COM6115: Text Processing

Natural Language Generation

Chenghua Lin

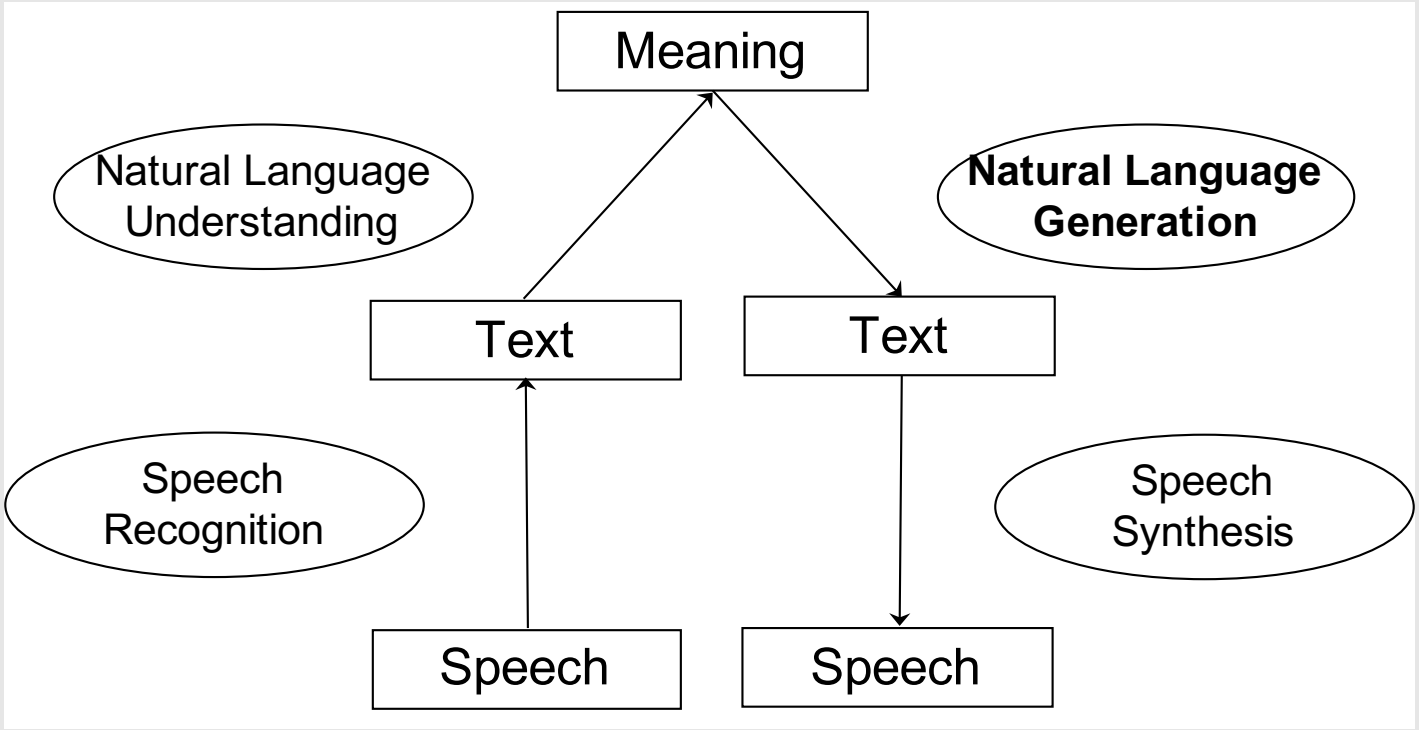
Department of Computer Science
University of Sheffield

What is NLG?

Background read: "*Building Natural Language Generation Systems*"
– Reiter and Dale

- Defined as the task of generating text (or speech) from non-linguistic input.
  - Data: numbers, RDF triples, etc
  - Output is documents, reports, explanations, help messages, and other kinds of texts
- Requires
  - Knowledge of language
  - Knowledge of the domain

