# Modelling the Task of Summarising Time Series Data Using KA Techniques

Somayajulu G. Sripada, Ehud Reiter, Jim Hunter, and Jin Yu
Dept of Computing Science
University of Aberdeen, Aberdeen, UK
{ssripada,ereiter,jhunter,jyu}@csd.abdn.ac.uk

Ian P. Davy
WNI Oceanroutes
Oceanroutes (U.K.) Ltd.
Aberdeen, UK
iand@abd.wni.com

Abstract

The SUMTIME project aims to develop better techniques for producing natural language summaries of time-series data. The initial phase of the project has focused on understanding how human experts perform this task, via knowledge acquisition and corpus analysis exercises. This has led to a number of observations, of which the most important is that producing human like summaries involves as much processing of domain knowledge as that of communicative and linguistic knowledge. Summarisation is not simply a verbalisation of how a set of numbers changes over time. Instead it is the process in which the author adds value to the raw data using his expertise and presents exactly the information that is relevant to the end-user.

## 1. Introduction

Time-series data (TSD) is ubiquitous in the modern world. Currently, such data is usually presented to humans either graphically or as tables of raw numbers. In the SUMTIME project at Aberdeen University (www.csd.abdn.ac.uk/research/sumtime/), we are attempting to develop better models and techniques for generating textual summaries of time-series data, using a combination of natural-language generation (NLG) and automatic time-series analysis techniques in the light of knowledge derived from observing humans summarising TSD. In this paper we discuss some of the observations we have made during our study of how humans produce textual summaries of time-series data (which has been our main activity to date), including the impact of domain and user knowledge on the summarisation task.

## 2. Background

### 2.1 Time-Series Data

A time-series data set is a collection of values of a set of parameters over time. For example, regular measurements of the temperature and blood pressure of a hospital patient would constitute a time series data set. Time-series data sets can range in size from 5-10 measurements on one parameter (for example, enrolment over the past ten years in CS1001) to millions of measurements on tens of thousands of parameters (for example, sensor data from a space shuttle mission).

Human beings frequently need to examine and make inferences from a time-series data set. For example, a hospital doctor will look at time-series patient data when trying to diagnose and treat a patient, and a university administrator will look at time-series data about course enrolment when estimating how much lab resources a particular course is likely to require. Currently, human examination of time-series data is generally done either by direct inspection of the data (for small data sets), by graphical visualisation, or by statistical analyses.

## 2.2 Textual Summaries of Time-Series Data

Sometimes people want textual summaries of time-series data as well as graphical summaries or statistical analyses. For example, newspapers publish textual summaries of weather predictions, the results of polls and surveys, and stock market activity, instead of just showing numbers and graphs. This may be because graphical depictions of time-series data require time and skill to interpret, which is not always available. For example, a doctor rushing to the side of a patient who is suffering from a heart attack may not have time to examine a set of graphs, and a newspaper reader may not have the statistical knowledge necessary to interpret raw poll results. Also, some data sets (such as the space shuttle one mentioned above) are simply too large to present graphically as a whole in any meaningful fashion.

Currently textual descriptions of time-series data must be produced manually, which makes them expensive and also means they can not be produced instantly. Graphical depictions of data, in contrast, can be produced quickly and cheaply using off-the-shelf computer software; this may be one reason why they are so popular. If textual summaries of time-series data could be automatically produced by software as cheaply and as quickly as graphical depictions, then they might be more widely used.

Summarising time-series data is also interesting from the perspective of natural-language generation because the input data is truly non-linguistic. Many NLG systems are components of text-to-text systems (such as machine translation and text summarisation), which means that the conceptual structure of their inputs is already linguistic; this is also often the case for NLG systems which use AI knowledge bases as their inputs. Such systems may, for example, need to decide whether the concept INCREASING should be expressed lexically as *increasing*, *rising*, or *going up*, but they do not need to determine whether the concept INCREASING is an appropriate description of a data set. A time-series summarisation system, in contrast, must make such decisions, which means that it must represent and reason about the meaning of 'close-to-linguistic' concepts such as INCREASING[1].

