a constraint-solver, theorem-prover ...

31

Research Questions

- When is text the best means to communicate with the user?
- When is NLG technology better than string concatenation?
- Is there an architecture which combines the theoretical elegance of integrated approaches with the engineering simplicity of the pipeline?
- How should document plans and text specifications be represented?

32

Overview

- 1 An Introduction to NLG
- 2 Requirements Analysis and a Case Study
- 3 The Component Tasks in NLG
- 4 NLG in Multimedia and Multimodal Systems
- 5 Conclusions and Pointers

33

A Case Study in Applied NLG

- Each month an institutional newsletter publishes a summary of the month's weather
- The summaries are based on automatically collected meteorological data
- The person who writes these summaries will no longer be able to
- The institution wants to continue publishing the reports and so is interested in using NLG techniques to do so

34

A Weather Summary

MARSHFIELD (Marquette University No 1)
On Campus, Square 79

TEMPERATURES (C)

Hottest Max for Mth: 18.1 Warmer than average
Hottest Max for June (20 yrs): 17.2
Highest Max (Hottest Day): 23.9 on 01
Lowest Min (Coldest Day): 13.0 on 12
Hottest Min for Mth: 08.2 Much warmer than ave
Hottest Min for June (20 yrs): 08.4
Lowest Min (Coldest Night): 02.6 on 09
Highest Min (Hottest Night): 11.9 on 24

RAINFALL (mm) (24 hrs to 09:00)

Total Rain for Mth: 95.4 on 12 days
Slightly below average.
Wettest Day (24h to 09:00): 26.4 on 11
Average for June (25 yrs): 109.0 on 10
Total for 06 mths so far: 345.0 on 75 days
Average for 06 mths (25 yrs): 792.0 on 131 days
Annual Average Rainfall (25 yrs): 1142.8 on 131 days

WIND RUN (at 2m height) (km) (24 hrs to 09:00)

Total Wind Run for Mth: 1660
Wettest Day (24 hrs to 09:00): 185 on 26, 172 on 27
Calmmest Day (24 hrs to 09:00): 09 on 16

SUNRISE & SUNSET

Date	Sunrise	Sunset	Difference
01 Jun	06:52	16:54	10:02
11 Jun	06:37	16:51	09:56
21 Jun	07:00	16:54	09:54
30 Jun	07:01	16:37	09:36

(Sunset times began to get later after about June 11)
(Sunrise times continue to get later until early July)
(Some we can take advantage of the later sunsets)

SUMMARY

The month was warmer than average with average rainfall, but the total rain so far for the year is still very depleted. The month began with mild to warm maximums, and became cooler as the month progressed, with some very cold nights such as June 09 with 02.6. Some other years have had much colder June nights than this, and minimums below zero in June are not very unusual. The month was mostly calm, but strong winds blew on 23, 24 and 26, 27. Fog occurred on 17, 18 after some rain on 17. Heavy rain fell on 11 June.

35

Output: A Weather Summary

The month was warmer than average with average rainfall, but the total rain so far for the year is still very depleted. The month began with mild to warm maximums, and became cooler as the month progressed, with some very cold nights such as June 09 with 02.6. Some other years have had much colder June nights than this, and minimums below zero in June are not very unusual. The month was mostly calm, but strong winds blew on 23, 24 and 26, 27. Fog occurred on 17, 18 after some rain on 17, heavy rain fell on 11 June.

36

The Input Data

- A set of 16 data elements collected automatically every 15 minutes: air pressure, temperature, wind speed, rainfall ...
- Preprocessed to construct DailyWeatherRecords:

```
((type dailyweatherrecord)
 (date ((day ...)
        (month ...)
        (year ...)))
 (temperature ((minimum ((unit degrees-centigrade)
                          (number ...)))
               (maximum ((unit degrees-centigrade)
                          (number ...)))))
 (rainfall ((unit millimetres)
            (number ...))))
```

37

Other Available Data

- Historical Data: Average temperature and rainfall figures for each month in the Period of Record (1971 to present)
- Historical Averages: Average values for temperature and rainfall for the twelve months of the year over the period of record

38

Requirements Analysis

The developer needs to:

- understand the client's needs
- propose a functionality which addresses these needs

39

Corpus-Based Requirements Analysis

A corpus:

- consists of examples of output texts and corresponding input data
- specifies 'by example' the functionality of the proposed NLG system
- is a very useful resource for design as well as requirements analysis

40

Corpus-Based Requirements Analysis

Four steps:

- assemble an initial corpus of (human-authored) output texts and associated input data
- analyse the information content of the corpus texts in terms of the input data
- develop a target text corpus
- create a formal functional specification

41

Step 1: Creating an Initial Corpus

- Collect a corpus of input data and associated (human-authored) output texts
- One source is archived examples of human-authored texts
- If no human-authored examples of the required texts exist, ask domain experts to produce examples
- The corpus should cover the full range of texts expected to be produced by the NLG system

42

Initial Text (April 1995)

SUMMARY

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again, after almost catching up during March. Mars Creek dried up again on 30th April at the waterfall, but resumed on 1st May after light rain. This is the fourth time it dried up this year.

