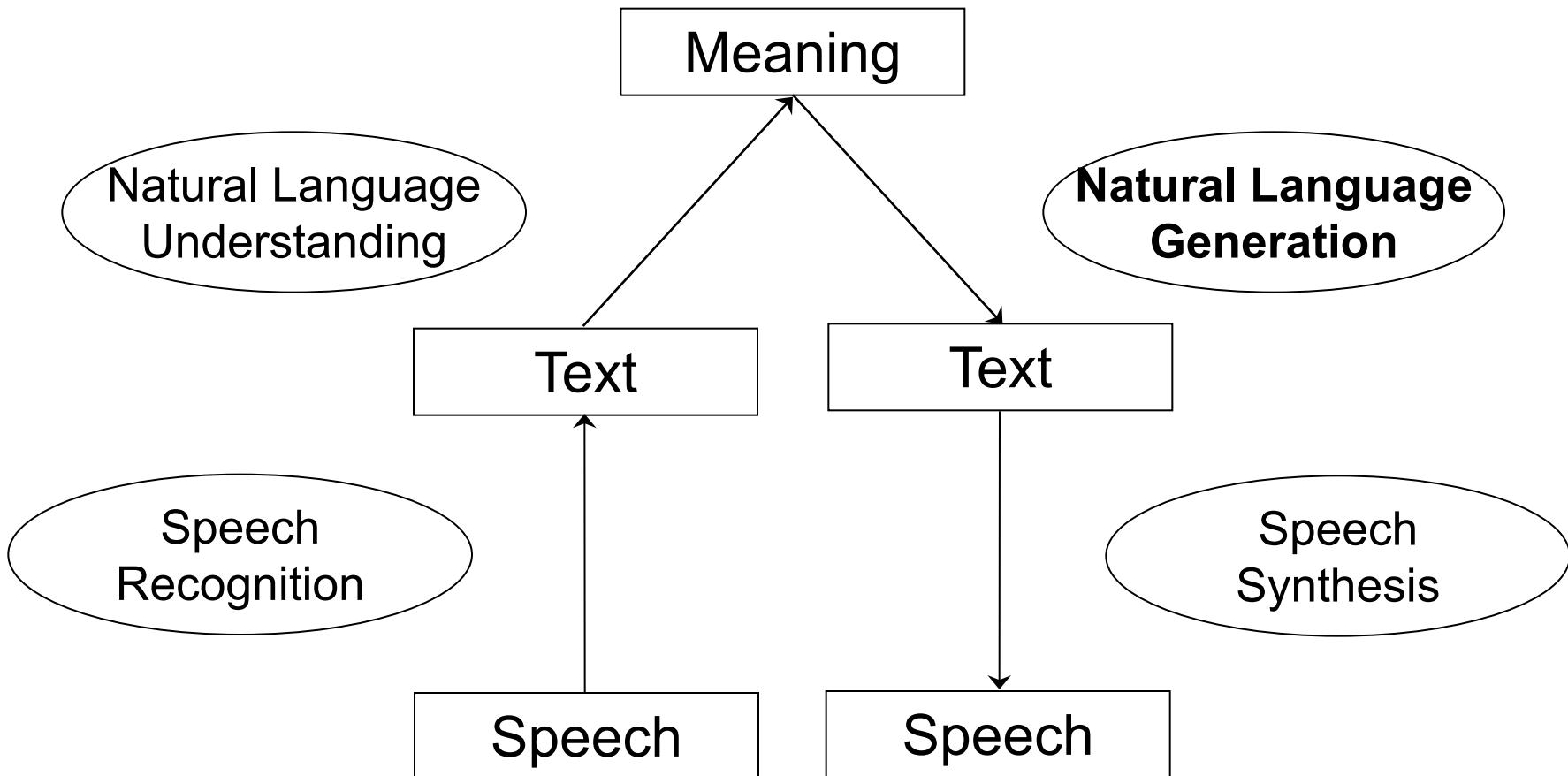

CS4025: Natural Language Generation

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

What is NLG?