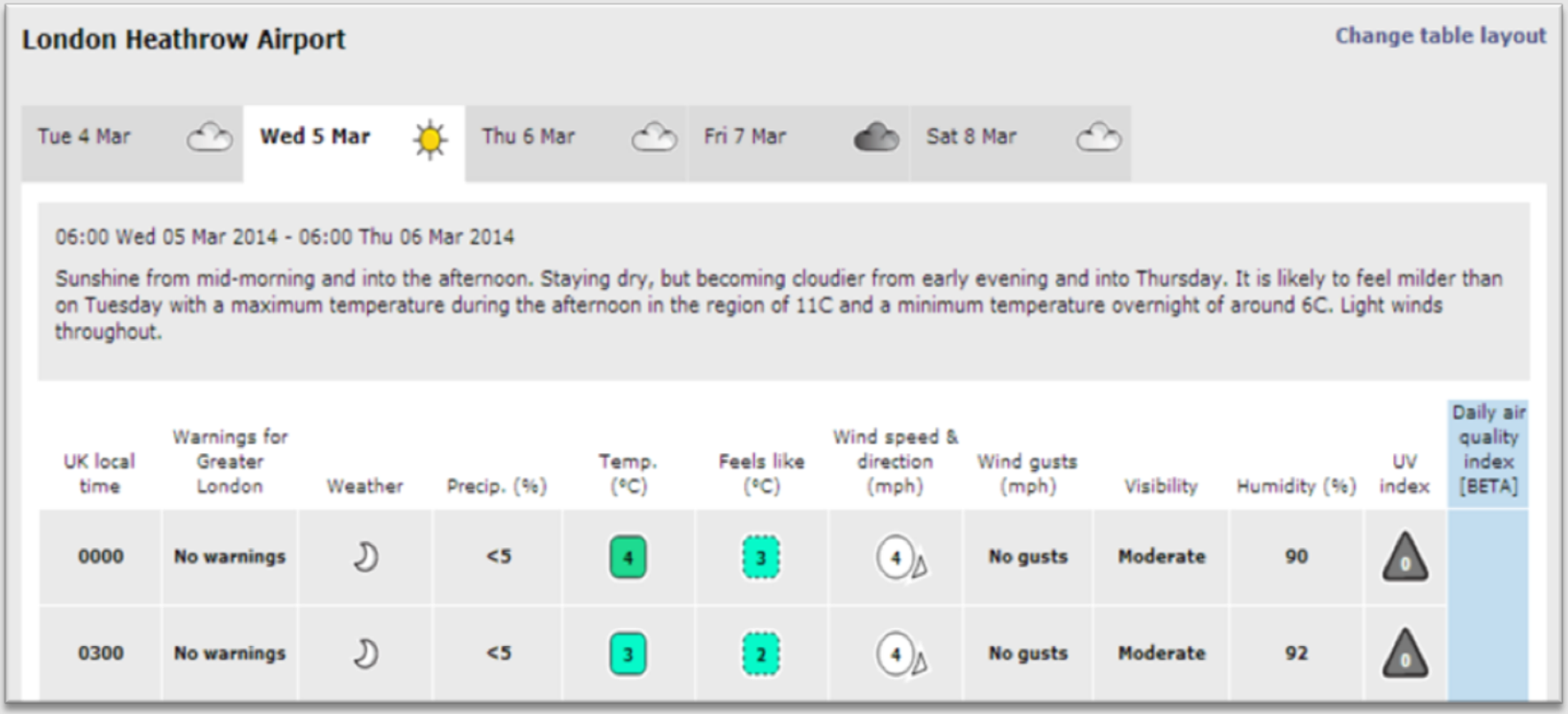
Language Technology



First Example: Weather Forecasts

- Input: numerical weather predictions
  - From supercomputer running a numerical weather simulation
- Output: textual weather forecast
  - Users prefer some NLG texts over human texts!
  - More consistent, better word choice

Simple example: Point weather forecast



London Heathrow Airport

Change table layout

Tue 4 Mar | **Wed 5 Mar** | Thu 6 Mar | Fri 7 Mar | Sat 8 Mar

06:00 Wed 05 Mar 2014 - 06:00 Thu 06 Mar 2014

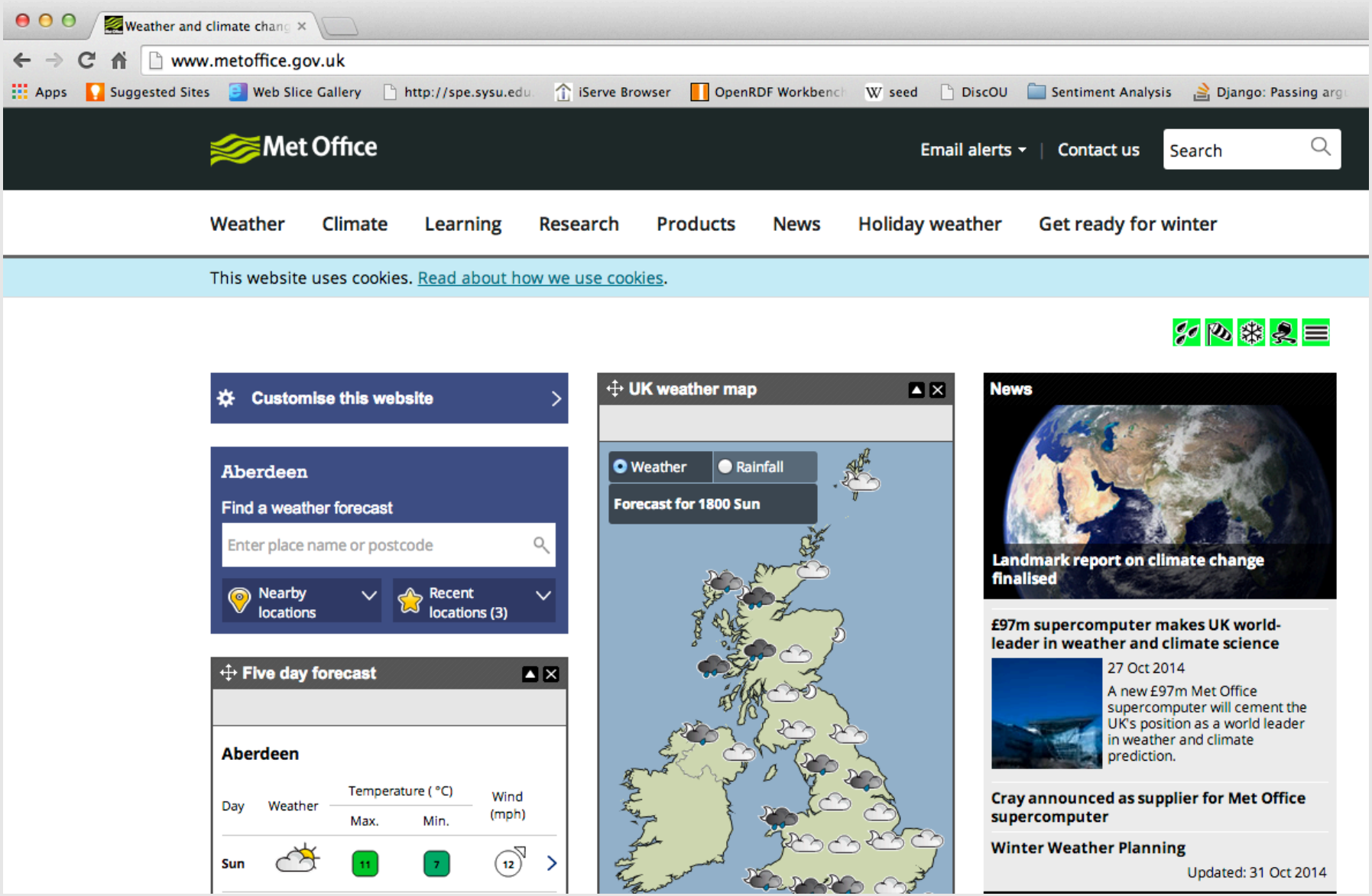
Sunshine from mid-morning and into the afternoon. Staying dry, but becoming cloudier from early evening and into Thursday. It is likely to feel milder than on Tuesday with a maximum temperature during the afternoon in the region of 11C and a minimum temperature overnight of around 6C. Light winds throughout.

UK local time	Warnings for Greater London	Weather	Precip. (%)	Temp. (°C)	Feels like (°C)	Wind speed & direction (mph)	Wind gusts (mph)	Visibility	Humidity (%)	UV index	Daily air quality index [BETA]
0000	No warnings		<5	4	3	4	No gusts	Moderate	90	0	
0300	No warnings		<5	3	2	4	No gusts	Moderate	92	0	

Example 1: Met Office NLG System

- Input:
  - Weather prediction data of temperature, wind speed and direction, precipitation and visibility, etc.;
  - Daily summary weather prediction data of average daily and nightly values for parameters as above; and
  - Seasonal averages (lows, highs and mean) for temperature.
- Output: weather forecast texts
- NLG system vs. manual report writing
  - Volume: satellite cloud data is gathered at a speed of 158M per second
  - Time: NLG system ($<$ 30 secs) vs. human expert (hours)