## 2.3 SUMTIME

SUMTIME is a new project at the University of Aberdeen, whose goal is to develop better models and techniques for generating textual summaries of time-series data, in part by trying to integrate leading edge techniques for natural language generation and time-series analysis. We are initially working in two domains:

- **Meteorology**: Producing weather forecasts from numerical weather simulations, in collaboration with WNI Oceanroutes, a private-sector meteorological services organisation.
- **Gas Turbines**: Summarising sensor readings from a gas turbine, in collaboration with Intelligent Applications, a leading developer of monitoring software for gas turbines.

We will start working on a third domain in the second half of the project. This is likely to be in the medical area, perhaps summarising clinical patient data from babies in neo-natal units, although this is not definite.

The meteorological and gas turbine domains differ in the amount of data that needs to be summarised. Generally a weather forecast is based on a time series data set which contains tens of parameters (the exact number depends on the type of forecast). The numerical simulation generally gives parameter values at hourly or 3-hourly intervals, so a 7-day weather forecast would be based on tens of values for each of these tens of parameters, or on the order of magnitude of a thousand numbers in all. The gas-turbine domain, in contrast, involves much more data. Again detailed numbers depend on the particular turbine being monitored, but as an order or magnitude we get hundreds of parameters measured at one second intervals over a period of a day or so, which means on the order of ten million numbers in all.

The first year of SUMTIME has primarily been devoted to trying to understand how human experts summarise time-series data, via a combination of knowledge acquisition techniques and corpus analysis. This is discussed below.

## 2.4 Previous Work

There have been a number of previous systems, which summarised time-series data, of which perhaps the best known is FOG [2]. FOG also produced weather forecasts from numerical weather simulations, and is in operational use in Environment Canada. At least as described in the published papers, FOG's research emphasis

---

[1] Note that we cannot simply say that a time-series is increasing if the last value is higher than the first value. The time series -1, 1, -1, 1, -1, 1, -1, 1, -1, 1, for example, satisfies this criteria but would not normally be described as *increasing*.

was on multi-lingual microplanning and realisation (using the terminology of Reiter and Dale [9]) using Meaning-Text Theory [8] models. The research emphasis in SUMTIME, in contrast, is on content determination and document planning. In other words, our research will focus on determining the conceptual content of a forecast, and we plan to use existing techniques where possible to decide how to express this content linguistically. FOG, in contrast, focused from a research perspective on the issue of how a conceptual representation of a forecast could be expressed in both English and French.

Closer in spirit to SUMTIME is TREND [1], which produced text summaries of historical weather data (for example, the weather over the past month). TREND, like SUMTIME, focused on content determination and conceptual issues, and it also, like SUMTIME, used sophisticated time-series analysis techniques (wavelets). Perhaps the main difference between TREND and SUMTIME is that TREND described the visual appearance of a graph without reference to a domain or user model, or an underlying task. That is, TREND's goal was to describe what a time-series graph ‘looked like’ to a human viewer, and it was partially based on models of how people visually perceived graphs. SUMTIME, in contrast, has the goal of producing a text summary that is useful to a human user, which means that the system, like a human meteorologist (see Section 4 below), will need to take into account knowledge about the domain and the user.

Systems, which summarise time-series data, have also been built in other domains, such as stock market reports [6] and statistical summaries [4]. Again to the best of our knowledge, these systems used relatively simple techniques for concept formation, although many of them used sophisticated techniques for microplanning and realisation.

There has also been a considerable amount of research on sophisticated techniques for analysing and abstracting time-series data, such as Shahar [13]. However, these systems have not been connected to text generators, instead their output has generally been displayed graphically [12] or as raw abstractions.

The goal of SUMTIME is in part to bridge the gap between these two strands of research, NLG and time-series analysis, and do this in a way which to some degree replicates the manner in which a human expert would summarise time-series data.