- Software which produces language (words and sentences) from data
 - » 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				No gusts	Moderate	90		
0300	No warnings		<5				No gusts	Moderate	92		

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)

Weather and climate change

www.metoffice.gov.uk

Apps Suggested Sites Web Slice Gallery http://spe.sysu.edu.iServe Browser OpenRDF Workbench seed DiscOU Sentiment Analysis Django: Passing arg

Met Office

Email alerts | Contact us Search

Weather Climate Learning Research Products News Holiday weather Get ready for winter

This website uses cookies. [Read about how we use cookies.](#)

Aberdeen
Find a weather forecast
Enter place name or postcode 
 

Five day forecast
Aberdeen
Day Weather Temperature (°C) Wind (mph)
Max. Min.
Sun  11 7 12 >

UK weather map
Weather Rainfall Forecast for 1800 Sun 

News

Landmark report on climate change finalised
£97m supercomputer makes UK world-leader in weather and climate science 
27 Oct 2014 A new £97m Met Office supercomputer will cement the UK's position as a world leader in weather and climate prediction.
Cray announced as supplier for Met Office supercomputer
Winter Weather Planning
Updated: 31 Oct 2014

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

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.

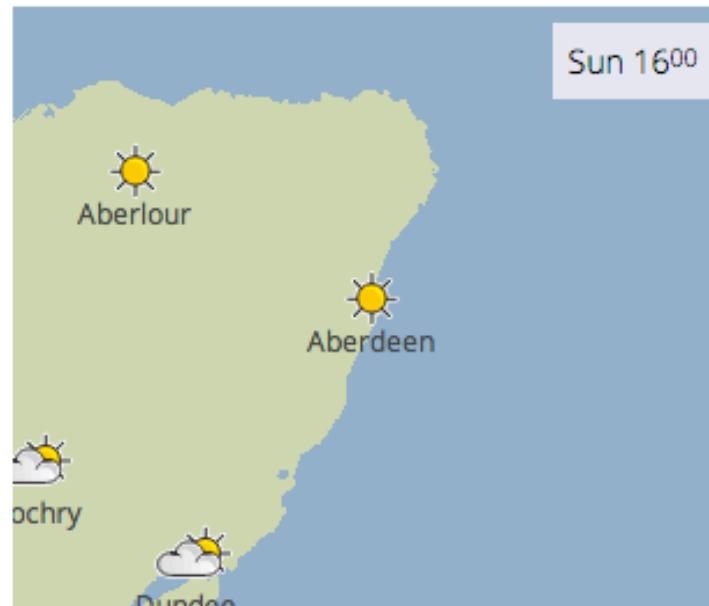
Monday:

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.

Weather map



Weather forecast map for Aberdeen

Location Details

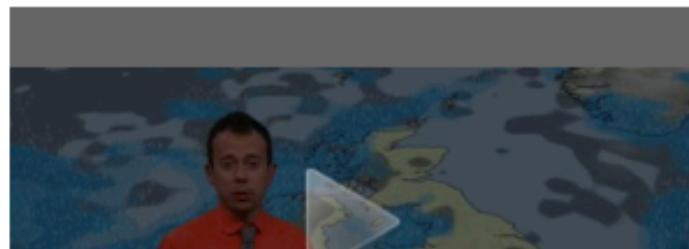
Aberdeen

Location: 57.1498, -2.0927

Altitude: 19m above mean sea level

Video forecast

Sunday's Forecast



Complex example: road maintenance

- Forecasts for gritting and other winter road maintenance procedures
- Input is 15 parameters over space and time
 - » Temperature, wind speed, rain, etc
 - » Over thousands of points on a grid
 - » Over 24 hours (20-min interval)

Points



Generated Text

Overview Road surface temperatures will reach marginal levels on most routes from this evening until tomorrow morning.

