NLG Use Case 1: Weather Forecasts

Summary: Weather services and companies use numerical simulations of the atmosphere to predict temperature, wind, precipitation, and other weather parameters. NLG systems can generate textual (written) weather forecasts from this data. Some of Arria's work in this area is described in http://aclweb.org/anthology/W14-4401

There are many types of weather forecasts (see https://www.metoffice.gov.uk/services).   The simplest kind to generate are point forecasts which describe how the weather will change over time at one place (such as Aberdeen Airport); more challenging are regional forecasts which describe how the weather will evolve over a region (such as Scotland).   Different forecasts are produced for different users, because different types of people care about different things. For example, public forecasts (for the general public) emphasise precipitation and temperature, while marine forecasts (for sailors and oil rigs) also talk about wind and wave conditions, and aviation forecasts focus on wind and visibility at altitude as well as on the ground.

Technology: Most commercial NLG forecast generators (including the Arria one) use rule-based techniques. However, Facebook is introducing a neural NLG system to respond to weather questions (https://www.aclweb.org/anthology/2020.coling-industry.7.pdf); their approach requires a very substantial effort in building high-quality training data. Academics have explored using statistical models (eg, https://doi.org/10.1017/S1351324907004664). Some of the research on using machine learning uses the output of rule-based systems for training data, which is dubious (https://ehudreiter.com/2017/05/09/weathergov/).

Data: You can download the SumTime-Meteo corpus of weather data and corresponding human-written forecasts from https://ehudreiter.com/resources/ .   Some of the data used by Facebook is available from https://github.com/facebookresearch/TreeNLG . You can get current weather data from many national weather offices, including UK Met Office (https://www.metoffice.gov.uk/datapoint).

Commercial: Simple rule/template techniques are widely used to generate point forecasts. For example, go to https://www.weather.gov/ and type a location into the box; you will see a template/rule based forecast under Detailed Forecast.   More sophisticated approaches have had limited commercial success.   One problem from a commercial perspective is that there is not a lot of money in weather forecasting (in part because because many types of forecast are delivered for free). Also liability concerns (eg, if a misleading forecast leads to a car crash) make providers cautious.

Projects: A team working in this area should choose a type of forecast, and discuss this with the course lecturers.

NLG use case 2: Automatic Journalism

Summary: There is a lot of interest in using NLG systems to automatically produce news articles (or perhaps article drafts which are revised by a human journalist). Interest is especially strong for local news and financial news (also sports stories, but this is covered under use case 3). For general background, see https://en.wikipedia.org/wiki/Automated_journalism

For this class, we encourage students to look at Urbs Media (https://urbsmedia.com/), which uses NLG (in fact Arria's NLG Studio) to generate draft articles for local media based on government data. For example, Urbs took government data about smoking while pregnant, which is presented for each NHS region, and generated stories for local media based on this data, using the relevant NHS region (eg, the Aberdeen story would be based on data for the NHS Grampian region).

The BBC also uses Arria NLG Studio to generate some stories (https://ehudreiter.com/2019/12/23/election-results-lessons-from-a-real-world-nlg-system/)

Technology: Most automated journalism uses rule-based techniques. Thomson Reuters has used machine learning to extract templates from corpora (http://www.aclweb.org/anthology/P13-1138).

Data: The source data used by Urbs comes from government sources (see https://data.gov.uk/). You can get current newspaper articles from websites. Many news organisations charge money to access old material, but some (including BBC) do not.

Commercial: There are a number of deployed automatic journalism systems (see Wikipedia page for some links). Most of these systems use simple rule/template approaches, in part because this enables journalists to modify and control the systems; it would be very hard for a journalist to modify or control a deep learning system, for example. Most systems also allow journalists to post-edit the computer-generated stories, in part to add local knowledge (eg, about relevant health news which is not in the target data set).

Projects: A team working in this area should choose a type of story, and discuss this with the course lecturers. Please note that although there has been a lot of media attention recently on algorithms which generate fake news, this is not an acceptable topic for this assessment.

NLG Use Case 3: Sports Reporting

Summary: NLG systems can be used to generate stories about sporting matches; see for example https://www.bbc.co.uk/news/technology-34204052 . As with automatic journalism in general, this is mainly used for sports matches of local interests; there are also apps for children's sports (https://gc.com/fan-features), which are used by parents and coaches, not journalists.

The content of the story depends on the sport. For example, football stories often focus on goals, while stories about American baseball tend to be full of statistics about players.   A major challenge in sports reporting is to add "colour" to articles to make them interesting, since sports news is primarily used for entertainment (whereas weather reports, for example, are mostly used to support decision making).   Sometimes NLG is used to generate live updates or tweets, as well as complete descriptions of matches. Reports can cover the match as a whole, or focus on a specific player (eg, "Sam Cosgrove rescues Aberdeen's hopes").

A variant is to generate summaries of games or players for coaches or talent scouts.

Technology: This is similar to automatic journalism.   All commercial deployed systems which I am aware of use rule-based approaches. An Aberdeen PhD student, Craig Thomson, is working on generating basketball summaries using a combination of neural NLG and rule-like domain knowledge.

Data: You can use Craig Thomson's data set, https://github.com/nlgcat/sport_sett_basketball . There is some information about a baseball data set on https://github.com/ratishsp/mlb-data-scripts . Unfortunately most commercial data sets must be paid for, but it maybe possible to scrape data and corpus texts from websites for other sports.