# 3. Knowledge Acquisition

The first year of SUMTIME has focused, after an initial literature review, on trying to understand how human experts summarise time-series data. The main objectives of the Knowledge Acquisition (KA) activities have been [3, 11]:

1. To determine the task model of summarisation

2. To determine the types of knowledge required for the task model.

3. To acquire all the required types of knowledge in detail.

The KA activities in the two domains (meteorology & gas turbine) have been different because the two domains differ in the following aspects:

1. Gas-turbine engineers do not currently write textual summaries of sensor readings, which makes corpus acquisition and expert observation difficult. On the other hand, in the domain of meteorology we have access to a large corpus of human-written forecasts.
2. The large amount of data in the gas-turbine domain meant that we had to spend a considerable amount of time developing tools for displaying data to experts in KA sessions. In the domain of meteorology the data sets are much smaller.

In SUMTIME, we have carried out the following KA activities tailored to the needs of the individual domains:

- Discussions with experts.
- Think-aloud sessions, in which experts were asked to 'think-aloud' as they examined and summarised a data set.
- Observation of experts as they wrote summaries in their workplace environment.
- Corpus analysis of a collection of time-series data sets and corresponding manual written summaries.
- Collaborative prototype development, in which we essentially built an initial system to an expert's specification, and then compared its output to manually written summaries of the same data sets.

| day | hour | wind dir | wind speed |
|---|---|---|---|
| 2-3-01 | 6 | W | 10 |
| 2-3-01 | 9 | W | 11 |
| 2-3-01 | 12 | W | 13 |
| 2-3-01 | 15 | WNW | 14 |
| 2-3-01 | 18 | WNW | 15 |
| 2-3-01 | 21 | WNW | 11 |
| 3-3-01 | 0 | NW | 9 |
| 3-3-01 | 3 | NW | 8 |
| 3-3-01 | 6 | WSW | 9 |
| 3-3-01 | 9 | SSW | 11 |
| 3-3-01 | 12 | SSW | 21 |
| 3-3-01 | 15 | SSW | 22 |
| 3-3-01 | 18 | SSW | 24 |
| 3-3-01 | 21 | SW | 26 |
| 4-3-01 | 0 | SW | 27 |
| 4-3-01 | 3 | SW | 24 |
| 4-3-01 | 6 | WSW | 22 |
| 4-3-01 | 9 | W | 21 |
| 4-3-01 | 12 | W | 21 |
| 4-3-01 | 15 | W | 16 |
| 4-3-01 | 18 | WSW | 10 |
| 4-3-01 | 21 | SSW | 14 |
| 5-3-01 | 0 | S | 22 |

**Table 1. Wind predictions from TAB model on 2-3-01**

```
06-24 GMT, 02-Mar 2001:
WSW 10-14 RISING 14-18 AND VEERING WNW THIS
AFTERNOON, VEERING NW 8-12 THIS EVENING

00-24 GMT, 03-Mar 2001:
NW 8-12 BACKING SSW 18-22 IN THE MORNING THEN RISING 26-30 IN
THE EVENING

00-24 GMT, 04-Mar 2001:
SSW 26-30 VEERING W 18-22 IN THE MORNING, BACKING WSW
10-14 BY EVENING THEN BACKING S 22-26 LATE EVENING

5-Mar 2001 and 6-Mar 2001:
SSW 22-26 RISING 35-40 ON MONDAY, VEERING WSW 28-32
TUESDAY MORNING THEN BACKING SW 20-24 LATER

7-Mar 2001 and 8-Mar 2001:
SW 20-24 BACKING SE 32-38 WEDNESDAY AFTERNOON,
VEERING S-SW 20-24 EARLY ON THURSDAY
```

**Figure 1. Wind texts from human-written forecast for 2-3-01**

# 4. Observations Made

A more complete description of the results of our knowledge acquisition activities is given in [14]. Here we just summarise some of the more important and interesting observations. We first describe the domain independent observations followed by domain dependent observations. Data sets in the Gas turbine domain are very large (nearly two hundred parameters sampled every second) and therefore for simplicity's sake we describe the domain independent observations with the help of example data set from the weather domain. We will focus on the particular task of producing textual summaries of wind speed and direction, in weather forecasts intended for offshore oilrigs.