43

Step 2: Analyzing the Content of the Texts

- Goal:
 - to determine where the information present in the texts comes from, and the extent to which the proposed NLG system will have to manipulate this information
- Result:
 - a detailed understanding of the correspondences between the available input data and the output texts in the initial corpus

44

Information Types in Text

- Unchanging text
- Directly-available data
- Computable data
- Unavailable data

45

Unchanging Text

SUMMARY

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again, after almost catching up during March. Mars Creek dried up again on 30th April at the waterfall, but resumed on 1st May after light rain. This is the fourth time it dried up this year.

46

Directly Available Data

SUMMARY

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again, after almost catching up during March. Mars Creek dried up again on 30th April at the waterfall, but resumed on 1st May after light rain. This is the fourth time it dried up this year.

47

Computable Data

SUMMARY

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again, after almost catching up during March. Mars Creek dried up again on 30th April at the waterfall, but resumed on 1st May after light rain. This is the fourth time it dried up this year.

48

Unavailable Data

SUMMARY

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again, after almost catching up during March. Mars Creek dried up again on 30th April at the waterfall, but resumed on 1st May after light rain. This is the fourth time it dried up this year.

49

Solving the Problem of Unavailable Data

- More information can be made available to the system: this may be expensive
 - add sensors to Mars Creek?
- If the system is an authoring-aid, a human author can add this information
 - system produces the first two sentences, the human adds the second two
- The target corpus can be revised to eliminate clauses that convey this information
 - only produce the first two sentences

50

Step 3: Building the Target Text Corpus

Mandatory changes:

- eliminate unavailable data
- specify what text portions will be human-authored

Optional changes:

- simplify the text to make it easier to generate
- improve human-authored texts
- enforce consistency between human authors

51

Target Text

The month was rather dry with only three days of rain in the middle of the month. The total for the year so far is very depleted again.

52

Step 4: Functional Specification

- Based on an agreed target text corpus
- Explicitly states role of human authoring, if present at all
- Explicitly states structure and range of inputs to be used

53

Initial Text #2

The month was our driest and warmest August in our 24 year record, and our first 'rainless' month. The 26th August was our warmest August day in our record with 30.1, and our first 'hot' August day (30). The month forms part of our longest dry spell 47 days from 18 July to 02 September 1995. Rainfall so far is the same as at the end of July but now is very deficient.

54

Target Text #2

The month was the driest and warmest August in our 24 year record, and the first rainless month of the year. 26th August was the warmest August day in our record with 30.1, and the first hot day of the month. Rainfall for the year is now very deficient.

55

The Case Study So Far

We'll assume that:

- We have located the source data
- We have preprocessed the data to build the DailyWeatherRecords
- We have constructed an initial corpus of texts
- We have modified the initial corpus to produce a set of target texts

56

Is it Worth Using NLG?

- For one summary a month probably not, especially given the simplifications required to the texts to make them easy to generate
- However, the client is interested in a pilot study:
 - in the future there may be a shift to weekly summaries
 - there are many automatic weather data collection sites, each of which could use the technology

57

Research Issues

- Development of an appropriate corpus analysis methodology
- Using expert system knowledge acquisition techniques
- Automating aspects of corpus analysis
- Integrating corpus analysis with standard requirements analysis procedures

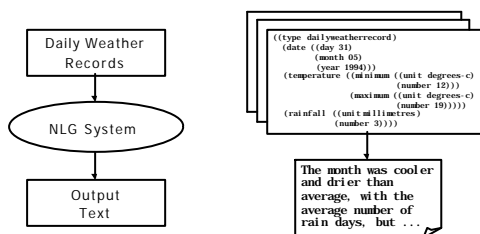
58

Overview

- 1 An Introduction to NLG
- 2 Requirements Analysis and a Case Study
- 3 The Component Tasks in NLG
- 4 NLG in Multimedia and Multimodal Systems
- 5 Conclusions and Pointers

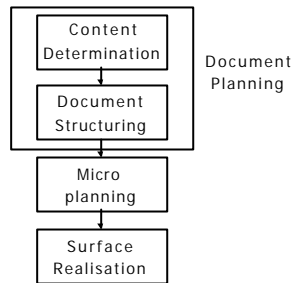
59

Inputs and Outputs



60

The Architectural View



61

Document Planning

Goals:

- to determine what information to communicate
- to determine how to structure this information to make a coherent text

Two Common Approaches:

- methods based on observations about common text structures
- methods based on reasoning about discourse coherence and the purpose of the text

62

Content Determination

Based on MESSAGES, predefined data structures which:

- correspond to informational elements in the text
- collect together underlying data in ways that are convenient for linguistic expression

Core idea:

- from corpus analysis, identify the largest possible agglomerations of informational elements that do not pre-empt required flexibility in linguistic expression