Commercial: Sports is a big business, and there is a lot of interest in automated reporting for sports. With this said, its notable that some NLG companies which started off doing sports reporting (Narrative Science and Automated Insights) no longer seem very interested in this area.   Perhaps one challenge is that many amateurs are happy to write sports stories for free. As with automated journalism in general, some post-editing by human journalists (eg, to add "colour") is common. We are however seeing more interest in using NLG to produce reports for coaches and talent scouts

Projects: A team working in this area should choose a type of sports, and discuss this with the course lecturers. You should also say whether you want to generate live updates or after-match summaries, and whether you want to cover the match as a whole or focus on an individual.

NLG Use Case 4: Product Descriptions

Summary: Online vendors often include descriptions of the things they sell, which summarise key attributes.   This applies to services (such as hotel rooms) as well as objects (such as TVs). NLG systems can automatically generate descriptions if they are based on product data; for example see the description of Trivago's Hotel Scribe system in https://www.aclweb.org/anthology/W19-8647.pdf.   Product descriptions are often used for search engine optimisation (SEO) as well as informing users about products. A real challenge with auto-generating product descriptions is to make a product description unique to that product.

It is often essential that product descriptions be available in different languages, so that users can read descriptions in their language. For this reason there is perhaps a stronger emphasis on multilinguality in product descriptions compared to other NLG applications.

Technology: Ax Semantics, the market leader (https://en.ax-semantics.com/automated-product-descriptions-online-shops) , uses template/rule technology, as does Trivago (they developed their system in-house, using simplenlg). However Ebay seems to be using statistical and neural approaches (http://aclweb.org/anthology/W18-6530). One of the attractions of learning-based approaches in this use case is that it is relatively easy to add additional languages.

Data: The E2E dataset (https://gem-benchmark.com/data_cards/E2E) contains simple restaurant descriptions. ToTTo (https://gem-benchmark.com/data_cards/ToTTo) includes summaries of data from Wikipedia tables, which often describe aspects of entities. It is usually easy to scrape data and descriptions from ecommerce websites.

Commercial: There is certainly a market for product descriptions, especially if they are multilingual. An ecommerce website may sell 100,000 items (or more!), and it is difficult to write descriptions of these items in ten languages, and impossible to keep the descriptions up-to-date. Quality is a concern, and some large companies are considering using NLG but need confidence that the output of the NLG system will always be accurate and readable, for all items and languages.

Projects: A team working in this area should choose a type of product, and discuss this with the course lecturers.   It is OK to just generate descriptions in English!

Use Case 5: Medical Reporting

Summary: NLG has been used for many kinds of medical reporting. The most common usage is automating routine clinical reporting (eg, http://www.aclweb.org/anthology/E03-2008), but there have been many others. Simone Balloccu (course demonstrator) is working with Phillips (www.philhumans.eu) on summarising information for people using personal health apps. Arria has worked with Tibco to produce a text-augmented dashboard which reports on Covid (https://ehudreiter.com/2020/05/21/adding-narrative-to-a-covid-dashboard/).

One challenge in medicine is that clinical reporting is very different in different health care systems (eg, US vs UK), and indeed can differ substantially between different hospitals in the same system. For this reason it is essential that medical NLG systems be flexible and customisable.

Technology: Medical NLG systems are usually based on rules and templates. In part this is because it is usually essential that the content be correct (otherwise a patient might get the wrong treatment), and hence hallucination (https://ehudreiter.com/2018/11/12/hallucination-in-neural-nlg/) is completely unacceptable.

Data: Medical data and texts are hard to get, since it is personal and confidential. However, there are some exceptions, such as https://www.physionet.org/physiobank/database/ . It may be possible to create example data sets based on information from personal acquaintances. I suggest that anyone looking for data talk to Simone.

Commercial: There is considerable commercial interest in this area (including from another Aberdeen spinoff company, https://mimetechnologies.com/), but so far no application has really taken off. The key challenges seem to be scalability across health care providers, and also confidence in the quality of the generated texts.

Projects: Teams working in this area should discuss with the course lecturers, in part to ensure that suitable data is available. It may be best to focus on patient information applications (such as PhilHumans), since clinical applications require a lot of medical domain knowledge.

Use Case 6: Financial Reporting

Summary: Financial reporting is the "hottest" area of commercial NLG in 2021. Almost every NLG company is pushing tools and products which describe and summarise data about sales, profits, and so forth. Financial reports are usually multi-modal, ie they contain charts and tables as well as words, and indeed the words are often used to supplement and explain the data described in the graphs. Arria has done a lot of work on integrating NLG with Business Intelligence tools; see https://www.arria.com/studio-for-bi/ and https://www.arria.com/pbi-gallery/ .

Technology: Rule/template systems dominate, in part because accuracy is often critical (similar to medicine), and also because financial analysts and journalists want to be able to control and modify the NLG system.

Data: There are some excellent providers of financial data but they tend to charge serious money for it. You can get some free data from stock markets and company annual reports. Detailed data about sales and profits is usually commercially sensitive and hence not available.

Commercial: This is currently seen as being very attractive for NLG, in part because of generality. Every company has sales, profits, and losses, and needs to summarise and analyse these, so an NLG system which can help people understand such data can be sold widely to many different customers. Also the data is somewhat consistent across customers, in part because of accounting standards and software packages.