Table 1 shows wind speed and wind direction predictions extracted from the 'TAB' model file produced by the numerical simulation on 2 March 2001,[2] and Figure 1 shows the text summaries produced by a human forecaster on 2 March 2001. This forecast is broken up into 5 forecast periods. The forecasts for the first three forecast periods (up to the end of 4 Mar) are based on the TAB model data shown in Table 1. The forecasts for the last two forecast periods (5 March to 8 March) are based on the less accurate MMO model, which we have not shown in this paper. Note that a complete forecast for an offshore oil rig would describe many other meteorological parameters, including wind at 50M altitude, visibility, waves, temperature, cloud cover, and precipitation.

[2] In fact the TAB file was extracted from a model built by a human forecaster from a numerical simulation; the forecaster used a graphical editor, which allowed him to interpolate and adjust the numbers produced by the numerical simulation.

## 4.1 Domain Independent Observations

### 4.1.1 Qualitative Overview

One of the most interesting observations in our KA sessions was that experts usually formed a qualitative overview of the underlying system as a whole before writing summaries. This overview is used to decide what to do about boundary cases or unusual situations. For example, if a forecaster sees an outlier in a data set, such as a N wind becoming NE for one time period and then reverting to N, he or she will use the qualitative overview to decide whether this is realistic and should be reported, or whether it is an artefact of the numerical simulation and should be ignored.

When asked about the role of an overview in writing the summaries the expert has replied that it's main role is to facilitate reasoning with the input data. This ability to reason with input data helps to draw inferences about the state of the underlying system. When included in the summaries, these inferences would be more useful to the end user than just the raw data. Details about our observations on qualitative overviews, and a 'two-stage' model for content determination based on these observations, are given in [15].

While inferring the overview, during a KA session, the expert focused entirely on domain reasoning as if he were solving a meteorological problem. He has ignored the communicative issues indicating that the creation of the overview is more like a problem solving activity than like a communicating one. The objective of this problem solving activity seems to be the determination of the overall state of the underlying system (gas turbine or atmosphere).

The above observation suggests that the task of overview creation is a knowledge-based activity, which can be studied in its own right as a task of building an expert system. In SUMTIME, we intend to scale down the problem and make it a component in a system SUMTIME-MOUSAM that we are currently building. SUMTIME-MOUSAM would act as a test bed for carrying out experiments with summarising weather data (initially) and gas turbine data (later). This is explained further in the Section 5.1.3 below.

We have observed overview formation by human experts in other domains as well, including the SUMTIME gas turbine domain and a previous project on generating personalised smoking-cessation letters [10].

### 4.1.2 Impact of User

Another fact that emerged from our knowledge acquisition efforts was that forecasters sometimes consider how users (rig staff) will use the forecast when deciding on its content (same with the gas turbine domain as well, where the personnel in charge of the particular turbine need to act based on the summaries and the summaries need to take this into consideration). Of course, anyone who has looked at weather forecasts, realises that the kind of information present in a forecast depends on the user; for example, oil rig forecasts include information about wind at 50m altitude which is not present in newspaper forecasts, because it is not of interest to the general public. However, forecasters may also adjust the detailed content of forecasts depending on user and task models.

For example, during one of our think-aloud sessions, where experts spoke their thoughts aloud while writing a forecast, the forecaster decided to use the phrase *19-24* to describe wind speed when the TAB file predicted a wind speed of 19kt. This is because the forecaster knew that even if the average wind speed in the period was 19kt, the actual speed was going to vary minute by minute and often be above 20kt, and he also knew that rig staff used different operational procedures (for example for docking supply boats) when the wind exceeded 20kt. Hence he decided to emphasise that the wind was likely at times to exceed 20kt by using a wind speed range of “19-24”, instead of a range centred around 19kt, such as “17-21”.