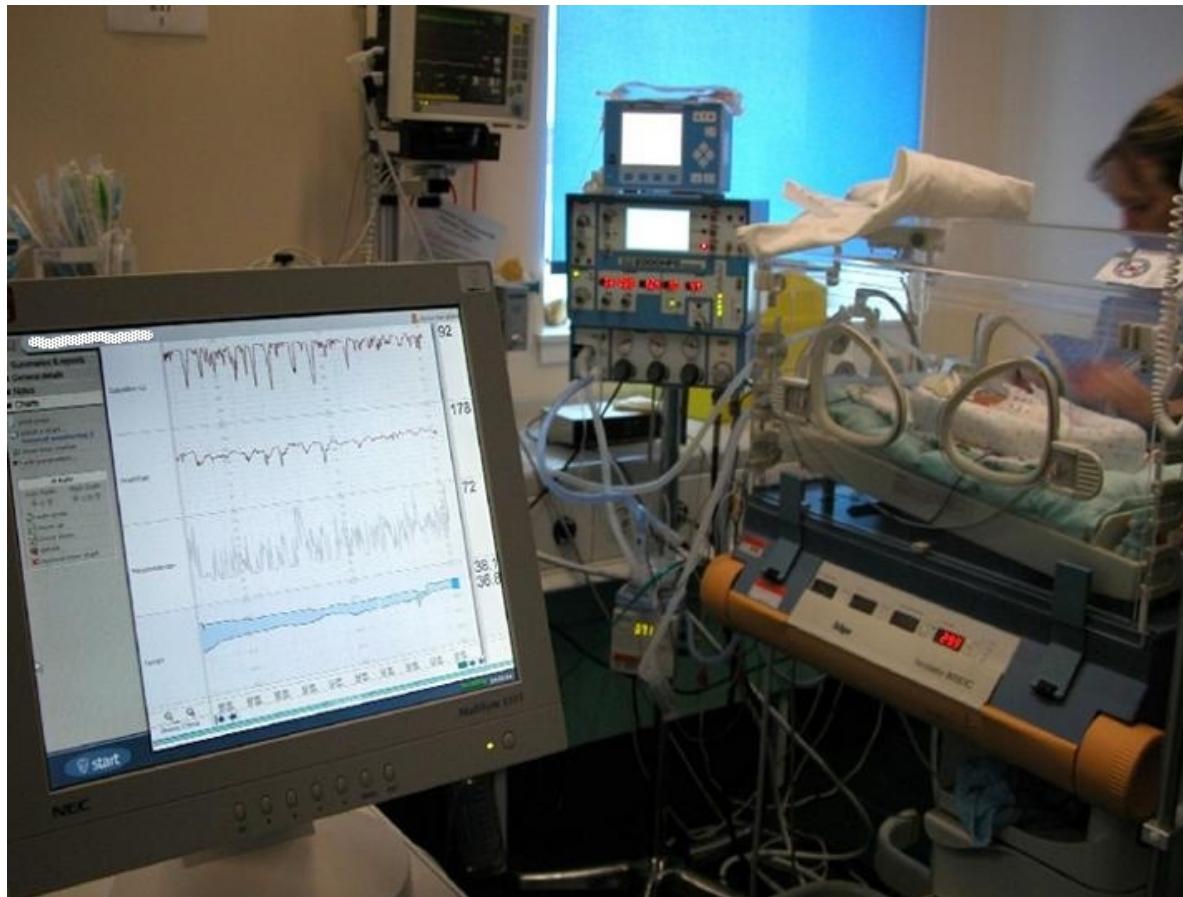
Wind (mph) NW 10-20 gusts 30-35 for a time during the afternoon and evening in some southwestern places, veering NNW then backing NW and easing 5-10 tomorrow morning.

