Deep Graph Convolutional Encoders for Structured Data to Text Generation

Diego Marcheggiani[1,2]   Laura Perez-Beltrachini[1]

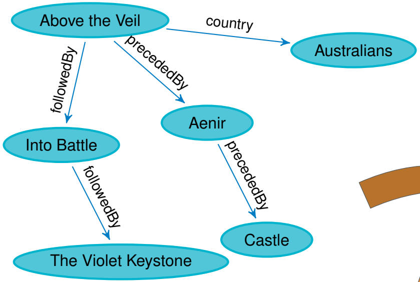
[1] ILCC, University of Edinburgh
[2] ILLC, University of Amsterdam

marcheggiani@uva.nl   lperez@inf.ed.ac.uk

University of Amsterdam

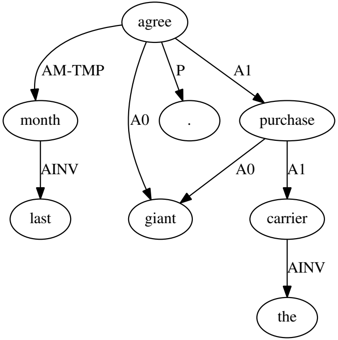
Edinburgh NLP
University of Edinburgh Natural Language Processing

Structured Data-to-Text Generation



Above the Veil is an Australian novel and the sequel to Aenir and Castle . It was followed by Into the Battle and The Violet Keystone .

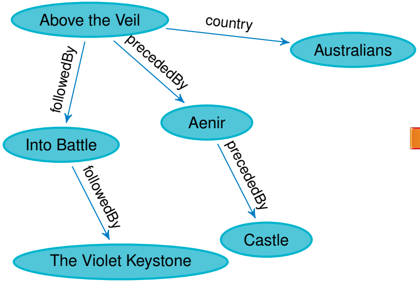
WebNLG [Gardent et al., 2017]



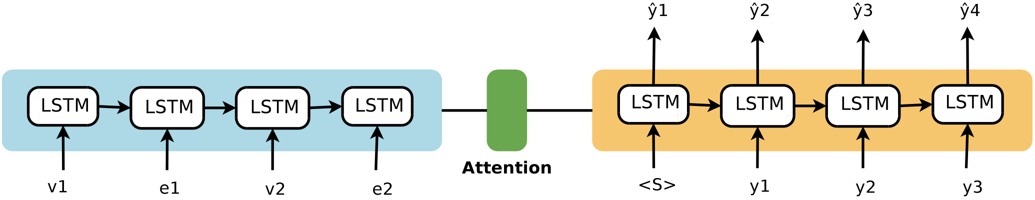
Giant agreed last month to purchase the carrier .

SR11Deep [Belz et al., 2011]

Sequential Encoder-Decoder Models



Above the Veil is an Australian novel and the sequel to Aenir and Castle . It was followed by Into the Battle and The Violet Keystone .



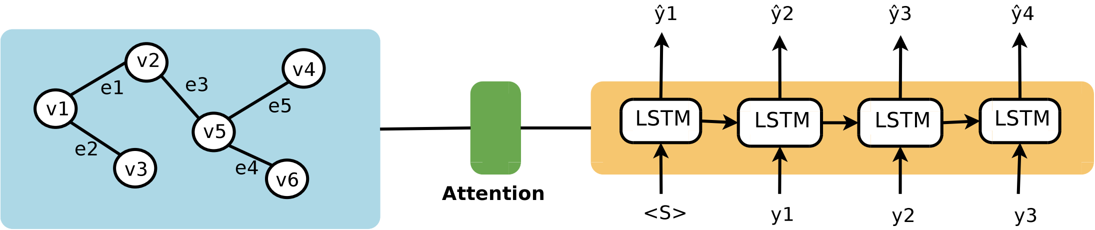
Directly Encoding Input Structure ?

- Sequential encoders, require a separate input **linearisation step**
- After training they will **learn a “structure”** representation
- However, **input explicit structure is NOT directly exploited**

Directly Encoding Input Structure ?

