Scoring Lexical Entailment with a Supervised Directional Similarity Network

Marek Rei, Daniela Gerz and Ivan Vulić

UNIVERSITY OF CAMBRIDGE

Lexical Relations

Task: Graded lexical entailment
To what degree is X a type of Y?

girl → person	9.85 / 10
guest → person	7.22 / 10
person → guest	2.88 / 10

Useful for query expansion, natural language inference, paraphrasing, machine translation, etc.

Lexical Relations

- Distributional vectors are not great for directional lexical relations

carrot ~ vegetable new ~ old

- Retro-fitting (Faruqui et al., 2015)
  Counter-fitting (Mrkšić et al., 2016)

BUT these mostly affect words that are in the training data

Main Idea

01 Specialized network for directional lexical relations

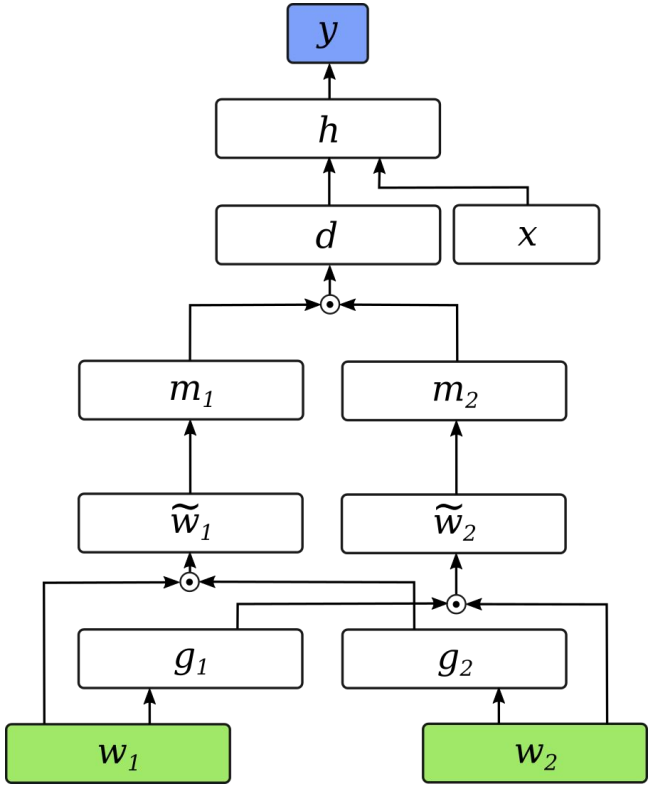
02 Off-the-shelf pre-trained embeddings

03 Train the network to discover task-specific regularities in the embeddings

Supervised Directional Similarity Network

Fixed pre-trained word embeddings as input

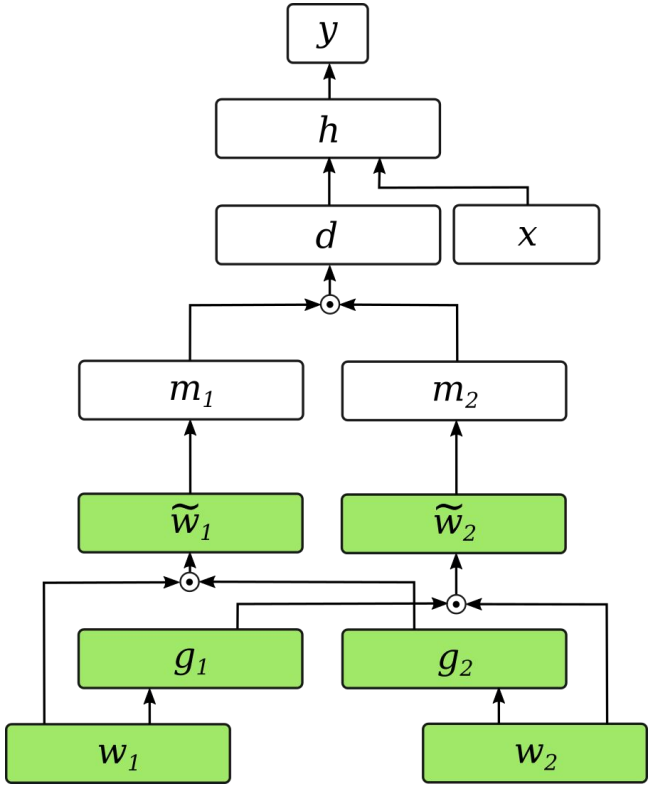
Predict a score indicating the strength of a specific lexical relation