Example 1: Met Office NLG System



Data Input

```
96,122,1,5,2.00,200,-14.41,-3.668,-1.431,.345,1023,15.41,15.82,20.07,-11.1,-2.878,104.2,.28,153.6,53.19,0,16.26
96,122,1,5,2.25,215,-10.72,-3.241,-1.35,.152,1023,15.3,15.78,20.07,-11.42,-2.762,105,.208,98.2,.822,0,17.05
96,122,1,5,2.50,230,-8.37,-1.282,-.904,2.15,1022,15.3,15.71,20.05,-11.66,-3.206,104.4,.2,141.6,42.96,0,17.7
96,122,1,5,2.75,245,-12.81,-2.11,-1.067,2.119,1022,15.33,15.79,19.99,-11.15,-3.093,104.8,.2,186.5,11.32,0,17.81
96,122,1,5,3.00,300,-13.68,-3,-1.35,1.075,1022,15.36,15.79,19.96,-10.63,-3.005,104.6,.402,285.8,61.45,0,18.47
96,122,1,5,3.25,315,-10.2,-2.457,-1.13,-.73,1022,15.32,15.66,19.92,-11.17,-3.263,103.6,.304,354.7,36.29,0,19.03
96,122,1,5,3.50,330,-9.33,-1.353,-.942,.902,1022,15.21,15.62,19.9,-10.95,-2.903,104.3,.313,302.2,34.69,0,19.16
96,122,1,5,3.75,345,-7.29,-.285,-.76,2.048,1022,15.24,15.63,19.87,-10.68,-3.27,104,.252,313,29.7,0,19.61
96,122,1,5,4.00,400,-6.822,-.365,-.653,1.531,1022,15.25,15.63,19.83,-9.93,-3.316,104,.331,274.2,52.98,0,20.42
96,122,1,5,4.25,415,-8.78,-.65,-.747,1.602,1023,15.35,15.66,19.79,-9.77,-2.656,103.3,.253,247.7,10.99,0,21.08
96,122,1,5,4.50,430,-8.73,-.641,-.741,1.785,1023,15.46,15.81,19.75,-9.16,-2.782,103.7,.2,295,29.15,0,21.3
96,122,1,5,4.75,445,-11.45,-2.671,-1.03,-.456,1022,15.46,15.82,19.74,-8.81,-2.464,103.7,.2,355.3,23.98,0,21.65
96,122,1,5,5.00,500,-13.12,-4.3,-1.306,-1.359,1022,15.42,15.75,19.76,-9.39,-2.49,103.4,.2,20.67,.188,0,21.83
96,122,1,5,5.25,515,-13.62,-4.621,-1.344,-.842,1022,15.32,15.67,19.81,-9.47,-2.703,103.7,.2,20.65,.183,0,21.98
96,122,1,5,5.50,530,-13.8,-3.534,-1.325,.943,1022,15.23,15.61,19.86,-10.92,-3.384,103.9,.2,20.65,.183,0,22.14
96,122,1,5,5.75,545,-14.7,-3.748,-1.419,.385,1022,15.06,15.47,19.9,-11.62,-2.868,104.4,.2,341.6,18.6,0,22.36
96,122,1,5,6.00,600,-13.61,-2.315,-1.287,2.038,1022,14.98,15.42,19.9,-12.37,-3.092,104.7,.2,298.6,5.173,0,22.54
96,122,1,5,6.25,615,-14,-2.894,-1.293,.669,1022,14.92,15.36,19.88,-12.48,-3.808,104.7,.591,320.3,21.07,0,22.87
```

Example 1: Met Office NLG System

Forecast summary

Regional | UK 5 days | UK 6-30 days

Regional forecast for Grampian

Rain edging northwards during Monday morning. Becoming drier later.

This Evening and Tonight:

Dry this evening and for most of the night with some clear spells. Becoming cold with a few mist or fog patches forming as winds fall light. Showers will spread up into southern Aberdeenshire by morning. Minimum Temperature 2C.

Monday:

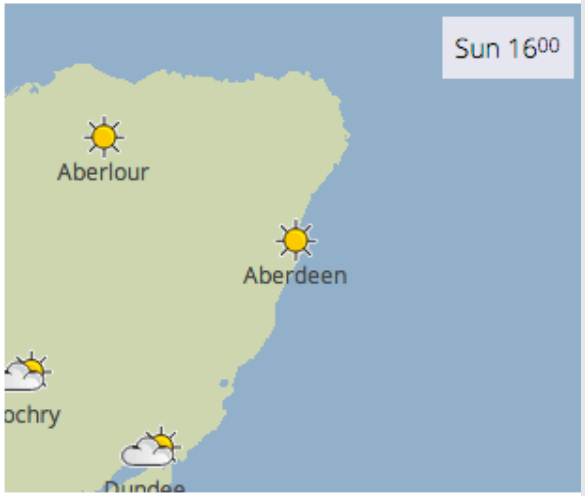
Dry, bright start, soon becoming cloudy. Showers or longer spells of over southern Aberdeenshire will edge northwards to all parts during the morning. Becoming drier from south during afternoon. Maximum Temperature 10C.

Outlook for Tuesday to Thursday:

Frequent showers Tuesday, wintry on hills later, as winds turn more northerly. Some showers early Wednesday, hill snow, then dry and bright. Frost overnight then bright start Thursday, rain later.