- Sequential encoders, require a separate input **linearisation step**
- After training they will **learn a "structure"** representation
- However, **input explicit structure is NOT directly exploited**

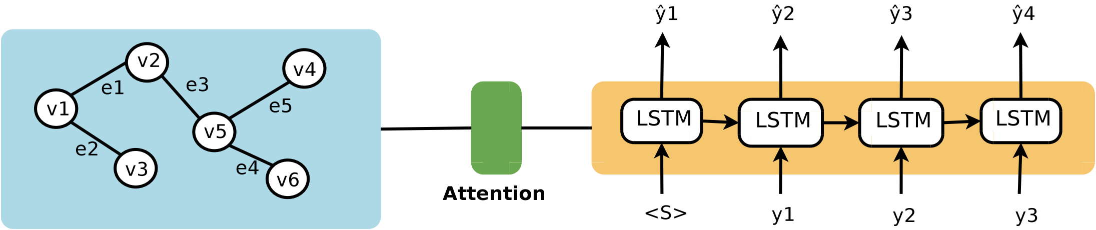
- **Graph Convolutional Network (GCN)** encoder



Directly Encoding Input Structure ?

- Sequential encoders, require a separate input **linearisation step**
- After training they will **learn a "structure"** representation
- However, **input explicit structure is NOT directly exploited**

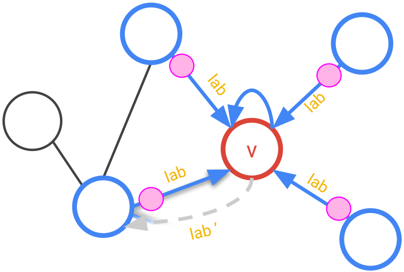
- **Graph Convolutional Network (GCN)** encoder



✓ Input encoding guided by the graph structure
✓ Explicit encoding long-distance dependencies given by the graph
✓ Requiere less amounts of data to learn them

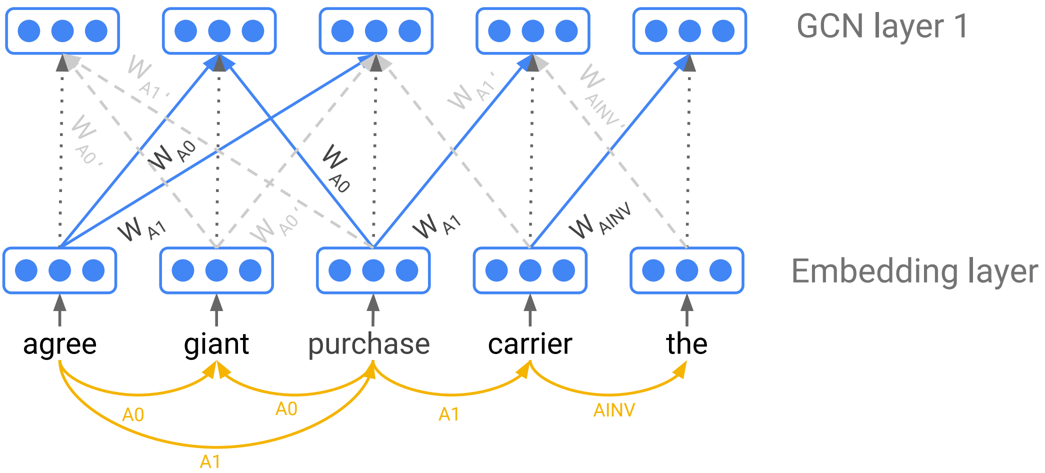
Directed-Labelled Graph Convolutional Networks

- Message passing [Kipf and Welling, 2016]
- Edge directions, labels and importance [Marcheggiani and Titov, 2017]