SDSN: Gating

Conditioning each word based on the other

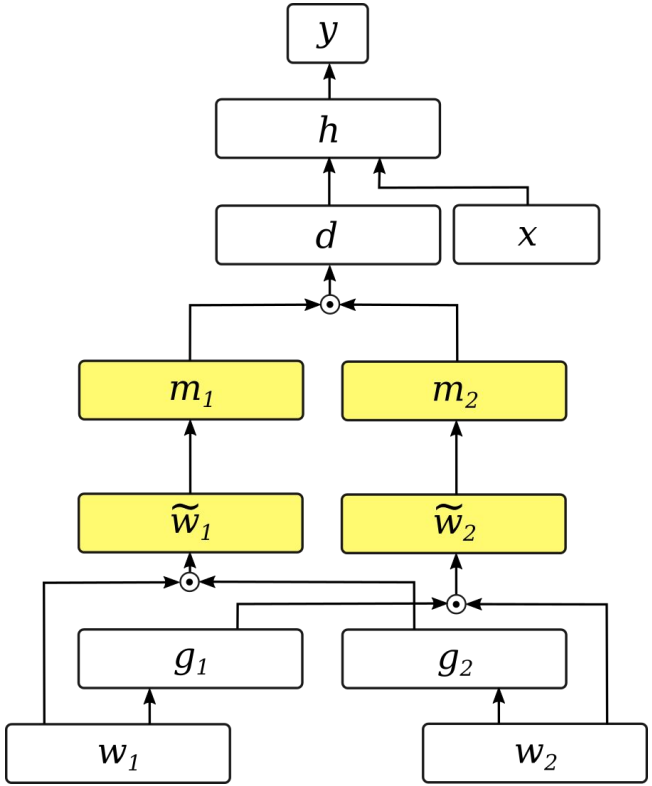
$$g_1 = \sigma(W_{g_1} w_1 + b_{g_1})$$
$$g_2 = \sigma(W_{g_2} w_2 + b_{g_2})$$
$$\widetilde{w}_1 = w_1 \odot g_2$$
$$\widetilde{w}_2 = w_2 \odot g_1$$



SDSN: Mapping

Mapping the representations to new spaces

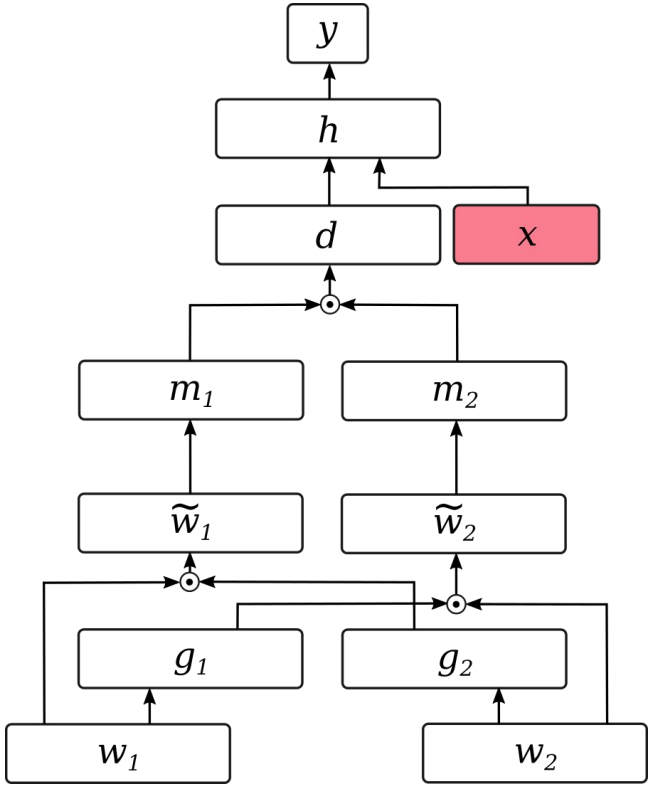
$$m_1 = tanh(W_{m_1}\widetilde{w}_1 + b_{m_1})$$
$$m_2 = tanh(W_{m_2}\widetilde{w}_2 + b_{m_2})$$