Issued at: 1600 on Sun 02 Nov 2014

Weather map



Weather forecast map for Aberdeen

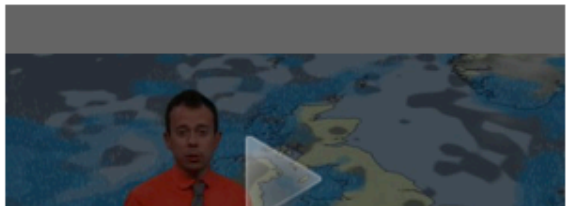
Location Details

Aberdeen

Location: 57.1498, -2.0927
Altitude: 19m above mean sea level

Video forecast

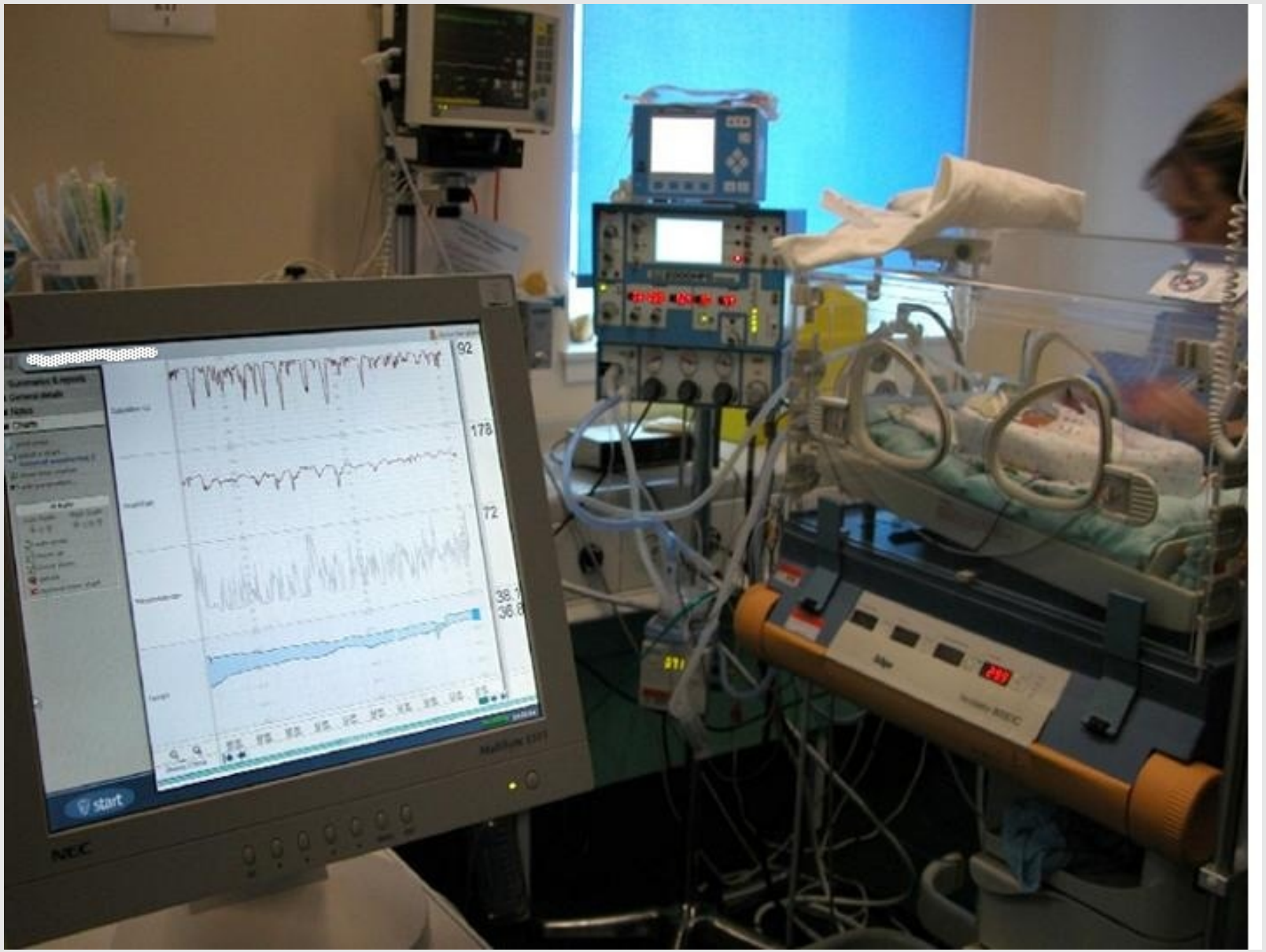
Sunday's Forecast



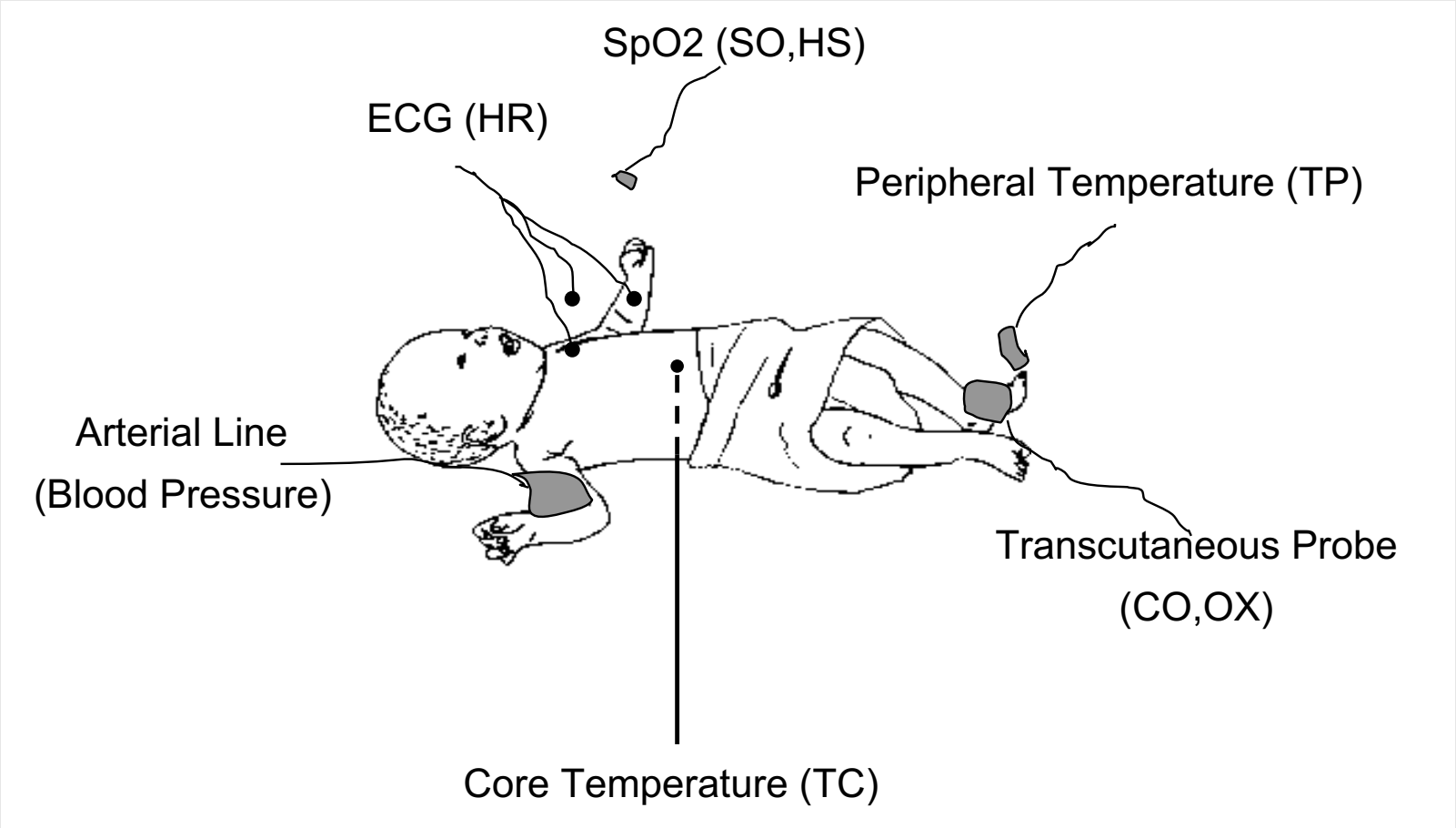
Example 2: BabyTalk

- Goal: Summarise clinical data about premature babies in neonatal ICU
- Input: sensor data; records of actions/observations by medical staff
- Output: multi-para texts, summarise
  - BT45: 45 mins data, for doctors
  - BT-Nurse: 12 hrs data, for nurses
  - BT-Family: 24 hrs data, for parents