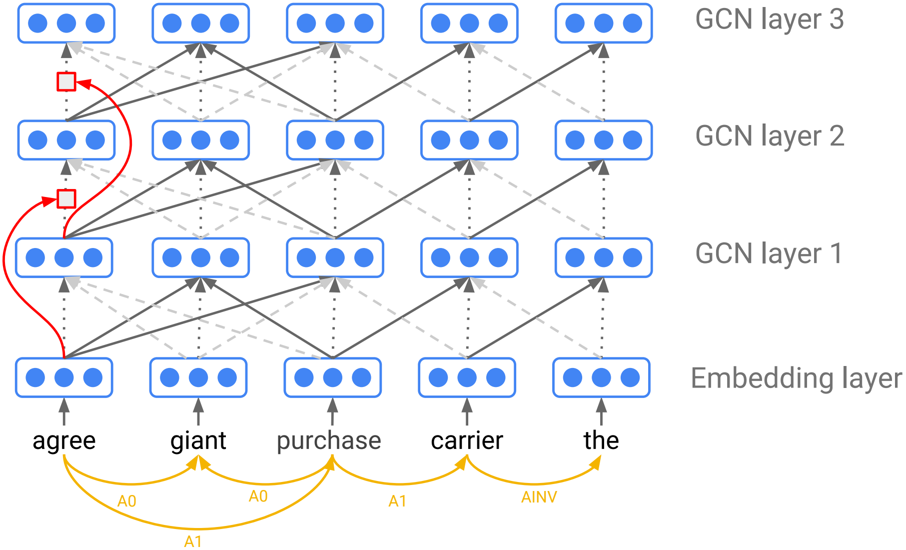


$$\mathbf{h}'_v = \rho\left(\sum_{u \in \mathcal{N}(v)} g_{u,v}\left(W_{dir(u,v)}\,\mathbf{h}_u + \mathbf{b}_{lab(u,v)}\right)\right)$$

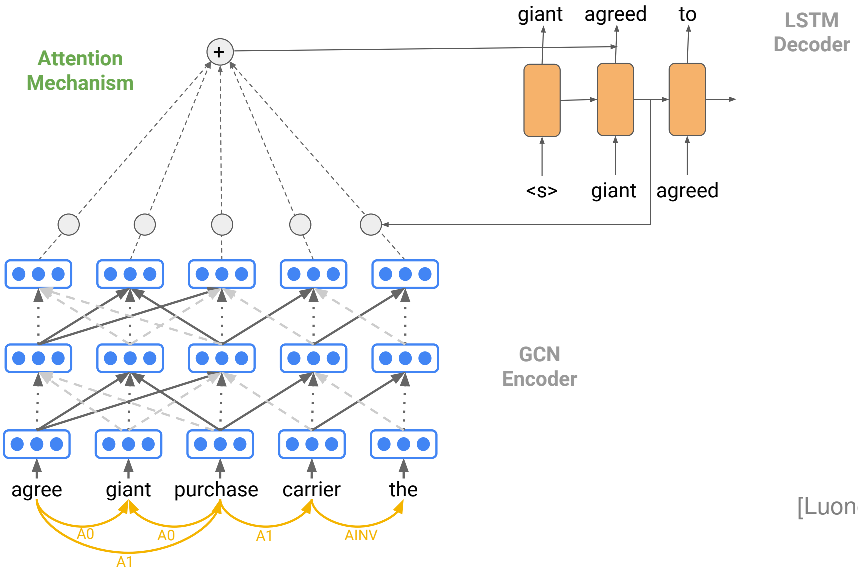
Sinlge Layer GCN Encoder



Stacked Layers and Skip-Connections



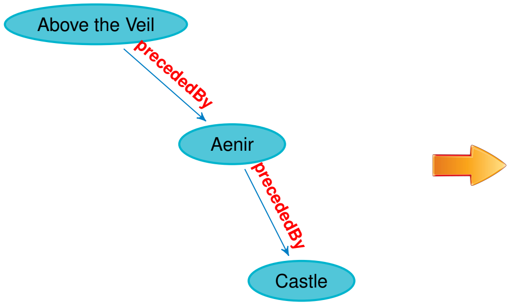
GCN Encoder-Decoder with Attention



[Luong et al., 2015]

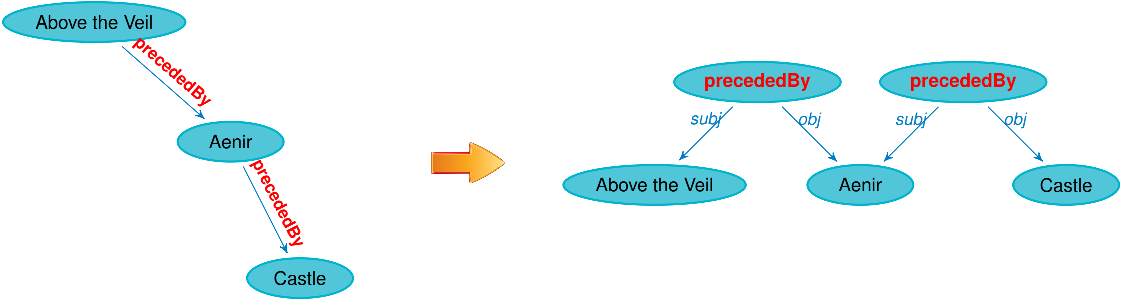
Reification on Knowledge Base (KB) Graphs