63

Content Determination in WeatherReporter

- Routine messages
 - MonthlyRainFallMsg,
MonthlyTemperatureMsg,
RainSoFarMsg,
MonthlyRainyDaysMsg
- Always constructed for any summary to be generated

64

Content Determination in WeatherReporter

A MonthlyRainfallMsg:

```
((message-id msg091)
 (message-type monthlyrainfall)
 (period ((month 04)
          (year 1996)))
 (absolute-or-relative relative-to-average)
 (relative-difference ((magnitude ((unit millimeters)
                                   (number 4)))
                       (direction +))))
```

65

Content Determination in WeatherReporter

- Significant Event messages
 - RainEventMsg,
RainSpellMsg,
TemperatureEventMsg,
TemperatureSpellMsg
- Only constructed if the data warrants their construction: eg if rain occurs on more than a specified number of days in a row

66

Content Determination in WeatherReporter

A RainSpellMsg:

```
((message-id msg096)
 (message-type rainspellmsg)
 (period ((begin ((day 04)
                  (month 02)
                  (year 1995)))
          (end ((day 11)
                (month 02)
                (year 1995))))
 (duration ((unit day)
            (number 8))))
 (amount ((unit millimetres)
          (number 120))))
```

67

Content Determination

Alternative strategies:

- Build all possible messages from the underlying data, then select for expression those appropriate to the context of generation
- Identify information required for context of generation and construct appropriate messages from the underlying data

The content determination task is essentially a domain-dependent expert-systems-like task

68

Document Structuring via Schemas

Basic idea (after McKeown 1985):

- texts often follow conventionalised patterns
- these patterns can be captured by means of 'text grammars' that both dictate content and ensure coherent structure
- the patterns specify how a particular document plan can be constructed using smaller schemas or atomic messages
- can specify many degrees of variability and optionality

69

Document Structuring via Schemas

Implementing schemas:

- simple schemas can be expressed as grammars
- more flexible schemas usually implemented as macros or class libraries on top of a conventional programming language, where each schema is a procedure
- currently the most popular document planning approach in applied NLG systems

70

Deriving Schemas From a Corpus

Using the Target Text Corpus:

- take a small number of similar corpus texts
- identify the messages, and try to determine how each message can be computed from the input data
- propose rules or structures which explain why message *x* is in text *A* but not in text *B* — this may be easier if messages are organised into a taxonomy
- discuss this analysis with domain experts, and iterate
- repeat the exercise with a larger set of corpus texts

71

Document Planning in WeatherReporter

A Simple Schema:

WeatherSummary →
MonthlyTempMsg
MonthlyRainfallMsg
RainyDaysMsg
RainSoFarMsg

72

Document Planning in WeatherReporter

A More Complex Set of Schemata:

```
WeatherSummary →
  TemperatureInformation RainfallInformation
TemperatureInformation →
  MonthlyTempMsg [ExtremeTempInfo] [TempSpellsInfo]
RainfallInformation →
  MonthlyRainfallMsg [RainyDaysInfo] [RainSpellsInfo]
RainyDaysInfo →
  RainyDaysMsg [RainSoFarMsg]
...
```

73

Schemas in Practice

Tests and other machinery are often made explicit:

```
(put-template maxwert-grenzwert "M01"
 (:PRECOND (:CAT DECL-E
            :TEST ((pred-eq 'maxwert-grenzwert)
                    (not (status-eq (theme) 'no))))
 :ACTIONS (:TEMPLATE (:RULE MAX-AVG-VALUE-E (self))
                    ". As a result, "
                    (:RULE EXCEEDS-THRESHOLD-E (self))
                    ". )))
```

74

Schemas: Pros and Cons

Advantages of schemas:

- Computationally efficient
- Allow arbitrary computation when necessary
- Naturally support genre conventions
- Relatively easy to acquire from a corpus

Disadvantages

- Limited flexibility: require predetermination of possible structures
- Limited portability: likely to be domain-specific

75

Document Structuring via Explicit Reasoning

Observation:

- Texts are coherent by virtue of relationships that hold between their parts — relationships like narrative sequence, elaboration, justification ...

Resulting Approach:

- segment knowledge of what makes a text coherent into separate rules
- use these rules to dynamically compose texts from constituent elements by reasoning about the role of these elements in the overall text

76

Document Structuring via Explicit Reasoning

- Typically adopt AI planning techniques:
 - Goal = desired communicative effect
 - Plan constituents = messages or structures that combine messages (subplans)
- Can involve explicit reasoning about the user's beliefs
- Often based on ideas from Rhetorical Structure Theory

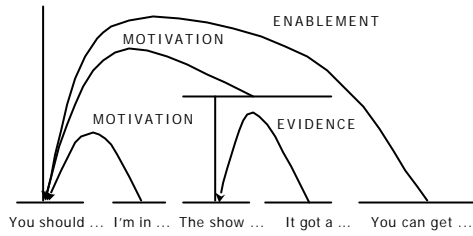
77

Rhetorical Structure Theory

- D1: You should come to the Northern Beaches Ballet performance on Saturday.
- D2: I'm in three pieces.
- D3: The show is really good.
- D4: It got a rave review in the Manly Daily.
- D5: You can get the tickets from the shop next door.

