A Case Study in Applied NLG • Each month an institutional newsletter publishes a summary of the month's weather **NLG Lecture 7:** The summaries are based on automatically collected **Content planning 1** meteorological data The person who writes these summaries will no longer be able to The institution wants to continue publishing the reports Slides adapted from Jon Oberlander and so is interested in using NLG techniques to do so With thanks to Robert Dale and Ehud Reiter. Informatics 1 **Content planning** A Weather Summary

- Discourse coherence meets generation
- Today:
 - Principles, example
- Next time:
 - Restaurants (+ museums)

MARSFIELD (Macquarie University No 1) On Campus, Square F9

TEMPERATURES (C)

 Mean Max for Mth:
 18.1 warmer than average

 Mean Max for June (20 yrs):
 17.2

 Highest Max (warmest Day):
 23.9 on 01

 Lowest Max (coldest Day):
 13. on 12

 Mean Min for Mth:
 08.2 Much warmer than average

 Mean Min for June (20 yrs):
 06.4

 Lowest Min (Coldest Night):
 02.6 on 09

 Highest Min (warmest Night):13.5 on 24

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

 Total Rain for Mth:
 90.4 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:
 542.0 on 72

Very depleted. Average for 06 mths (25 yrs): 762.0 on 71 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

 windiest Day (24 hrs to 09:00):
 189 on 24, 185 on 26, 172 on 27

 Calmest 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	06:57	16:53	09:56
21 Jun	07:00	16:54	09:54
30 Jun	07:01	16:57	09:56

(Sunset times began to get later after about June 11) (Sunrise times continue to get later until early July) (Soon 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 or 23, 24 and 26, 27. Fog occurred on 17, 18 after some rain on 17, heavy rain fell on 11 June.

Output: A Weather Summary The Input Data The month was warmer than average with average A set of 16 data elements collected automatically rainfall, but the total rain so far for the every 15 minutes: air pressure, temperature, vear is still very depleted. The month began wind speed, rainfall ... with mild to warm maximums. and became cooler Preprocessed to construct DailyWeatherRecords: as the month progressed, with some very cold nights such as June 09 with 02.6. Some other ((type dailyweatherrecord) vears have had much colder June nights than (date ((day ...) this, and minimums below zero in June are not (month ...) verv unusual. The month was mostly calm. but (vear ...))) (temperature ((minimum ((unit degrees-centigrade)))) strong winds blew on 23, 24 and 26, 27. Fog (number ...))) occurred on 17, 18 after some rain on 17, heavy (maximum ((unit degrees-centrigrade)) rain fell on 11 June. (number ...))))) (rainfall ((unit millimetres) (number ...)))) 6 **Other Available Data Inputs and Outputs** Historical Data: **Dailv Weather** - Average temperature and rainfall figures for each month in Records the Period of Record (1971 to present) + ((type dailyweatherrecord) Historical Averages: (date ((day 31) Historical Data (month 05) - Average values for temperature and rainfall for the twelve (year 1994))) (temperature ((minimum ((unit degrees-c) months of the year over the period of record (number 12))) (maximum ((unit degrees-c) (number 19))))) (rainfall ((unit millimetres) (number 3)))) NLG System The month was cooler and drier than Output average, with the average number of Text rain days, but ... 8





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

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

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 that 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, users, and other stakeholders, and iterate
- repeat the exercise with a larger set of corpus texts

Document Planning in WeatherReporter

A Simple Schema:

WeatherSummary → MonthlyTempMsg MonthlyRainfallMsg RainyDaysMsg RainSoFarMsg



Document Structuring via Explicit Reasoning Rhetorical Structure Theory Typically adopt AI planning techniques: D1: You should come to the Northern Beaches Ballet performance on Saturday. - Goal = desired communicative effect D2: I'm in three pieces. - Plan constituents = messages or structures that combine D3: The show is really good. messages (subplans) D4: It got a rave review in the Manly Daily. • Can involve explicit reasoning about the user's beliefs D5: You can get the tickets from the shop next door. Often based on ideas from Rhetorical Structure Theory An RST Relation Definition: Motivation **Rhetorical Structure Theory** Relation name: Motivation Constraints on N: Presents an action (unrealised) in which the hearer is the actor Constraints on S: ENABLEMENT MOTIVATION 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 MOTIVATION **EVIDENCE** increased

You should ... I'm in ... The show ... It got a ... You can get ...



Document Structuring in WeatherReporter

ELABORATION

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

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

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

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 = {} or no operators apply



One content plan, many text plans

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



Document Planning

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

A Simple Realiser

41

42

- 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

Example: 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

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