It is possible that the fact that the forecast in Figure 1 starts with WSW instead of W (which is what the input data in Table 1 suggests) is due to user factors. In particular, if the forecaster is unsure whether the wind will be coming from W or WSW, it probably makes sense for him to state the broadest variety of wind directions in the phrase (that is *WSW ….. veering WNW* instead of *W ….. veering WNW*), in order to warn rig staff that the wind might be WSW and they shouldn't count on it being W or WNW.

User needs also influence the level of detail of forecasts. For example, if the wind is less than 10kt, then the exact wind speed and direction usually have no impact on rig operations (with some exceptions, such as when a rig is flaring gas); this is one reason why many forecasters use generic phrases such as *less than 10* for light winds (Sect 4.2.1).

#### 4.1.3 Data Reliability

One of the first observations we made is that forecasters do not treat the numbers in the TAB and MMO files as gospel. For example they may adjust the numbers based on their knowledge of local geographical effects which are not considered in the numerical simulation, such as the effect of peninsulas on wind speed at nearby off-shore sites. Forecasters may also adjust numbers based on their own experience; for example, the specification given to us by one of the forecasters for wind text production adjusted the wind speed in the TAB file using another parameter, lapse rate (a measure of the stability of the atmosphere), because the forecaster felt that this adjustment improved the accuracy of the prediction.

Perhaps more subtly, forecasters also made judgements about the reliability of the numbers, especially temporally. For example, it might be clear that a storm is going to move through an area, thus increasing wind speeds, and a forecaster might agree with the model's prediction of how high wind speeds will rise but feel less confident about the model's prediction of when this will happen, that is when the storm will actually reach the oil rig. Thus forecasts usually use relatively precise (given the 3-hour granularity of the TAB file) temporal terms such as *early afternoon*, but sometimes use less precise terms such as *later* or *for a time*. This can again be observed in Figure 1 where the forecast for 5 and 6 March uses vague terms such as *on Monday* and *later*.

In the gas turbine domain, the issue of data reliability may be either due to noise in data transmission or due to sensor failure.

## 4.2 Domain Dependent Observations

### 4.2.1 Forecaster Variations

One surprise to us was the degree of variation between individual forecasters. This included

- lexical variations: For example, some forecasters describe an increase in wind speed as *increasing*, whereas others use the term *rising*. Another example is that some forecasters always use a four knot range for wind speed, such as *12-16*, while others vary the size of the range depending on the circumstances. The forecaster who wrote the text in figure 1 for example, generally used four knot ranges but used some larger ranges (*35-40*, *32-38*) in later forecast periods.
- content variations: For example, some forecasters use phrases such as *less than 10* for low wind speeds, without going into details, whereas others give detailed descriptions such as *2-6*. Another example is that when the wind is varying a lot, some forecasters report every change explicitly, while others use general terms such as *occasionally* or *mainly*.

Variations among individual forecasters are, in fact, one of the reasons why forecasting organisations are interested in computer text generation. For example forecaster variations means that an offshore oil rig could one day get a forecast predicting a wind speed of *less than 10* and the next day get a forecast predicting a wind speed of *2-6*, when the wind in fact is identical on both days. This could perhaps confuse oil rig staff if they did not realise the cause of the variation.

### 4.2.2 Algorithmic Issues

The above sections present non-numeric factors that influence the content of a forecast, such as qualitative overview, impact of user, data reliability, and forecaster variations. But what about the numbers themselves, what type of time-series analysis is best for producing summaries?

The comparison with the initial algorithm suggested by the forecasters themselves was particularly useful here. This algorithm essentially used a threshold model, where wind speed and direction were mentioned at the beginning of a forecast period and then whenever they changed by a threshold amount (typically 5 kt for speed and 22.5 degrees for direction, although the actual threshold amount depended on the wind speed). Corpus analysis and observation of forecasters as they worked suggested that

- A linear segmentation model [5,7] might be a better fit to what forecasters wrote than a threshold one.
- If wind speed or direction is changing slowly but consistently, this will often be reported even if the overall change is small.