78

Rhetorical Structure Theory



79

An RST Relation Definition

Relation name: Motivation

Constraints on N:

Presents an action (unrealised) in which the hearer is the actor

Constraints on S:

Comprehending S increases the hearer's desire to perform the action presented in N

The effect:

The hearer's desire to perform the action presented in N is increased

80

Document Structuring in WeatherReporter

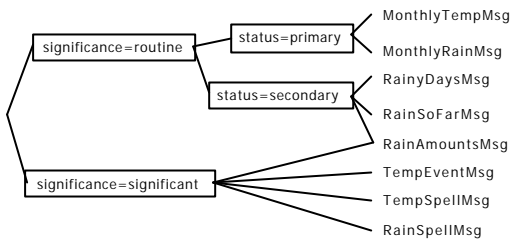
Three basic rhetorical relationships:

- SEQUENCE
- ELABORATION
- CONTRAST

Applicability of rhetorically-based planning operators determined by attributes of the messages

81

Message Attributes



82

Document Structuring in WeatherReporter

SEQUENCE

- Two messages can be connected by a SEQUENCE relationship if both have the attribute message-status = primary

83

Document Structuring in WeatherReporter

ELABORATION

- Two messages can be connected by an ELABORATION relationship if:
 - they are both have the same message-topic
 - the nucleus has message-status = primary

84

Document Structuring in WeatherReporter

CONTRAST

- Two messages can be connected by a CONTRAST relationship if:
 - they both have the same message-topic
 - they both have the feature absolute-or-relative = relative-to-average
 - they have different values for relative-difference:direction

85

Document Structuring in WeatherReporter

- Select a start message
- Use rhetorical relation operators to add messages to this structure until all messages are consumed or no more operators apply
- Start message is any message with message-significance = routine

86

Document Structuring using Relation Definitions

The algorithm:

```
DocumentPlan = StartMessage
MessageSet = MessageSet - StartMessage
repeat
  – find a rhetorical operator that will allow attachment
    of a message to the DocumentPlan
  – attach message and remove from MessageSet
until MessageSet = 0 or no operators apply
```

87

Target Text #1

The month was cooler and drier than average, with the average number of rain days, but the total rain for the year so far is well below average. Although there was rain on every day for 8 days from 11th to 18th, rainfall amounts were mostly small.

88

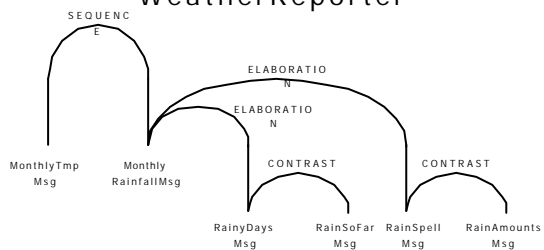
Document Structuring in WeatherReporter

The Message Set:

MonthlyTempMsg ("cooler than average")
MonthlyRainfallMsg ("drier than average")
RainyDaysMsg ("average number of rain days")
RainSoFarMsg ("well below average")
RainSpellMsg ("8 days from 11th to 18th")
RainAmountsMsg ("amounts mostly small")

89

Document Structuring in WeatherReporter



90

More Complex Algorithms

Adding complexity, following [Marcu 1997]:

- Assume that multiple DocumentPlans can be created from a set of messages and relations
- Assume that a desirability score can be assigned to each DocumentPlan
- Determine the best DocumentPlan

91

Document Planning

- Result is a DOCUMENT PLAN: a tree structure populated by messages at its leaf nodes
- Next step: realising the messages as text

92

Research Issues

- The use of expert system techniques in content determination -- for example, case based reasoning
- Principled ways of integrating schemas and relation-based approaches to document structuring
- A better understanding of rhetorical relations
- Knowledge acquisition -- eg, methodologies for creating content rules, schemas, and relation applicability conditions for a particular application

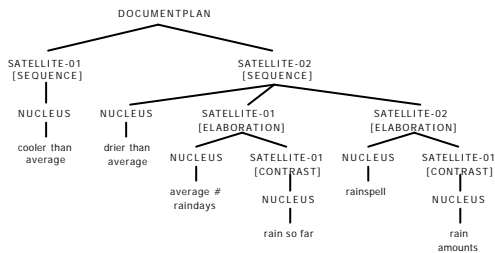
93

A Simple Realiser

- We can produce one output sentence per message in the document plan
- A specialist fragment of code for each message type determines how that message type is realised