[Baader 2003]



Reification on Knowledge Base (KB) Graphs

[Baader 2003]



- The **representation of KB relations** as entities enables **Attention** over them
- Reduces the number of KB relations to be modelled as network parameters

Experimental Setup

✓ Encoders Comparison

✓ Existing Systems Comparison

WebNLG [Gardent et al., 2017]

- **GCN** (4 layers +residual encoder, 1 layer decoder, 256 dim)
- **LSTM** (1 layer encoder, 1 layer decoder, 256 dim) + linearisation
- GCN + pre-trained Embeddings and Copy (**GCN**$^{++}$)

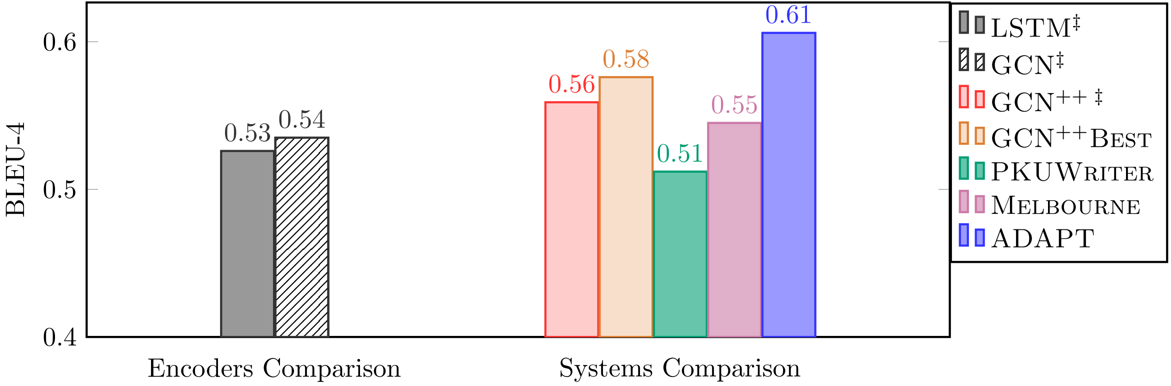
SR11Deep [Belz et al., 2011]

- **GCN** (7 layers +dense encoder, 1 layer decoder, 500 dim)
- **LSTM** (1 layer encoder, 1 layer decoder, 500 dim) + linearisation
- Encode morphological features present in the input (**GCN**$_{morph}$)

*model selection done on the development set

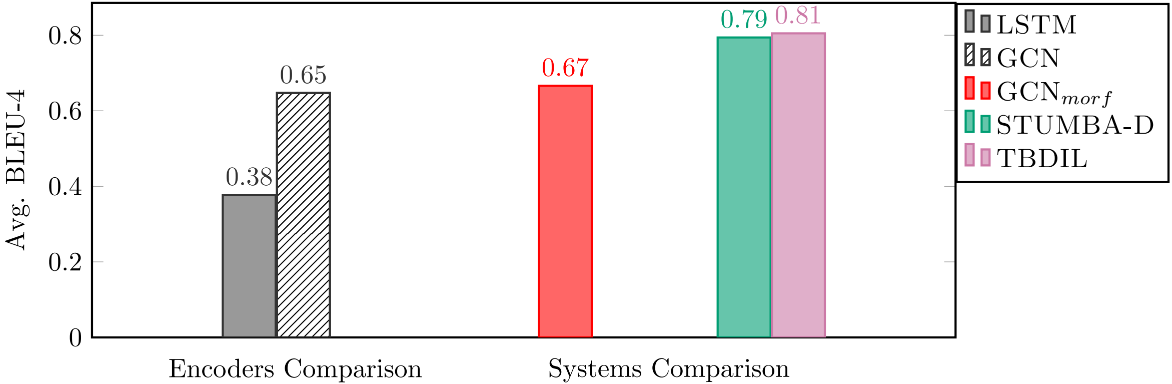
GCN performance on WebNLG

‡ Average BLEU-4 on 3 runs



*PKUWriter, Melbourne and ADAPT neural systems participating on the WebNLG challenge

GCN performance on SR11Deep



* STUMBA-D and TBDIL non-neural systems with pipeline of classifiers

Example Outputs WebNLG

Input graph:
(William Anders dateOfRetirement 1969 - 09 - 01)
(William Anders was a crew member of Apollo 8)
(Apollo 8 commander Frank Borman)
(Apollo 8 backup pilot Buzz Aldrin)

(LSTM) William Anders was a crew member of the OPERATOR operated Apollo 8 and retired on September 1st 1969 .