- Human forecasters like to report wind direction and speed as changing at the same time, even if the model suggests that in fact the speed change happens a bit before or after the direction change.

The above behaviour would emerge from an algorithm, which was based on linear segmentation and tried to optimise a combination of the accuracy of the segmentation and the length of the text. The length optimisation would bias the algorithm towards reporting speed and direction changes as occurring at the same time.

#### 4.2.3 Non-event Reporting

In the domain of gas turbines, the main task is reporting unexpected behaviour. One interesting observation here was that summaries are required to report a non-event – an event that should have occurred based on the state of affairs inferred from the input data, but never actually occurred. For example, if a controller attempts to increase fuel flow, but fuel flow does not change. This clearly indicates that the summarising system must be in a position to compute the state of the gas turbine based on the input data, and reason with that state to make predictions about the data. This once again points indirectly towards the need of an overview explained in 4.1.1.

## 5. Evaluation and Implementation

Although KA will continue, the focus of the project now includes testing the hypotheses about summarising time-series data, which emerged from our KA activities and which were (partially) described above. In particular, we are currently implementing a test-bed, which includes a core system, into which new modules can easily be inserted, and a testing framework that automatically runs the system with the new modules on corpus input data and compares the result against the corresponding manually written texts in the corpus. Since we are primarily interested in content issues, the comparison to corpus texts will be made on a conceptual level (using a conceptual mark-up scheme we have developed) as well as on a text level.

### 5.1 SUMTIME-MOUSAM

SUMTIME-MOUSAM (Figure 2) is a framework system that makes it easy to implement and evaluate new ideas for summarising weather data (initially) and gas turbine data (at a later stage). The framework consists of

- "Infrastructure" software for accessing data files, automatically comparing software output to human-written texts, regression testing of new software versions.
- An ontology which defines a conceptual level of representation of weather forecasts
- A corpus of human-written forecasts with their corresponding conceptual representations

- An initial system based on observations from KA, for generating conceptual representations and then forecast texts from data.

#### 5.1.1 Data/Corpus

This component manages input data/corpus for the entire system. SUMTIME has collected a large corpus of weather data and its corresponding forecast texts from WNI Oceanroutes. Currently the size of the corpus is 500 (data set - forecast text pairs) and is growing as we receive data and its corresponding forecast text twice daily from WNI Oceanroutes. In the gas turbine domain, we have collected continuous one-month data with 250 (approx.) analog channels and 700 (approx.) digital channels sampled every second. SUMTIME-MOUSAM views the corpus data as a set of 'projects', each project consisting of a set of input data sets and an output forecast text. Project stores the configuration information for a unique experiment carried out on SUMTIME-MOUSAM. An experiment here means the generation of a summary text (or conceptual representation of the summary) from a definite input data set by using one of the models we want to test.