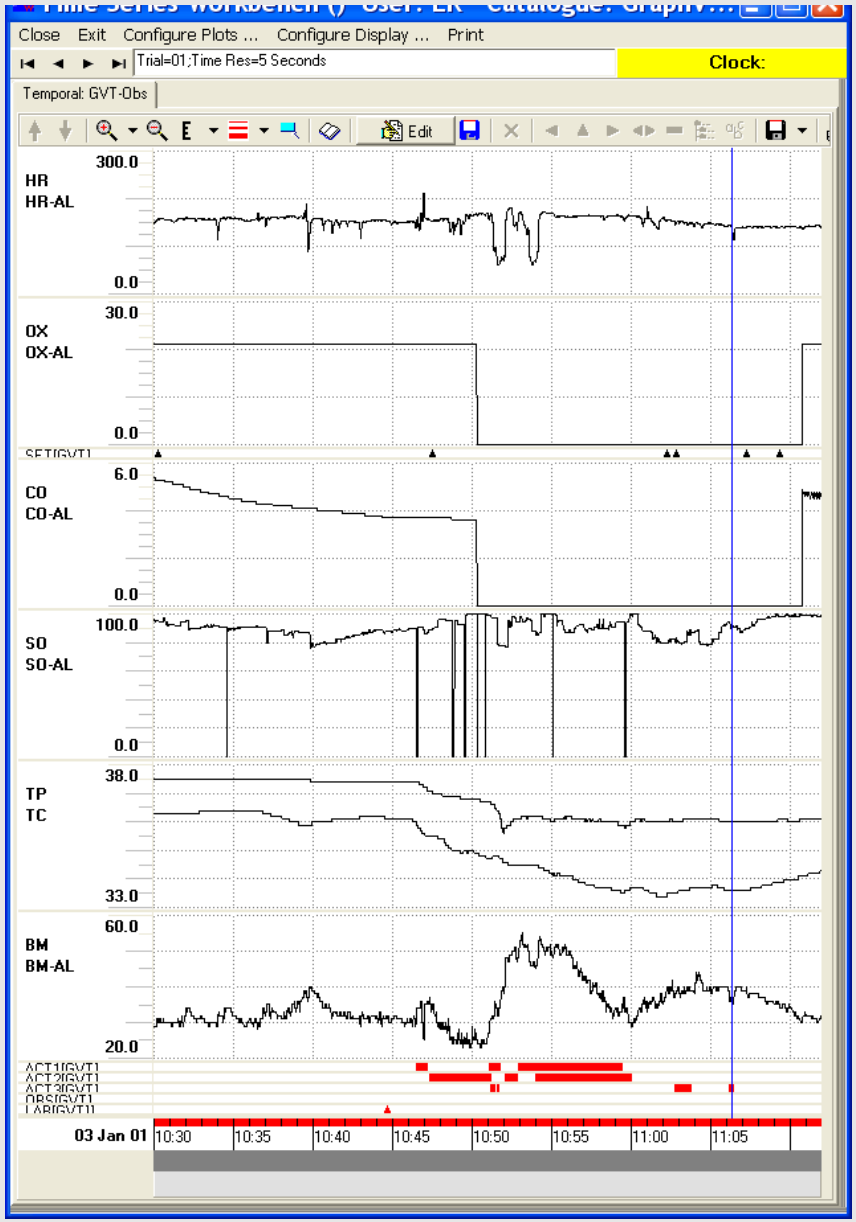
Neonatal ICU



Baby Monitoring



Input: Sensor Data



Input: Action Records

FullDescriptor	Time
SETTING;VENTILATOR;FiO2 (36%)	10.30
MEDICATION;Morphine	10.44
ACTION;CARE;TURN/CHANGE POSITION;SUPINE	10.46-10.47
ACTION;RESPIRATION;HAND-BAG BABY	10.47-10.51
SETTING;VENTILATOR;FiO2 (60%)	10.47
ACTION;RESPIRATION;INTUBATE	10.51-10.52

BT45 text (extract)

Computer-generated text

- By 11:00 the baby had been hand-bagged a number of times causing 2 successive bradycardias. She was successfully reintubated after 2 attempts. The baby was sucked out twice.
  At 11:02 FIO2 was raised to 79%.

BT-Nurse text (extract)

Respiratory Support

Current Status
Currently, the baby is on CMV in 27 % O2. Vent RR is 55 breaths per minute. Pressures are 20/4 cms H2O. Tidal volume is 1.5.

SaO2 is variable within the acceptable range and there have been some desaturations.

...

Events During the Shift
A blood gas was taken at around 19:45. Parameters were acceptable. pH was 7.18. CO2 was 7.71 kPa. BE was -4.8 mmol/L.

...

BT-Family text (extract)

John was in intensive care. He was stable during the day and night. Since last week, his weight increased from 860 grams (1 lb 14 oz) to 1113 grams (2 lb 7 oz). He was nursed in an incubator.

Yesterday, John was on a ventilator. The mode of ventilation is Bilevel Positive Airway Pressure (BiPAP) Ventilation. This machine helps to provide the support that enables him to breathe more comfortably. Since last week, his inspired Oxygen (FiO2) was lowered from 56 % to 21 % (which is the same as normal air). This is a positive development for your child.

During the day, Nurse Johnson looked after your baby. Nurse Stevens cared for your baby during the night.

Example 3: Summary of the NBA Game statistics

- Goal: Generate documents to summarise the game statistics of NBA.
  - ◇ In addition to capturing the writing style, a generation system should select record content, express it clearly, and order it appropriately.
- Input: structured data table capturing statistics info of games
- Output: a game summary document

Exam 3: NBA - Training Corpus

TEAM	WIN	LOSS	PTS	FG_PCT	RB	AS ...
Heat	11	12	103	49	47	27
Hawks	7	15	95	43	33	20

PLAYER	AS	RB	PT	FG	FGA	CITY ...
Tyler Johnson	5	2	27	8	16	Miami
Dwight Howard	4	17	23	9	11	Atlanta
Paul Millsap	2	9	21	8	12	Atlanta
Goran Dragic	4	2	21	8	17	Miami
Wayne Ellington	2	3	19	7	15	Miami
Dennis Schroder	7	4	17	8	15	Atlanta
Rodney McGruder	5	5	11	3	8	Miami
Thabo Sefolosha	5	5	10	5	11	Atlanta
Kyle Korver	5	3	9	3	9	Atlanta
...						