Weather Light rain will affect all routes this afternoon, clearing by 17:00. Fog will affect some central and southern routes after midnight until early morning and light rain will return to all routes. Road surface temperatures will fall slowly during this afternoon until tonight, reaching marginal levels in some places above 200M by 17:00.

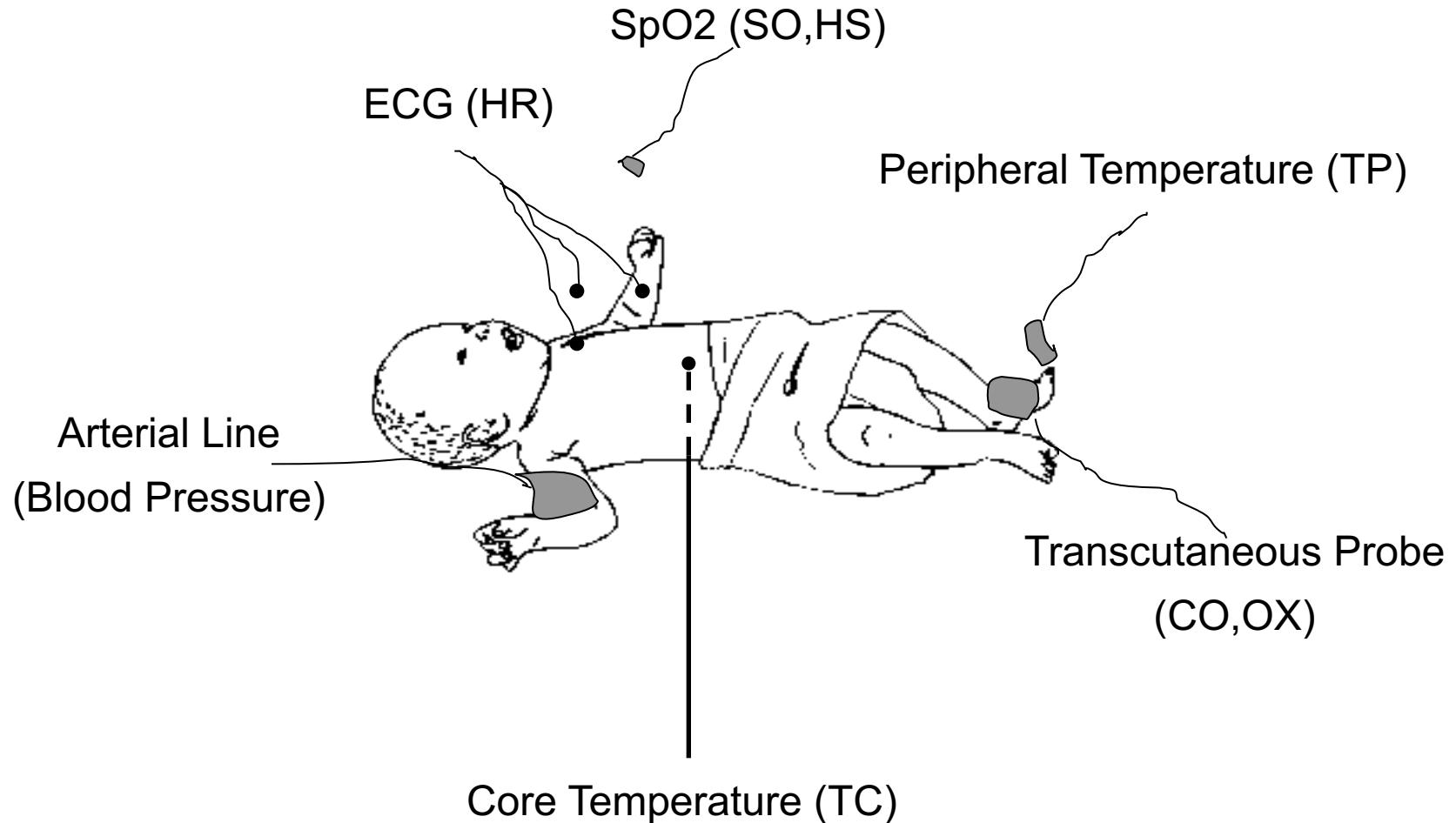
Example 2: BabyTalk

- 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

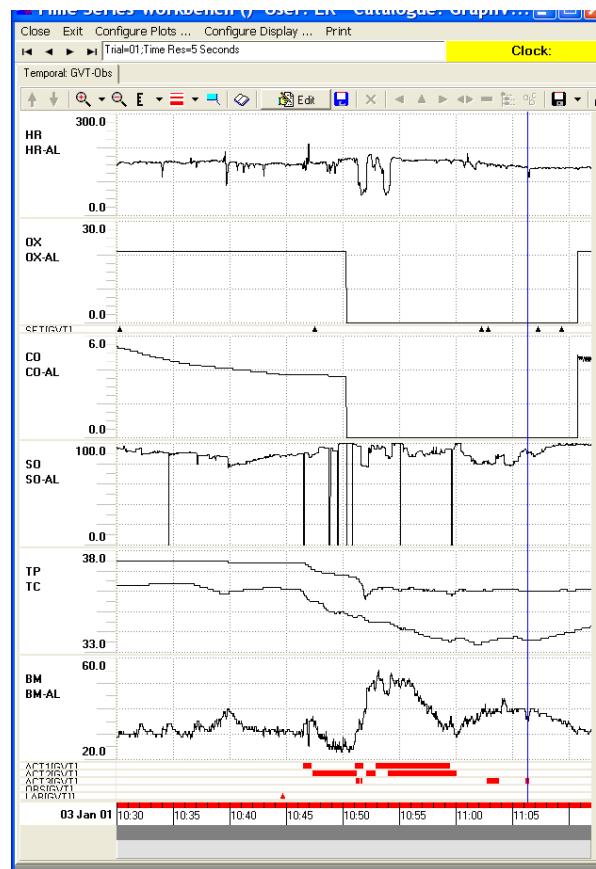
Neonatal ICU



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 texts (extract)

Computer-generated text

- By 11:00 the baby had been hand-bagged a number of times causing 2 successive bradycardias. She was successfully re-intubated 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.

Exam 3: Blogging Birds

The image consists of two parts. On the left is a map of the Scottish Highlands, specifically the Moray Firth and surrounding areas. It shows several locations marked with red pins and labeled with abbreviations: Su, Mo, W, We, Fr, Th, Sa, and Mo. A large yellow oval highlights a cluster of points around Dingwall and Inverness. On the right is a screenshot of a website titled "Eyes to the Skies". The header features a blue sky with clouds and the text "Eyes to the Skies". Below the header, there are tabs for "Wyvis", "Moray", "Millie", and "Ussie", with "Millie" being the active tab. A section titled "Millie's journeys from 2013-04-08 to 2013-04-14" contains text about Millie's movements and a "read more...." link. To the right of this is a "Select the week:" calendar for April 2013, showing dates from 1 to 30. Below the calendar is a "Click here for daily blog:" link and a "Bio:" section featuring a photograph of a bird of prey.

This week Millie was feeling restless. She predominantly flew around Easter Kinkell and Farraline and made several excursions clocking up about 264 kms. During this week, Millie's foraging patterns have been varied and she roosted in many woodlands on the move.

On Monday morning amid overcast conditions she was observed flying past the Beauly Firth to reach Teavarran 17.0 km away from where she started. On Tuesday and Wednesday she did not travel much and stayed in the Rootfield area. On Thursday morning she was observed flying down to Farraline, passing Loch Ness and Loch Ruthven. In the afternoon she was spotted in rough grassland near Easter Aberchalder, maybe feeding on small mammals, before retiring to the roost in woodland near Errogie.

[read more....](#)

2012 female bird, which fledged from a nest near Culbokie. Named after the scientific name for the Red kite *Milvus milvus*.

