Story Highlight Generation

Pedro Balage, Mariana Almeida, Afonso Mendes, André Martins,
Nikos Papasarantopoulos, Shashi Narayan, Shay Cohen

pedro.balage@priberam.pt, mla@priberam.pt, amm@priberam.pt, afm@priberam.pt
nikos.papasa@ed.ac.uk, shashi.narayan@ed.ac.uk, scohen@inf.ed.ac.uk

Goals

- To synthesise a coherent summary of all events that occur in the story.
- To enable a person monitoring the media to see a global picture and evolution of a story.

## Coveraged-based Extractive Summarisation
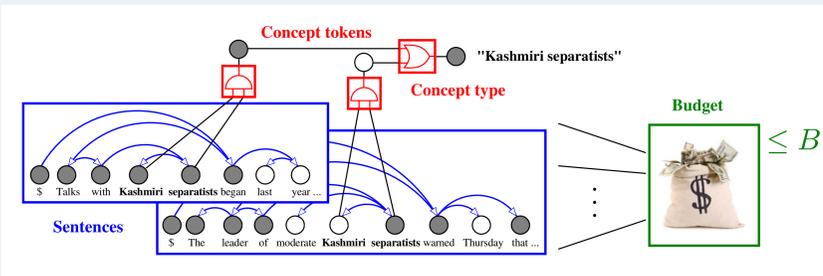
- The goal is to maximise the score of the summary given the constraints for concept coverage and maximum number of words allowed (budget).
- Uses an integer linear programming (ILP) solver.
- Concepts may be scored with standard frequency features or with semantic graph features.

Summariser	R1	R2	RL
Cov-based Extractive	36.43	9.9	32.4
Cov-based Extractive + Feat.	37.0	10.6	33.0
Cov-based Extractive + Sem.Feat	**37.1**	**10.7**	**33.1**

ROUGE score (recall) for multi-document summarisation with 100 words in TAC 2008 corpus.

## Summarisation using AMR-to-text Generation
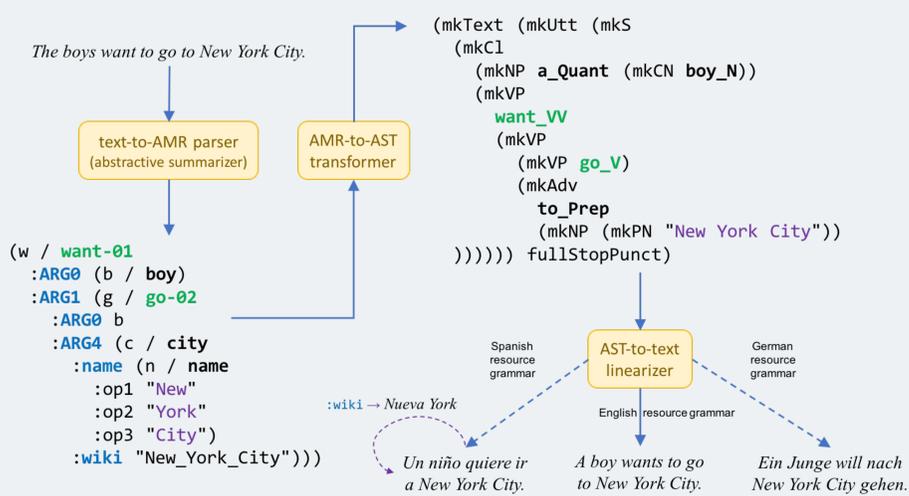
- An NLG component for an abstractive AMR-based summariser.
- Transforms AMR graphs (story highlights) into abstract syntax trees (AST), leaving the linearization of the ASTs to the existing English resource grammar and lexicon.
- Can be extended for multilingual generation of story highlights.

What are the Challenges?

- Leverage summarisation with semantic information.
- Extend neural approaches to multi-document summarisation.

Neural Extractive Summarisation with Side Information
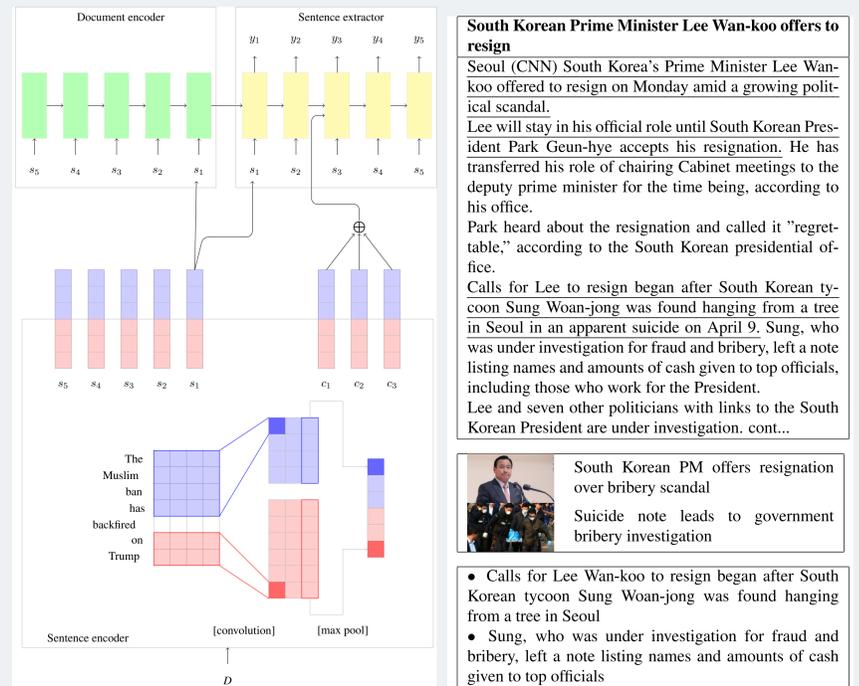
- We explore the side information of the document, such as title and image captions, where gist of the document often lies.
- Model includes a neural network-based hierarchical document encoder and a hierarchical attention-based sentence extractor.
- We report state-of-the-art results without using any human-engineered features.

MODELS	R1	R2	R3	R4	RL
LEAD	49.3	19.5	10.7	6.9	43.8
POINTERNET	51.7	19.7	10.6	6.6	45.7
SIDENET	**54.2**	**21.6**	**12.0**	**7.9**	**48.1**

Rouge score (recall) for the full length summaries on the CNN test set. POINTERNET is the sentence extraction system of Cheng and Lapata.

Coverage-based Extractive + Seq2seq Abstractive

- First step extracts the most relevant sentences (extractive).
- Second step performs a sequence-to-sequence neural network to convert each sentence into a highlight (abstractive).

Model	R1	R2	RL	Avg. size
Extractive	30.90	10.46	26.82	50
Nallapati et al. (2016)	35.46	13.30	32.65	-
Our Method	**35.97**	**14.26**	**33.40**	44.63

Performance on CNN/Daily Mail test set using full-length Rouge-F1 metric.

Demo

SideNet summariser: `http://kinloch.inf.ed.ac.uk/public/direct/sandbox/_site-20170628/sidenet.html`