(GCN) William Anders was a crew member of OPERATOR ' s Apollo 8 alongside backup pilot Buzz Aldrin and backup pilot Buzz Aldrin .

(GCN^{++}) william anders , who retired on the 1st of september 1969 , was a crew member on apollo 8 along with commander frank borman and backup pilot buzz aldrin .

Example Outputs SR11Deep

Reference:
The economy 's temperature will be taken from several vantage points this week , with readings on trade , output , housing and inflation .

(LSTM) the economy 's accords will be taken from several phases this week , housing and inflation readings on trade , housing and inflation .

(GCN) the economy 's temperatures will be taken from several vantage points this week , with reading on trades output , housing and inflation .

Concluding Remarks

- GCN-based generation architecture that directly encodes explicit structure in the input
- GCN -based models outperform a sequential baseline on automatic evaluation
  - improve on over- and under- generation cases
- Relational inductive bias of the GCN encoder produces more informative representations of the input [Battaglia et al., 2018]

Concluding Remarks

- GCN-based generation architecture that directly encodes explicit structure in the input
- GCN -based models outperform a sequential baseline on automatic evaluation
  - improve on over- and under- generation cases
- Relational inductive bias of the GCN encoder produces more informative representations of the input [Battaglia et al., 2018]

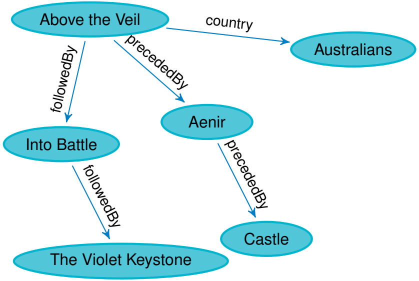
Future work

- Other input graph representations
  - Abstract Meaning Representations (AMR; [Banarescu et al., 2013])
  - Scoped semantic representations [Van Noord et al., 2018]
  - Scene graphs [Schuster, et al., 2015]
- Multi-lingual training of GCN layers with universal dependencies [Mille, et al., 2017]

Code (PyTorch) + Data: `github.com/diegma/graph-2-text`

Thank you!
Questions?

Separate Input Linearisation Step



Avobe the veil followedBy Into Battle | Avobe the veil followedBy the violet keystone | ...

Avobe the veil country Australians | Avobe the veil followedBy Into Battle | ...

Avobe the veil precededBy Aenir | Avobe the veil precededBy Castle | ...

Automatic Evaluation

Encoder	BLEU	METEOR	TER
LSTM	.526±.010	.38±.00	.43±.01
GCN	.535±.004	.39±.00	.44±.02
ADAPT	.606	.44	.37
GCN^{++}	.559±.017	.39±.01	0.41±.01
MELBOURNE	.545	.41	.40
PKUWRITER	.512	.37	.45

Table : Test results WebNLG task.

Encoder	BLEU	METEOR	TER
LSTM	.377±.007	.65±.00	.44±.01
GCN	.647±.005	.77±.00	.24±.01
GCN+feat	.666±.027	.76±.01	.25±.01

Table : Test results SR11Deep task.

Ablation Study

	BLEU			SIZE		
Model	none	res	den	none	res	den
LSTM	.543±.003	-	-	4.3	-	-
GCN						
1L	.537±.006	-	-	4.3	-	-
2L	.545±.016	.553±.005	.552±.013	4.5	4.5	4.7
3L	.548±.012	.560±.013	.557±.001	4.7	4.7	5.2
4L	.537±.005	.569±.003	.558±.005	4.9	4.9	6.0
5L	.516±.022	.561±.016	.559±.003	5.1	5.1	7.0
6L	.508±.022	.561±.007	.558±.018	5.3	5.3	8.2
7L	.492±.024	.546±.023	.564±.012	5.5	5.5	9.6