SDSN: Sparse Features

Features based on sparse distributional representations

- cosine
- weighted cosine (Rei & Briscoe, 2014)
- ratio of shared contexts

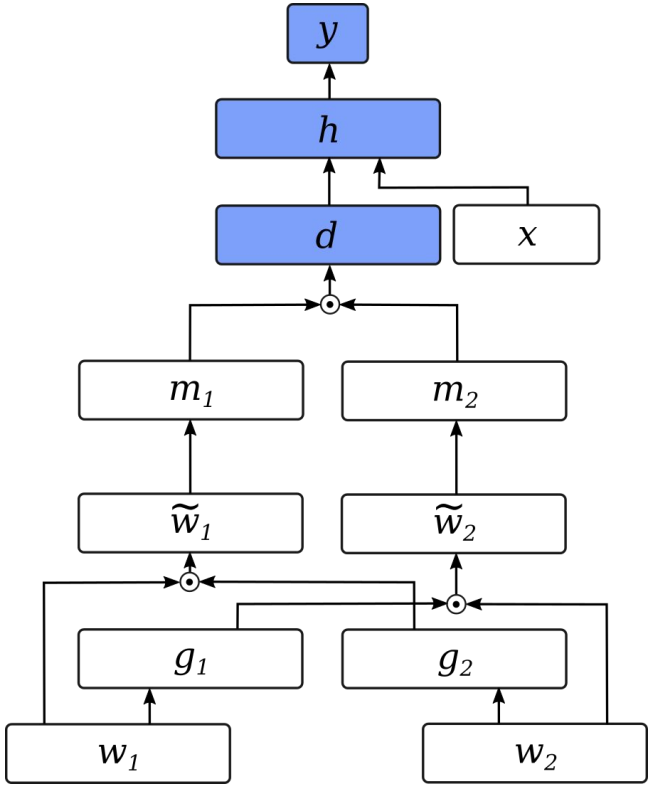


SDSN: Scoring

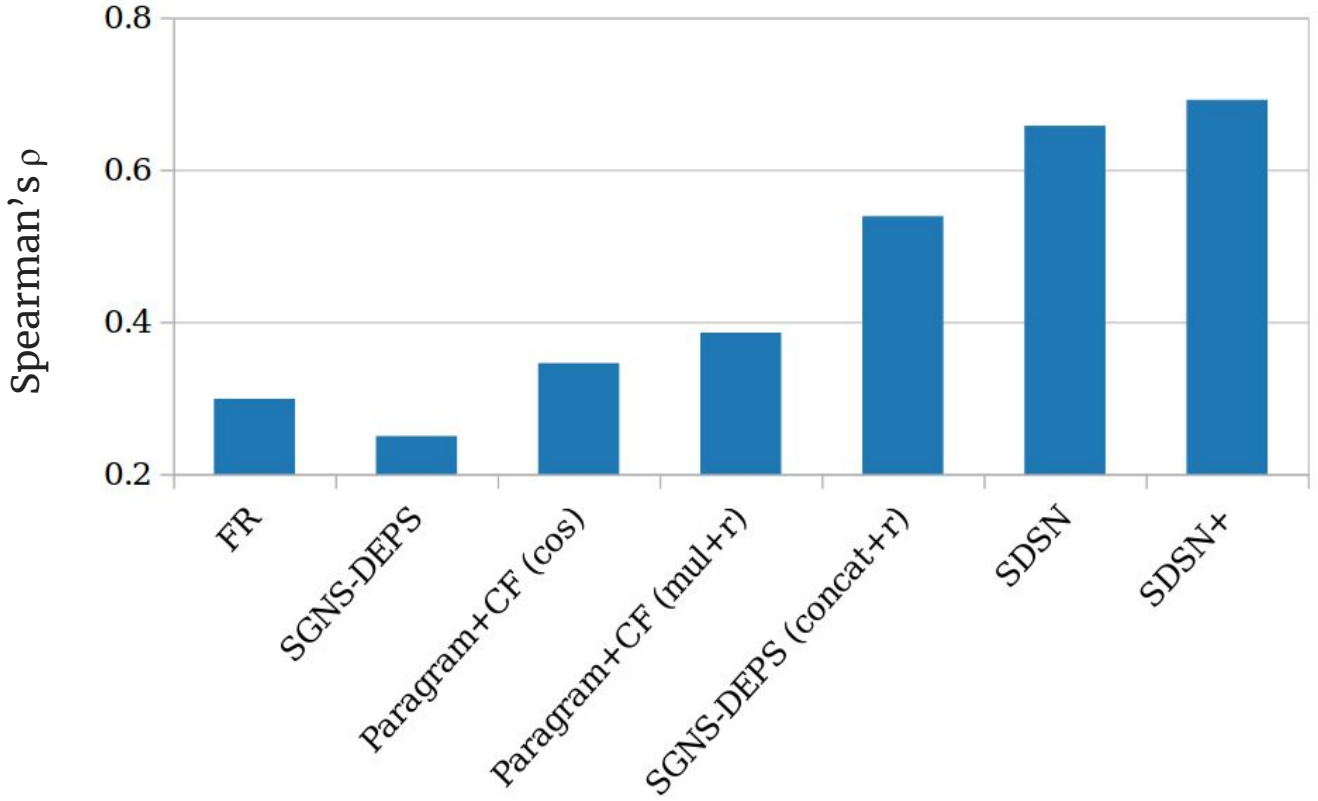
Mapping the representations to a score

Optimize the network with labeled examples

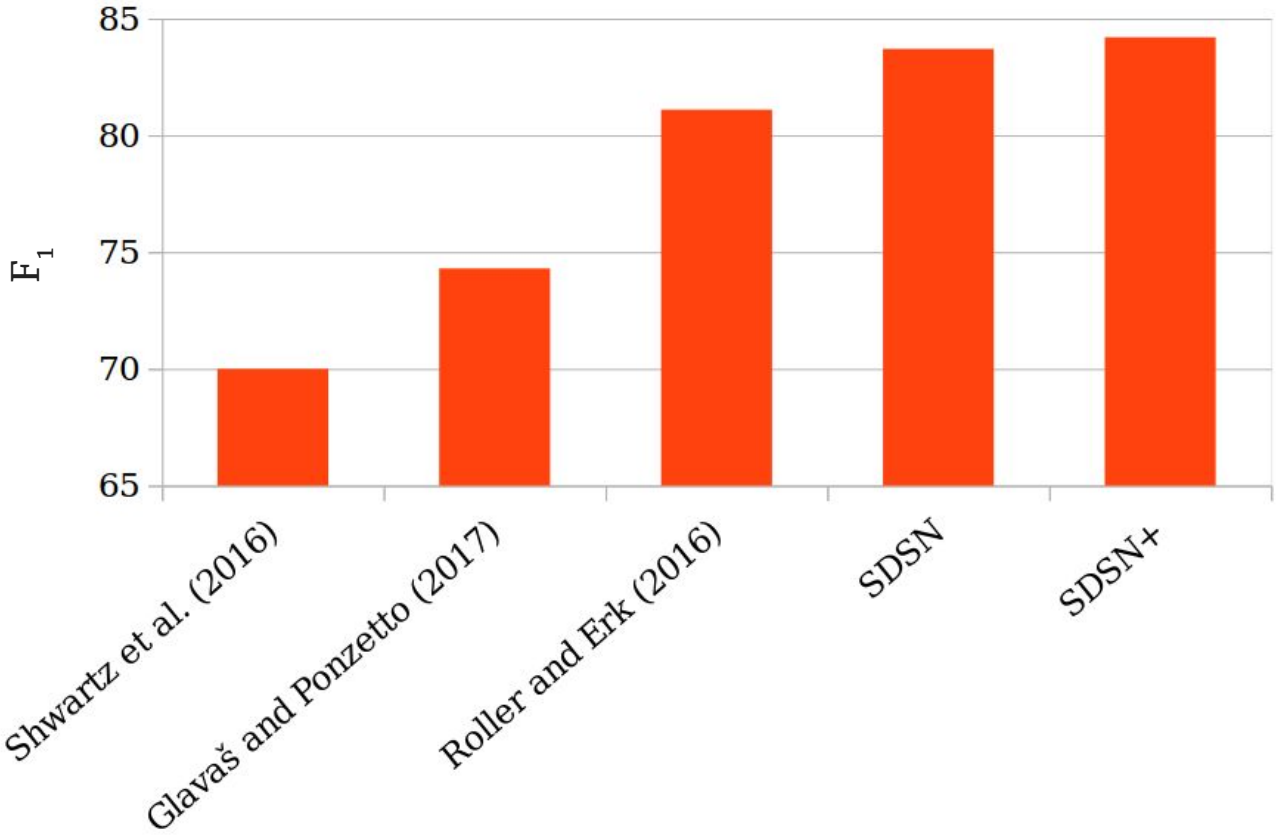
$$d = m_1 \odot m_2$$
$$h = tanh(W_h d + W_x x + b_h)$$
$$y = S \cdot \sigma(a(W_y h + b_y))$$



HyperLex: Graded Lexical Entailment



HypeNet: Hyponym Detection



Conclusion

01 Can train a neural network to find specific regularities in off-the-shelf word embeddings

02 Traditional sparse embeddings still provide complementary information

03 Achieves state-of-the-art on graded lexical entailment

Thank you!
Any questions?

Examples

Premise	Hypothesis	Gold	Predicted
captain	officer	8.22	8.17
celery	food	9.3	9.43
horn	bull	1.12	0.94
wing	airplane	1.03	0.84
prince	royalty	9.85	4.71
autumn	season	9.77	3.69
kid	parent	0.52	8.00
discipline	punishment	7.7	3.2