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 (eg, Siri, Cortana) and dialogue systems
- Etc, etc

Image labelling example

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

Arria/Data2text

- Spinout company
 - » Financial reporting
 - » Weather forecasts
 - » Monitoring equipment on oil platforms
- Based in MacRobert building
 - » Always looking for good developers

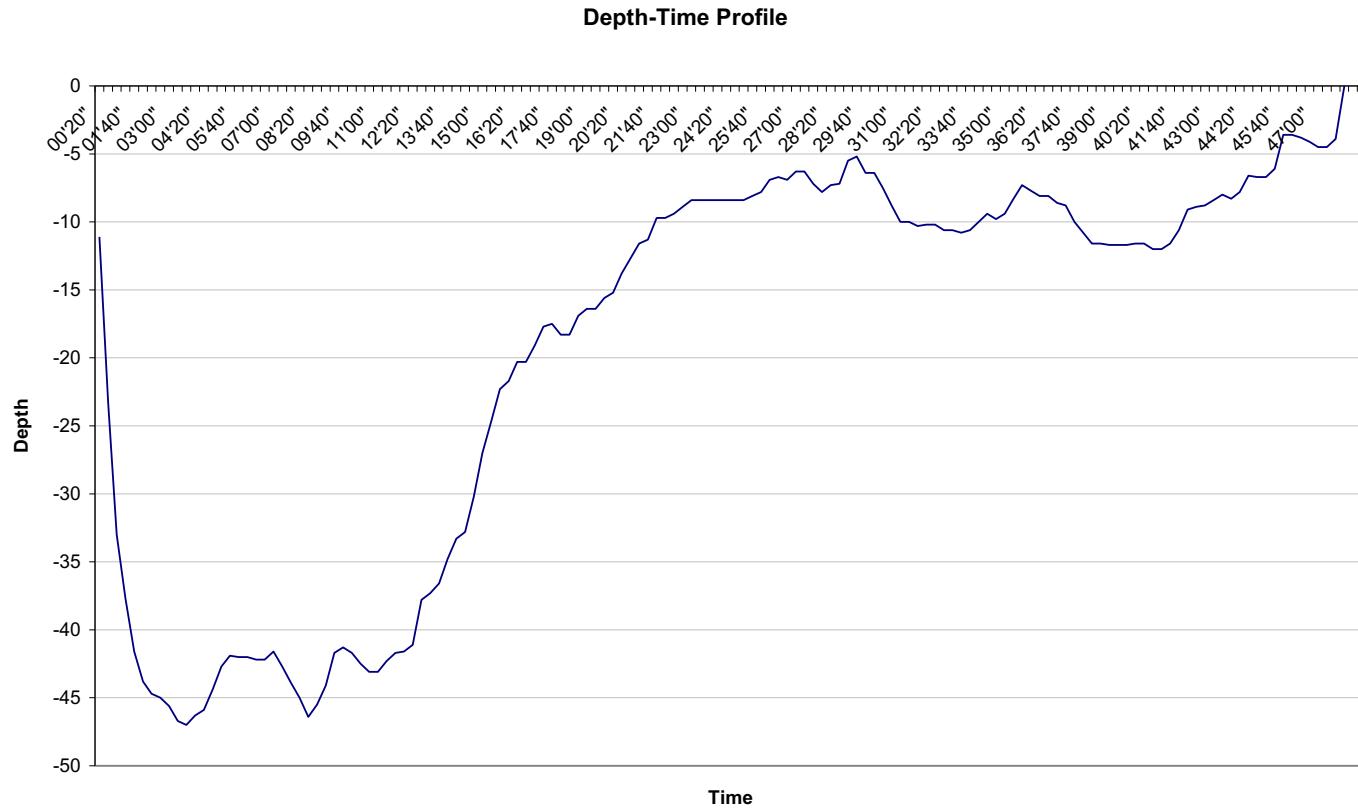
How do NLG Systems Work?

