Exploiting Semantics in Neural Machine Translation with Graph Convolutional Networks

Diego Marcheggiani¹ ², Joost Bastings², Ivan Titov¹ ²

{marcheggiani, bastings}@uva.nl, titov@inf.ed.ac.uk

¹ILLC, School of Informatics, University of Edinburgh
²ILLC, University of Amsterdam

We use graph convolutional networks (GCNs) to inject semantic structures into a neural machine translation.

The semantic-aware model outperforms the linguistic-agnostic one (EN-DE WMT16).

Syntax and semantics are complementary.

Semantic Role Labeling

• A0: the giver
• A1: the thing being given
• A2: an entity given to

• Predicate-argument structures define an abstraction which helps to generalize over different surface realizations of the same underlying “meaning”.
• We explicitly encode semantic role labeling (SRL) graphs using GCNs.

Graph Convolutional Networks

• Nodes are represented as vectors
• Messages are summed up following the graph structure
• Each W for every edge but self loop

• GCNs are used to encode syntactic dependencies for SRL and NMT.
• SRL graphs have directed and labeled edges.
• Messages passed through different edges have different meaning.
• Parameterize each message according to syntactic label and edge direction.

Output

\[ h_i^{k+1} = \text{ReLU} \left( \sum_{j \in \mathcal{N}(i)} W_{ij} h_j^k + b_i \right) \]

Self loop

Input

Hidden layer

Output

GCN Encoder-Decoder Model

• We thank Stella Frank and Wilker Aziz for their suggestions and comments.
• The project was supported by ERC StG BroadSem 679254 grant, NWO Vidi 639.022.518 grant.
• We thank NVIDIA for donating the GPUs used for this research.

Acknowledgments

References

• Diego Marcheggiani and Ivan Titov. Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling. EMNLP 2017.