The Atlanta Hawks defeated the Miami Heat , 103 - 95 , at Philips Arena on Wednesday . Atlanta was in desperate need of a win and they were able to take care of a shorthanded Miami team here . Defense was key for the Hawks , as they held the Heat to 42 percent shooting and forced them to commit 16 turnovers . Atlanta also dominated in the paint , winning the rebounding battle , 47 - 34 , and outscoring them in the paint 58 - 26.The Hawks shot 49 percent from the field and assisted on 27 of their 43 made baskets . This was a near wire - to - wire win for the Hawks , as Miami held just one lead in the first five minutes . Miami (7 - 15) are as beat - up as anyone right now and it 's taking a toll on the heavily used starters . Hassan Whiteside really struggled in this game , as he amassed eight points , 12 rebounds and one blocks on 4 - of - 12 shooting ...

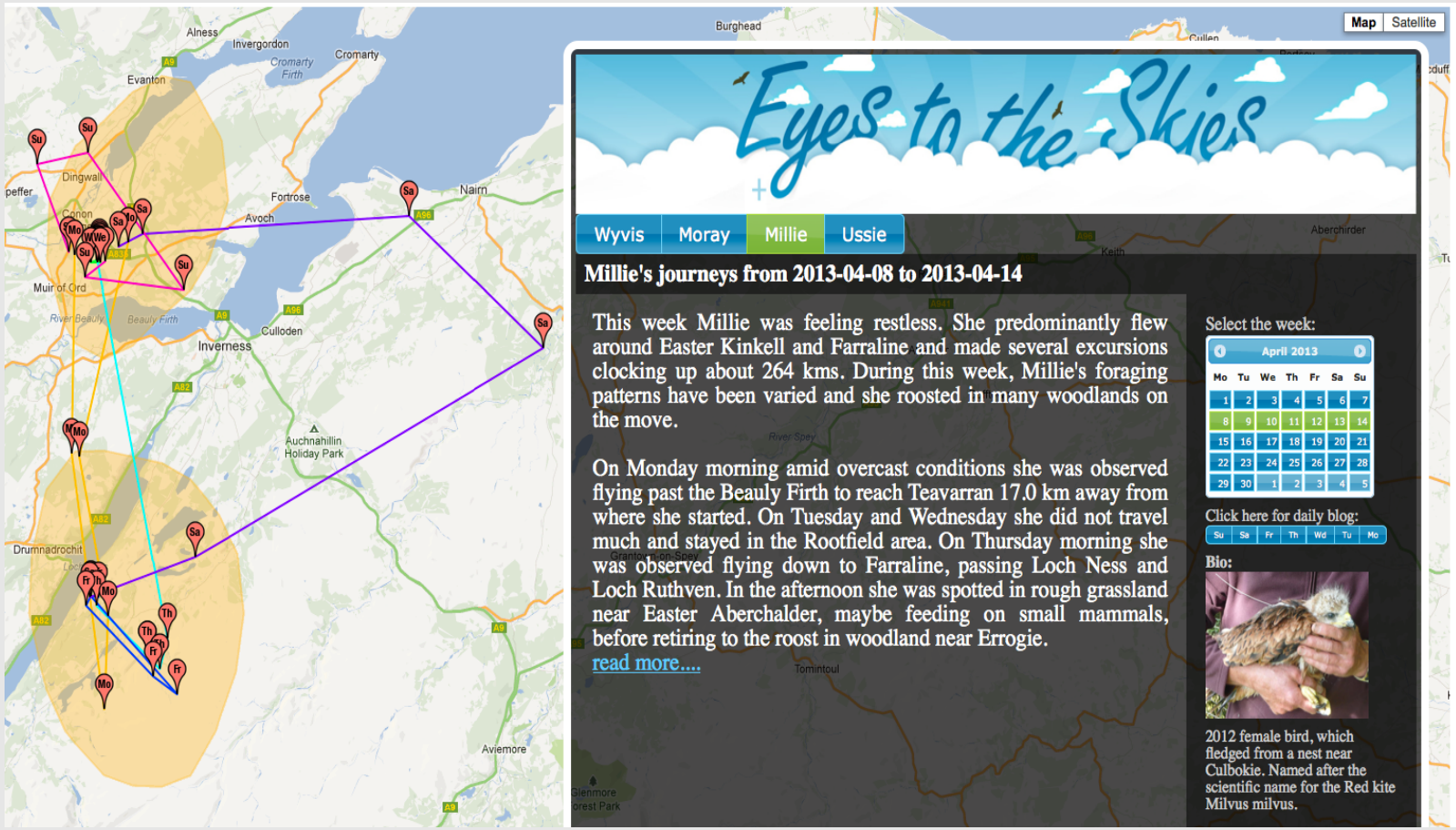
"An example data-record and document pair from the ROTOWIRE dataset. We show a subset of the game's records (there are 628 in total), and a selection from the gold document. The document mentions only a select subset of the records, but may express them in a complicated manner."

Exam 3: NBA - Generation Example

The Utah Jazz (38 - 26) defeated the Houston Rockets (38 - 26) 117 - 91 on Wednesday at Energy Solutions Arena in Salt Lake City . The Jazz got out to a quick start in this one , out - scoring the Rockets 31 - 15 in the first quarter alone . Along with the quick start , the Rockets were the superior shooters in this game , going 54 percent from the field and 43 percent from the three - point line , while the Jazz went 38 percent from the floor and a meager 19 percent from deep . The Rockets were able to out - rebound the Rockets 49 - 49 , giving them just enough of an advantage to secure the victory in front of their home crowd . The Jazz were led by the duo of Derrick Favors and James Harden . Favors went 2 - for - 6 from the field and 0 - for - 1 from the three - point line to score a game - high of 15 points , while also adding four rebounds and four assists ....

Example document generated by the Conditional Copy system with a beam of size 5. Text that accurately reflects a record in the associated box- or line-score is highlighted in blue, and erroneous text is highlighted in red.