- 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
 - » Eg *children* vs *childs*, *an apple* vs *a apple*

Scubatext example

- Demo system (Dr Sripada) 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



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: eg, depth increases fairly steadily over first 3 minutes
 - » Patterns: eg, 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 is easy to say, what makes good narrative
- *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
 - » 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 *700 sec* vs *714.56 sec*
- Aggregation: 1 sentence or 2 sent?
 - » “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 sublanguage, genre constraints
- Structure: Form legal HTML, RTF, or whatever output format is desired

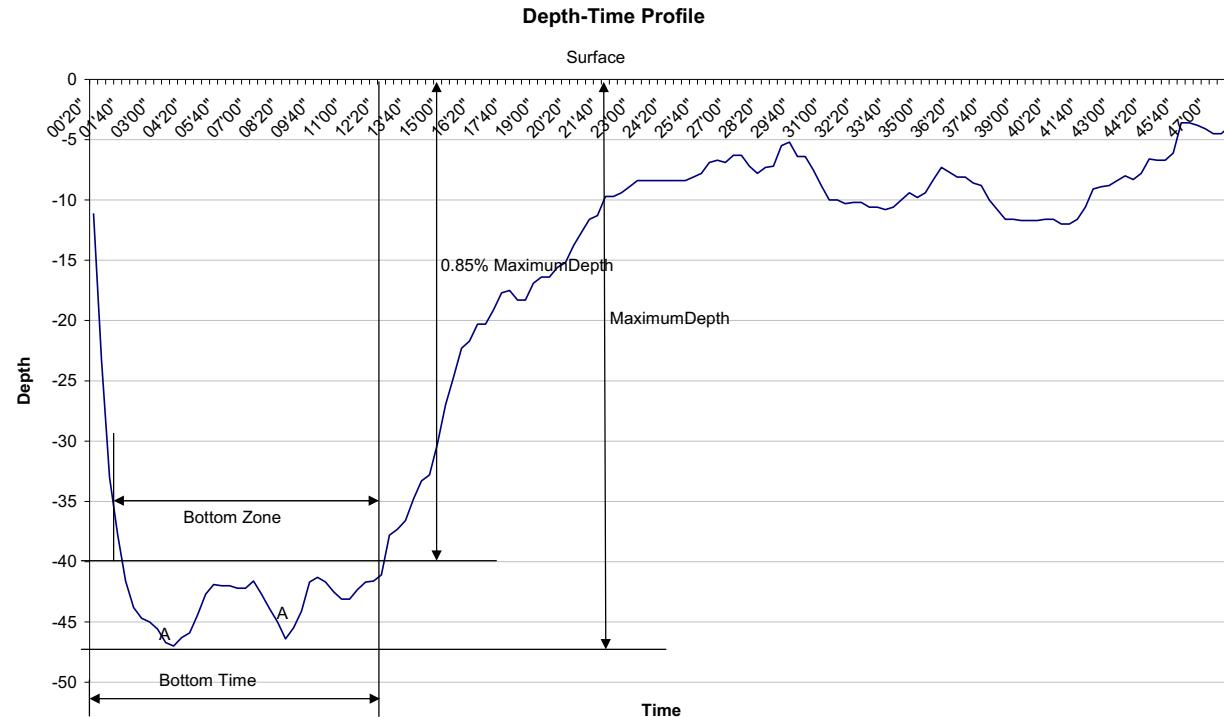
Scuba: Realisation

- 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
 - Eg, 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 buoyancy 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
 - » Manually examine; will focus on
 - » 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

Systems

- Ideally should be able to plug knowledge into NLG framework
 - » Unfortunately good NLG frameworks not available publicly to students and researchers
 - Arria/Data2text has, but proprietary

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
- To date less usage of statistical tech
 - » Focus on high-quality systems in limited domains