94

The Document Plan



95

A Simple Realiser

For the MonthlyTemperatureMsg:

```

TempString = case (TEMP - AVERAGETEMP)
[2.0 ... 2.9]: ' very much warmer than average.'
[1.0 ... 1.9]: ' much warmer than average.'
[0.1 ... 0.9]: ' slightly warmer than average.'
[-0.1 ... -0.9]: ' slightly cooler than average.'
[-1.0 ... -1.9]: ' much cooler than average.'
[-2.0 ... -2.9]: ' very much cooler than average.'
endcase
Sentence = 'The month was' + TempString
  
```

96

One Message per Sentence

- The Result:
The month was cooler than average.
The month was drier than average.
There were the average number of rain days.
The total rain for the year so far is well below average.
There was rain on every day for 8 days from 11th to 18th.
Rainfall amounts were mostly small.
- The Target Text:
The month was cooler and drier than average, with the average number of rain days, but the total rain for the year so far is well below average.
Although there was rain on every day for 8 days from 11th to 18th, rainfall amounts were mostly small.

97

Simple Templates

Problems with simple templates in this example:

- MonthlyTemp and MonthlyRainfall don't always appear in the same sentence
- When they do appear in the same sentence, they don't always appear in the same order
- Each can be realised in different ways: eg 'very warm' vs 'warmer than average'
- Additional information may or may not be incorporated into the same sentence

98

Microplanning

Goal:

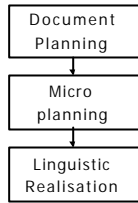
- To convert a document plan into a sequence of sentence or phrase specifications

Tasks:

- Paragraph and Sentence Aggregation
- Lexicalisation
- Reference

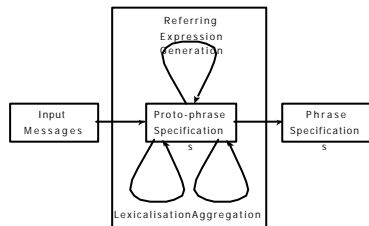
99

The Architectural View



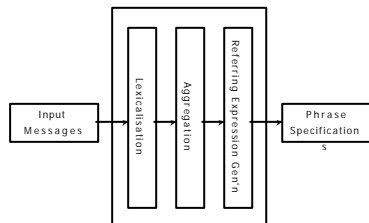
100

Interactions in Microplanning



101

Pipelined Microplanning



102

Aggregation

Combinations can be on the basis of

- information content
- possible forms of realisation

Some possibilities:

- Simple conjunction
- Ellipsis
- Embedding
- Set introduction

103

Some Examples

Without aggregation:

- **Heavy rain fell on the 27th.**
Heavy rain fell on the 28th.

With aggregation via simple conjunction:

- **Heavy rain fell on the 27th and heavy rain fell on the 28th.**

With aggregation via ellipsis:

- **Heavy rain fell on the 27th and [] on the 28th.**

With aggregation via set introduction:

- **Heavy rain fell on [the 27th and 28th].**

104

An Example: Embedding

Without aggregation:

- **March had a rainfall of 120mm.**
It was the wettest month.

With aggregation:

- **March, which was the wettest month, had a rainfall of 120mm.**

105

Choice Heuristics

There are usually many ways to aggregate a given message set: how do we choose?

- conform to genre conventions and rules
- observe structural properties
 - for example, only aggregate messages which are siblings in the document plan tree
- take account of pragmatic goals

106

Pragmatics: STOP Example

- Making the text friendlier by adding more empathy:
 - **It's clear from your answers that you don't feel too happy about being a smoker and it's excellent that you are going to try to stop.**
- Making the text easier for poor readers:
 - **It's clear from your answers that you don't feel too happy about being a smoker. It's excellent that you are going to try to stop.**

107

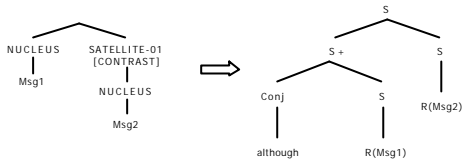
Aggregation in WeatherReporter

Sensitive to rhetorical relations:

- If two messages are in a SEQUENCE relation they can be conjoined at the same level
- If one message is an ELABORATION of another it can either be conjoined at the same level or embedded as a minor clause or phrase
- If one message is a CONTRAST to another it can be conjoined at the same level or embedded as a minor clause or phrase

108

An Aggregation Rule



Lexicalisation

- The process of choosing words to communicate the information in messages
- Methods:
 - templates
 - decision trees
 - graph-rewriting algorithms

110

Lexical Choice

If several lexicalisations are possible, consider:

- user knowledge and preferences
- consistency with previous usage
 - in some cases, it may be best to vary lexemes
- interaction with other aspects of microplanning
- pragmatics
 - **It is encouraging that you have many reasons to stop.** (more precise meaning)
 - **It's good that you have a lot of reasons to stop.** (lower reading level)

111

WeatherReporter: Variations in Describing Rainfall

Variations in syntactic category:

S: [rainfall was very poor indeed]

NP: [a much worse than average rainfall]

AP: [much drier than average]

Variations in semantics:

Absolute: [very dry]

[a very poor rainfall]

Comparative: [a much worse than average rainfall]

[much drier than average]

112

WeatherReporter: Aggregation and Lexicalisation

Many different results are possible:

- The month was cooler and drier than average. There were the average number of rain days, but the total rain for the year so far is well below average. There was rain on every day for 8 days from 11th to 18th, but rainfall amounts were mostly small.
- The month was cooler and drier than average. Although the total rain for the year so far is well below average, there were the average number of rain days. There was a small amount of rain on every day for 8 days from 11th to 18th.