Exam 3: Blogging Birds



Blogging Birds

Wyvis's Journeys (20/09/12 to 26/09/12) , Female bird born in 2012

DOW	Hour	Habitat	Significant Weather	Temp (C)	Visibility (m)	Wind Speed (mph)	Location	Features	Distance Flown	Other Kites
Friday	8	coniferous woodland	overcast	13.0	24000	3.0	East Croachy	Loch Ruthven,	0.0	
	10	rough grassland	heavy rain	13.9	5000	2.0	Torness	Loch Ruthven,	4.0	
	12	rough grassland	heavy rain	16.0	3600	1.0	Torness	Loch Ruthven,	2.0	
	14	rough grassland	heavy rain	16.0	3600	1.0	Torness	Loch Ruthven,	2.0	Merida
	16	improved grassland	overcast	18.4	45000	5.0	Torness		3.0	

Other NLG projects

- Automatic journalism
  - http://www.bbc.co.uk/news/technology-34204052
- Assistive technology: help people with learning disabilities, blind people, deaf people, ...
- Education: computerised tutoring systems, feedback on assessments
- Image labelling
- Agent and dialogue systems (e.g., Siri, Cortana)
- Etc, etc

Image labelling example

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image
 A person riding a motorcycle on a dirt road.	 Two dogs play in the grass.	 A skateboarder does a trick on a ramp.	 A dog is jumping to catch a frisbee.
 A group of young people playing a game of frisbee.	 Two hockey players are fighting over the puck.	 A little girl in a pink hat is blowing bubbles.	 A refrigerator filled with lots of food and drinks.
 A herd of elephants walking across a dry grass field.	 A close up of a cat laying on a couch.	 A red motorcycle parked on the side of the road.	 A yellow school bus parked in a parking lot.

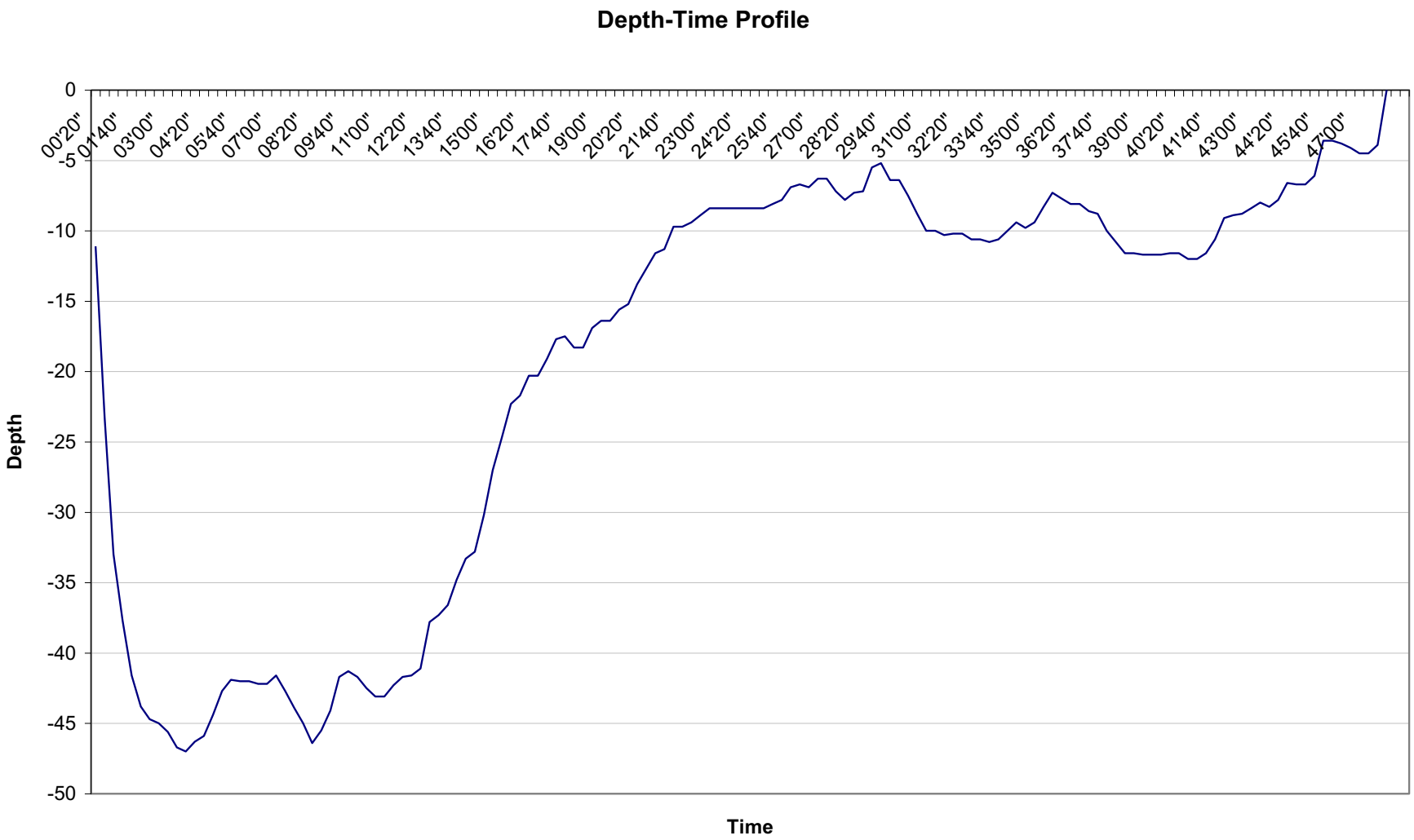
How do NLG Systems Work?

- Usually three stages
  - Not including data analysis
- **Document planning**: decide on content and structure of text
- **Microplanning**: decide how to linguistically express text (which words, sentences, etc to use)
- **Realisation**: grammatical details
  - E.g. *children* vs. *childs*, *an apple* vs. *a apple*

Scubatext example

- Demo system for scuba divers
- Input is *dive computer data*
  - Depth-time profile of scuba dive
- Output is feedback to diver
  - Mistakes, what to do better next time
  - Encouragement of things done well

Scuba - input


Depth-Time Profile
Depth
0
-5
-10
-15
-20
-25
-30
-35
-40
-45
-50
00'20"
01'40"
03'00"
04'20"
05'40"
07'00"
08'20"
09'40"
11'00"
12'20"
13'40"
15'00"
16'20"
17'40"
19'00"
20'20"
21'40"
23'00"
24'20"
25'40"
27'00"
28'20"
29'40"
31'00"
32'20"
33'40"
35'00"
36'20"
37'40"
39'00"
40'20"
41'40"
43'00"
44'20"
45'40"
47'00"
Time

Scuba – output

- "Risky dive with some minor problems. Because your bottom time of 12 min exceeds no-stop limit by 4 min this dive is risky. But you performed the ascent well. Your buoyancy control in the bottom zone was poor as indicated by 'saw tooth' patterns."