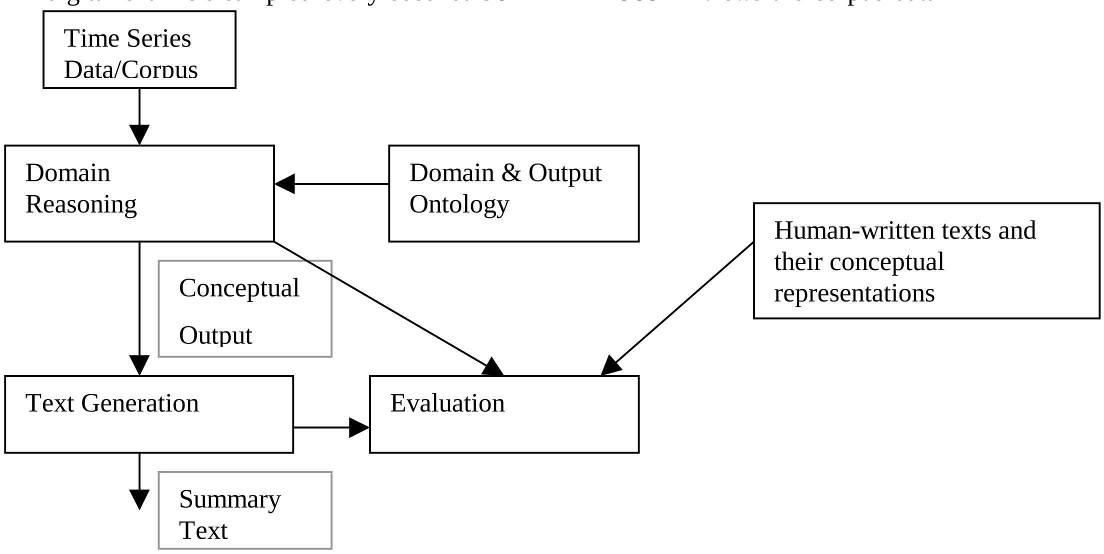


**Figure 2. Architecture of SUMTIME-MOUSAM**

#### 5.1.2 Conceptual Representation

The output of the domain reasoning component is a conceptual representation, not text. A later module converts this conceptual representation to text. The conceptual representation will be defined in output ontology (see below), perhaps at a low level. This conceptual representation gives us at least two advantages:

- It allows us to separately evaluate content and linguistic expression. We also derive conceptual representations from our corpus of human generated summary, and we compare computer summaries to human summaries.
- This representation gives us a convenient way of coding the corpus texts, which facilitates machine learning exercises to be carried out on the corpus.

At the moment we have conceptual representation for wind text and weather text. We define a “conceptual level” of a wind forecast to consist of a tuple with the following elements

- Time
- Wind speed lower range
- Wind speed upper range
- Wind direction (in degrees)
- Modifiers (gusts, shower, steady-change, gradual-change)

#### *5.1.2.1 Example*

1) Forecast period 1 from 26Dec2000_15.prn

50M: N 18-23 EASING 12-18 TONIGHT, GUSTS 40 IN SHOWERS

This is represented as 2 tuples

(1500, 18, 23, 0, nil)

(2100, 12, 18, 0, gusts+showers)

### 5.1.3 Domain Reasoning

This component is responsible for handling the domain reasoning to be performed on the data both while abstracting the numerical data into their conceptual representations and also while deriving new concepts from the input data. The domain reasoning referred to above could be a simple numerical computation such as computing total cloud cover from individual components of cloud, or a more complex inference such as determining the stability of an air mass.

### 5.1.4 Text Generation

This component is responsible for converting the output generated in the conceptual form into English text. Here the major issues are micro planning and lexical selection. As of today, this module is still not implemented. But like the domain reasoning, the SUMTIME-MOUSAM text generation module will be structured so that, new components can easily be inserted, and will have clear APIs.

### 5.1.5 Domain & Output Ontology

In SUMTIME, studies on corpus analysis and observations from KA sessions both suggest that the domain ontology, that is the specification of underlying concepts, underlying the overview may be quite different from the ontology underlying the actual output texts. The domain ontology includes concepts used by experts when reasoning about a domain (such as air masses or motivation), while the output ontology includes concepts useful for communicating information to the end user (such as wind speed, or longer life expectancy).

### 5.1.6 Output from the Initial Prototype

Figure 3 shows output from an initial prototype of SUMTIME-MOUSAM. Note that the text is more verbose than the human forecast text shown in Figure 1. We are currently working on improving the machine output based on comparisons with human-written texts.

```
06-24 GMT, 02-Mar 2001:
W 8-13, veering WNW in the mid-afternoon.

00-24 GMT, 03-Mar 2001:
WNW 8-13 soon NW 8-13, backing WSW in the early morning, then
SSW during the morning, becoming 18-23 around midday, veering
SW during the night, becoming 25-30 around midnight.

00-24 GMT, 04-Mar 2001:
SW 25-30, veering WSW in the early morning, then W 18-23
during the morning, becoming 8-13 in the early evening,
backing SSW during the night, then S 20-25 but increasing in
squally showers at times to 35 around midnight.
```