113

Referring Expression Generation

How do we identify specific domain objects and entities?

Two issues:

- Initial introduction of an object
- Subsequent references to an already salient object

114

Initial Reference

Introducing an object into the discourse:

- use a full name
 - **Jeremy**
- relate to an object that is already salient
 - **Jane's goldfish**
- specify physical location
 - **the goldfish in the bowl on the table**

Poorly understood; more research is needed

115

Subsequent Reference

Some possibilities:

- Pronouns
 - **It swims in circles.**
- Definite NPs
 - **The goldfish swims in circles.**
- Proper names, possibly abbreviated
 - **Jeremy swims in circles.**

116

Choosing a Form of Reference

Some suggestions from the literature:

- use a pronoun if it refers to an entity mentioned in the previous clause, and there is no other entity in the previous clause that the pronoun could refer to
- otherwise use a name, if a short one exists
- otherwise use a definite NP

Also important to conform to genre conventions -- for example, there are more pronouns in newspaper articles than in technical manuals

117

Example

I am taking the Caledonian Express tomorrow. It is a much better train than the Grampian Express. The Caledonian has a real restaurant car, while the Grampian just has a snack bar. The restaurant car serves wonderful fish, while the snack bar serves microwaved mush.

118

Referring Expression Generation in WeatherReporter

- Referring to months:
 - June 1999
 - June
 - the month
 - next June
- Relatively simple, so can be hardcoded in document planning

119

Research Issues

- How do we make microplanning choices?
- How do we perform higher-level aggregation, such as forming paragraphs from sentences?
- How do we lexicalise if domain concepts do not straightforwardly map into words?
- What is the best way of making an initial reference to an object?

120

Realisation

Goal:

to convert text specifications into actual text

Purpose:

to hide the peculiarities of English (or whatever the target language is) from the rest of the NLG system

121

Realisation Tasks

Structure Realisation :

- Choose markup to convey document structure

Linguistic Realisation :

- Insert function words
- Choose correct inflection of content words
- Order words within a sentence
- Apply orthographic rules

122

Structure Realisation

- Add document markup
- An example: means of marking paragraphs:
 - HTML `<P>`
 - LaTeX (blank line)
 - RTF `\par`
 - SABLE (speech) `<BREAK>`
- Depends on the document presentation system
- Usually done with simple mapping rules

123

Linguistic Realisation

Techniques:

- Bi-directional Grammar Specifications
- Grammar Specifications tuned for Generation
- Template-based Mechanisms

124

Bi-directional Grammar Specifications

- Key idea: one grammar specification used for both realisation and parsing
- Can be expressed as a set of declarative mappings between semantic and syntactic structures
- Different processes applied for generation and analysis
- Theoretically elegant
- To date, sometimes used in machine-translation systems, but almost never used in other applied NLG systems

125

Problems with the Bi-directional Approach

- Output of an NLU parser (a semantic form) is very different from the input to an NLG realiser (a text specification)
- Debatable whether lexicalisation should be integrated with realisation
- Difficult in practice to engineer large bidirectional grammars
- Difficulties handling fixed phrases

126

Grammar Specifications tuned for Generation

- Grammar provides a set of choices for realisation
- Choices are made on the basis of the input text specification
- Grammar can *only* be used for NLG
- Working software is available
 - Bateman's KPML
 - Elhadad's FUF/SURGE
 - CoGenTex's RealPro

127

Example: KPML

- Linguistic realiser based on Systemic Functional Grammar
- Successor to Penman
- Uses the Nigel grammar of English
- Smaller grammars for several other languages available
- Incorporates a grammar development environment

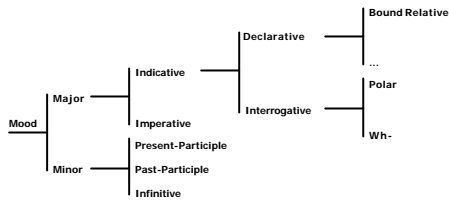
128

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

129

A Systemic Network



130

K P M L

How it works:

- choices are made using INQUIRY SEMANTICS
- for each choice system in the grammar, a set of predicates known as CHOOSERS are defined
- these tests are functions from the internal state of the realiser and host generation system to one of the features in the system the chooser is associated with

131

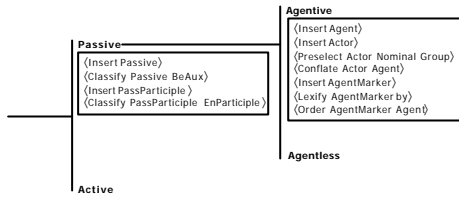
K P M L

Realisation Statements:

- small grammatical constraints at each choice point build up to a grammatical specification
- ⟨Insert SUBJECT⟩: an element functioning as subject will be present
- ⟨Conflate SUBJECT ACTOR⟩: the constituent functioning as SUBJECT is the same as the constituent that functions as ACTOR
- ⟨Order FINITE SUBJECT⟩: FINITE must immediately precede SUBJECT

132

Realisation Statements



133

An SPL input to KPML

```

(l / greater-than-comparison
 :tense past
 :exceed-q (l a) exceed
 :command-offer-q notcommandoffer
 :proposal-q notproposal
 :domain (n / one-or-two-d-time :lex month :determiner the)
 :standard (a / quality :lex average :determiner zero)
 :range (c / sense-and-measure-quality :lex cool)
 :inclusive (r / one-or-two-d-time
  :lex day
  :number plural
  :property-ascription (r / quality :lex rain)
  :size-property-ascription
  (av / scalable-quality :lex the-av-no-of)))

The month was cooler than average with the average number of rain days.
  
```

134

Observation

- These approaches are geared towards broad-coverage linguistically sophisticated treatments
- Most current applications don't require this sophistication: much of what is needed can be achieved by templates or even canned text
- But many applications will need deeper treatments for some aspects of the language
- Solution: integrate canned text, templates and "realisation from first principles" in one system

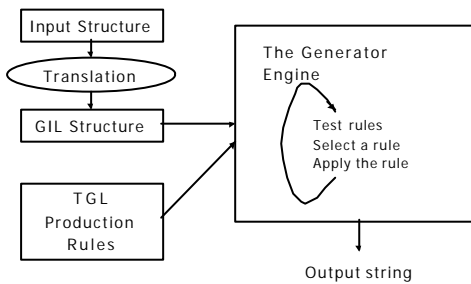
135

Busemann's TG/2

- Integrates canned text, templates and context free rules
- All expressed as production rules whose actions are determined by conditions met by the input structure
- Input structures specified in GIL, the Generation Interface Language; allows a portable interface
- Output can easily include formatting elements

136

TG/2 Overview



137

A GIL Input Structure

```
[ (COOP wertueberschreitung)
  (TIME [(PRED dof)
        (NAME [(DAY 31)
              (MONTH 12)
              (YEAR 1996)])])
  (POLLUTANT so2)
  (SITE "Völklingen-City")
  (THRESHOLD-VALUE [(AMOUNT 1000)
                   (UNIT mkg-m3)])
  (DURATION [(DAY 30)])
  (SOURCE [(LAW-NAME vdi-richtlinie-
            (THRESHOLD-TYPE mikwert)])]
  (EXCEEDS [(STATUS yes)
            (TIMES 4)])]
```

138

A TGL Rule

```
(defproduction threshold-exceeding "WU01"  
  (:PRECOND (:CAT DECL  
    :TEST ((coop-eq 'threshold-exceeding)  
            (threshold-value-p)))  
  :ACTIONS (:TEMPLATE  
    (:OPTRULE PPTIME (get-param 'time)  
    (:OPTRULE SITEV (get-param site)  
    (:RULE THTYPE (self)  
    (:OPTRULE POLL (get-param 'pollutant)  
    (:OPTRULE DUR (get-param 'duration)  
    " (" (:RULE VAL (get-param 'threshold-value)  
        (:OPTRULE LAW (get-param 'law-name)) " )"  
    (:RULE EXCEEDS (get-param 'exceeds)) " , "  
    :CONSTRAINTS (:GENDER (THTYPE EXCEEDS) :EQ))))))
```

139

TG/2 Output

On 31-12-1996 at the measurement station at Völklingen-City, the MIK value (MIK-Wert) for sulphur dioxide over a period of 30 days (1000 µg/m³ according to directive VDI 2310 (VDI-Richtlinie 2310)) was exceeded four times.

140

Pros and Cons

- No free lunch: every existing tool requires a steep learning curve
- Approaches like TG/2 may be sufficient for most applications
- Linguistically-motivated approaches require some theoretical commitment and understanding, but promise broader coverage and consistency

141

Morphology and Orthography

Realiser must be able to:

- inflect words
- apply standard orthographic spelling changes
- add punctuation
- add standard punctuation rules

142

Research Issues

- How can different techniques (and linguistic formalisms) be combined?
- What are the costs and benefits of using templates instead of 'deep' realisation techniques?
- How do layout issues affect realisation?