Scuba: data analytics

- Look for trends and patterns in data
  - Trends: e.g., depth increases fairly steadily over first 3 minutes
  - Patterns: e.g., sawtooth between 3 and 15 minutes
- Will not further discuss here

Document Planning

- Content selection: of the zillions of things I could say, which should I say?
  - Depends on what is important
  - What makes good narrative
  - What is easy to say
- Structure: How should I organise this content as a text?
  - What order do I say things in?
  - Rhetorical structure?

Scuba: content

- Probably focus on patterns indicating dangerous activities
  - E.g., most important thing to mention
- How much should we say about these?
  - Detail? Explanations?
- Encourage/praise good diving
  - Positive feedback is important

Scuba: structure

- Mention most dangerous thing first?
  - Or should we just order by time?
  - Start with overview?
- Linking words (cue phrases)
  - Also, but, because, ...

Microplanning

- Lexical/syntactic choice: Which words and linguistic structures to use?
- Aggregation: How should information be distributed across sentences and paras
- Reference: How should the text refer to objects and entities?

SCUBA: microplanning

- Lexical/syntactic choice:
  - *risky* vs. *dangerous* vs. *unwise* vs. ...
  - *performed the ascent* vs. *ascended* vs ...
  - *12 min* vs. *720 sec* vs. *714.56 sec*
- Aggregation: 1 sentence or 2 sentences?
  - "Because your bottom time of 12 min exceeds no-stop limit by 4 min this dive is risky, but you performed the ascent well."

Scuba: Microplanning

- Aggregation (continued)
  - Phrase merging
    - "Your first ascent was fine. Your second ascent was fine" vs.
    - "Your first and second ascents were fine."
  - Reference
    - Your ascent vs.
    - Your first ascent vs.
    - Your ascent from 33m at 3 min

Realisation

- Grammars (linguistic): Form legal English sentences based on decisions made in previous stages
  - Obey sub-languages, i.e., language of a restricted domain, particularly a technical domain.
  - genre constraints, e.g., scientific writing vs. social media text, etc.
- Structure: Form legal HTML, RTF, or whatever output format is desired

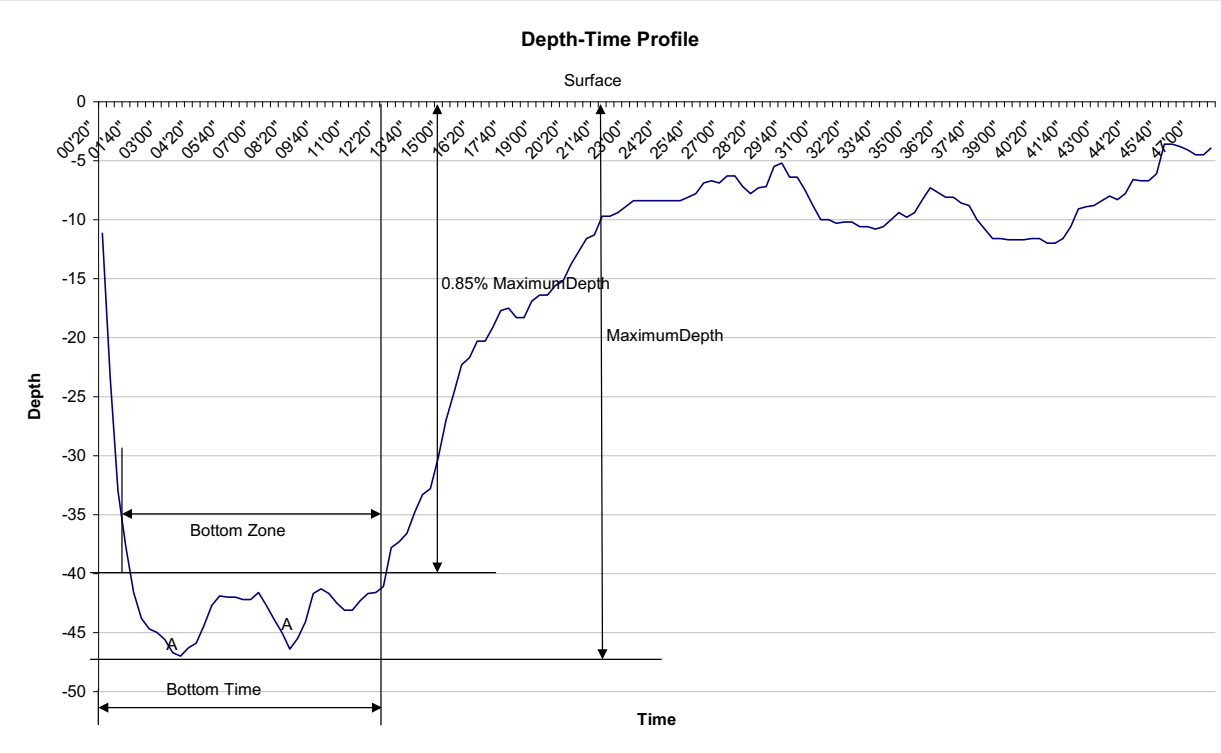
Scuba: Realisation

- Simple linguistic processing
  - Capitalise first word of sentence
  - Subject-verb agreement
    - Your first ascent was fine
    - Your first and second ascents were fine
- Structure
  - Inserting line breaks in text
  - Add HTML markups, eg, <P>

Multimodal NLG

- Speech output
- Text and visualisations
  - Produce separately, OR
  - Tight integration
    - E.g., text refers to graphic, OR
    - graphs has text annotations

Combined (Preferred)



Risky dive with some minor problems. Because your bottom time of 12.0min exceeds no-stop limit by 4.0min this dive is risky. But you performed the ascent well. Your buoyanc control in the bottom zone was poor as indicated by 'saw tooth' patterns marked 'A' on the depth-time profile.

Building NLG Systems

- Knowledge and corpus analysis
- Evaluation

Building NLG Systems: Knowledge

- Need knowledge
  - Which patterns most important?
  - What order to use?
  - Which words to use?
  - When to merge phrases?
  - Etc.
- Where does this come from?

Knowledge Sources

- Imitate a *corpus* of human-written texts
  - Most straightforward
  - Manually examine
  - Use learning if corpus is large enough
- Ask domain experts
  - Experts bad at explaining what they do
  - Better at critiquing what system does
- Experiments with users
  - Very nice in principle, but a lot of work

Scuba: Corpus

- See which patterns humans mention in the corpus, and have the system mention these
- See the words used by humans, and have the system use these as well
- etc.

Evaluation

- Does system help people?
  - Do divers dive more safely when they use Scuba NLG system
- Do people like the texts
  - Do divers consider Scuba to be useful?
- Comparison to human texts
  - Are Scuba texts similar to corpus texts

NLG vs NLP

- Producing rather than understanding language
- Focus on content and AI issues as well as linguistic issues
- Increasing uptake of statistical and deep learning techniques