**Figure 3. Wind texts from SUMTIME-MOUSAM for Data shown in Table 1**

## 6. Summary and Future

An overall message emerging from the KA activities carried out so far in SUMTIME is that textual summaries are not just descriptions of the visual appearance of time-series data. Instead they are communicative artefacts in which the authors use their knowledge of the domain (overall weather situation, local weather variations, reliability of the data) and the user (especially typical user tasks and procedures) to craft short texts which will help users make decisions or perform tasks. And it is perhaps their incorporation of domain and user knowledge, which makes textual weather forecasts more valuable than a table or graph, as well as the fine control which text allows over exactly what information is and is not communicated. We plan to continue KA activities for eliciting more detailed knowledge about summarisation. SUMTIME-MOUSAM is being used to evaluate the observations from the KA. Subsequent stages of the project will include applying our time-series summarisation ideas to a new domain, and also hopefully performing a user-task evaluation of generated forecasts and gas turbine summaries, to see if they actually help real users perform real tasks.


## Acknowledgements

Many thanks to our collaborators at WNI/Oceanroutes and Intelligent Applications, especially Dave Selway, Rob Milne, and Jon Aylett; this work would not be possible without them! This project is supported by the UK Engineering and Physical Sciences Research Council (EPSRC), under grant GR/M76881.


## References


1. Boyd S (1998). TREND: a system for generating intelligent descriptions of time-series data. In *Proceedings of the IEEE International Conference on Intelligent Processing Systems (ICIPS-1998).*
2. Goldberg, E; Driedger, N; and Kittredge, R (1994). Using Natural-Language Processing to Produce Weather Reports. *IEEE Expert* **9**:45-53.

3. Hayes-Roth, F., Waterman, D., and Lenat, D. (Eds) (1983). *Building Expert Systems.* Addison-Wesley.
4. Lidija Iordanskaja, Myunghee Kim, Richard Kitregde, Benoit Lavoie, and Alian Polguere. 1992. Generation of extended bilingual statistical reports. In *Proceedings of the 14th International Conference on Computational Linguistics (COLING-1992)*, Volume 3, pages 1019-1023.
5. Eammon Keogh (1997). A fast and robust method for pattern matching in time-series data. In *Proceedings of WUSS-97*.
6. Karen Kukich. (1983). Design and implementation of a knowledge-based report generator. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics (ACL-1983),* pages 145-150.
7. Hunter, J and McIntosh, N (1999). Knowledge-Based event detection in complex time series data. *LNAI 1620*, pp. 271-280.
8. Igor Mel'cuk. (1988). *Dependency Syntax: Theory and Practice.* State University of New York Press, Albany, NY.
9. Reiter, E and Dale, R (2000). *Building Natural Language Generation Systems*. Cambridge University Press, 2000.
10. Reiter, E; Robertson, R; and Osman, L (2000). Knowledge Acquisition for Natural Language Generation. *Proceedings of the First International Conference on Natural Language Generation (INLG-2000)*, pages 217-224.
11. Scott,A; Clayton, J; and Gibson, E (1991). *A Practical Guide to Knowledge Acquisition*. Addison-Wesley
12. Shahar, Y and Cheng, C. (1999). Intelligent visualization and exploration of time-oriented clinical data. *Topics in Health Information Management*, 20:15-31.
13. Shahar, Y (1997). Framework for Knowledge-Based Temporal abstraction. *Artificial Intelligence* **90**:79-133.
14. Sripada, S, Reiter, E, Hunter, J and Yu, J. (2001). SUMTIME: Observation from KA for Weather Domain. Technical Report AUCS/TR0102. Dept. of Computing Science, University of Aberdeen.
15. Sripada, S, Reiter, E, Hunter, J and Yu, J. (2001). A Two-stage Model for Content Determination. In *Proceedings of ENLGW-2001 pp3-10*.