143

Overview

- 1 An Introduction to NLG
- 2 Requirements Analysis and a Case Study
- 3 The Component Tasks in NLG
- 4 NLG in Multimedia and Multimodal Systems
- 5 Conclusions and Pointers

144

Document Types

- Simple ASCII (eg email messages)
 - relatively straightforward: just words and punctuation symbols
- Printed documents (eg newspapers, technical documents)
 - need to consider typography, graphics
- Online documents (eg Web pages)
 - need to consider hypertext links
- Speech (eg radio broadcasts, information by telephone)
 - need to consider prosody
- Visual presentation (eg TV broadcasts, MS Agent)
 - need to consider animation, facial expressions

145

Typography

- Character attributes (italics, boldface, colour, font)
 - can be used to indicate emphasis or other aspects of use: typographic distinctions carry meaning
- Layout (itemised lists, section and chapter headings)
 - allows indication of structure, can enable information access
- Special constructs provide sophisticated resources
 - boxed text, margin notes, ...

146

Typographically-flat Text

When time is limited, travel by limousine, unless cost is also limited, in which case go by train. When only cost is limited a bicycle should be used for journeys of less than 10 kilometers, and a bus for longer journeys. Taxis are recommended when there are no constraints on time or cost, unless the distance to be travelled exceeds 10 kilometers. For journeys longer than 10 kilometers, when time and cost are not important, journeys should be made by hire car.

147

Structured Text

When only time is limited:

travel by Limousine

When only cost is limited:

travel by Bus if journey more than 10 kilometers

travel by Bicycle if journey less than 10 kilometers

When both time and cost are limited:

travel by Train

When time and cost are not limited:

travel by Hire Car if journey more than 10 kilometers

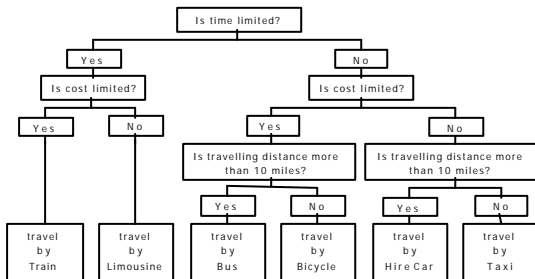
travel by Taxi if journey less than 10 kilometers

148

Tabular Presentation

149

Diagrammatic Presentation



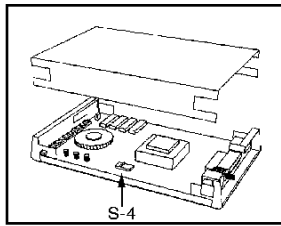
150

Text and Graphics

- Which is better? Depends on:
 - type of information communicated
 - expertise of user
 - delivery medium
- Best approach: use both!

151

An Example from WIP



Push the code switch S-4 to the right.

152

Similarities between Text and Graphics

- Both contain sublanguages
 - Both permit conversational implicature
 - Structure is important in both
 - Both express discourse relations
- Do we need a media-independent theory of communication?

153

Hypertext

- Generate Web pages!
 - Dynamic hypertext
- Models of hypertext
 - browsing
 - question-space
 - dialogue
- Model-dependent issues
 - click on a link twice: same result both times?
 - discourse models for hypertext

154

Example: Peba



155

Speech Output

- The NLG Perspective: enhances output possibilities
 - communicate via spoken channels (eg, telephone)
 - add information (eg emotion, importance)
- The speech synthesis perspective: intonation carries information
 - Need information about syntactic structure, information status, homographs
 - Currently obtained by text analysis
 - Could be obtained from an NLG system automatically: the idea of concept-to-speech

156

Examples

- *John took a bow*
 - Difficult to determine which sense of *bow* is meant, and therefore how to pronounce it, from text analysis
 - But an NLG system knows this
- *John washed the dog*
 - Should stress that part of the information that is new
 - An NLG system will know what this is

157

Overview

- 1 An Introduction to NLG
- 2 Requirements Analysis and a Case Study
- 3 The Component Tasks in NLG
- 4 NLG in Multimedia and Multimodal Systems
- 5 Conclusions and Pointers

158

Applied NLG in 1999

- Many NLG applications being investigated
- Few actually fielded
 - but at least there are some: in 1989 there were none
- We are beginning to see reusable software, and specialist software houses
- More emphasis on mixing simple and complex techniques
- More emphasis on evaluation
- We believe the future is bright

159

Resources: SIGGEN

SIGGEN (ACL Special Interest Group for Generation)

- Web site at <http://www.dynamimulti media.com.au/siggen>
 - resources: software, papers, bibliographies
 - conference and workshop announcements
 - job announcements
 - discussions
 - NLG people and places

160

Resources: Conferences and Workshops

- International Workshop on NLG: every two years
- European Workshop on NLG: every two years, alternating with the International Workshops
- NLG papers at ACL, EACL, ANLP, IJCAI, AAAI ...
- See SIGGEN Web page for announcements

161

Resources: Companies

Some software groups with NLG experience:

- CoGenTex: see <http://www.cogentex.com>
 - FoG, LFS, ModelExplainer, CogentHelp
 - RealPro and Exemplars packages available free for academic research
- ERLI: see <http://www.erli.com>
 - AlethGen, MultiMeteo system

162

Resources: Books

Out soon:

Ehud Reiter and Robert Dale [1999]
Building Natural Language Generation Systems
Cambridge University Press

Don't miss it!

163
