Sushrut Thorat & Varad Choudhari

Implementing a Reverse Dictionary, based on word definitions, using a Node-Graph Architecture

Read more at:
https://goo.gl/Wh2Sxh

The Problem: Computing words that semantically represent a phrase

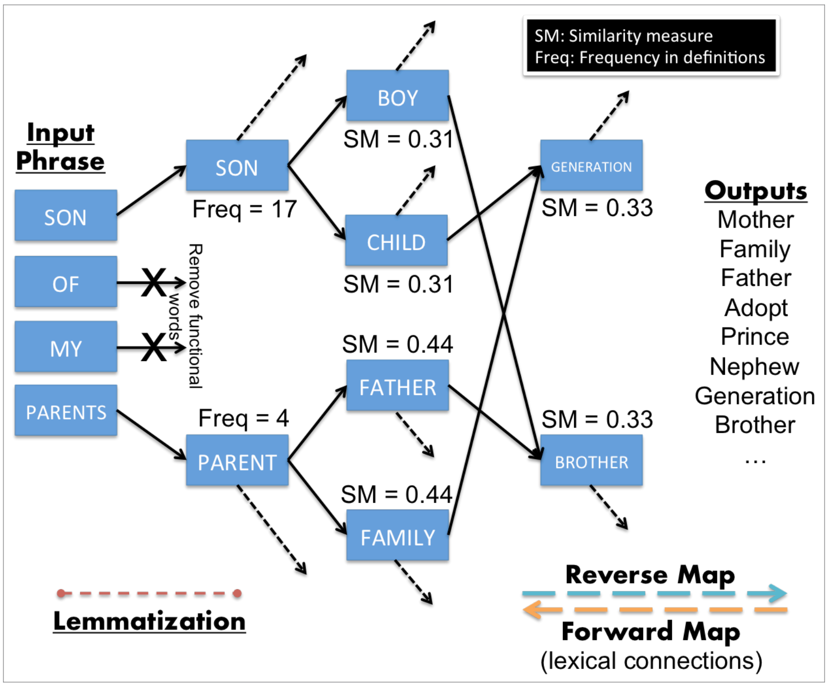
A reverse dictionary maps all phrases to semantically similar words (not just definitions)
Examples:
1. Son of my parents - Brother
2. A hard outer covering - Shell

Previous Work: (Systematic vs *Jugaad*)
1. Onelook Reverse Dictionary
   - Proprietary algorithm
2. Shaw, R., et.al., 2013
   - Shallow reverse search using lexical relations
3. Hill, F., et.al., 2015
   - RNN trained on definitions-word sets

The **Similarity Measure** between a Phrase and a Word:

$$E_{W,P} = \frac{\sum_i \left(\nu_{P_i} \times d_{W,P_i}\right)^{-1}}{\sum_i \nu_{P_i}^{-1}}$$

where, P_is are the lemmatised content words of P
, ν_w is the frequency of w in definitions
, d is the distance on the reverse map



Performance Testing

Test Set Generation:
- 179 user generated phrases given target words
  - The phrases are all different from dictionary definitions
- 179 definitions from Macmillian word dictionary
  - These definitions were not used in generating the reverse map

Measure of performance:
- Rank of target word in the RD output

Performance Results

Test Type →	Macmillian Word Definitions (179)			User Concept Descriptions (179)		
Evaluation → Models ↓	**Accuracy** @1/10/100	Rank Median	Rank σ	**Accuracy** @1/10/100	Rank Median	Rank σ
Onelook	.19/.41/.65	5	24	.04/.21/.40	10	26
Onelook, *corr**	.20/.46/.68	3	20	.07/**.26**/.52	13	30
W2V	.01/.06/.20	23	30	.01/.05/.18	34	28
W2V, *corr**	.02/.11/.29	21	26	.01/.08/.26	21	29
Chance, 3k	$10^{-4}/10^{-3}/.03$	50	29	$10^{-4}/10^{-3}/.03$	50	29
Fusion, FLM	.02/.10/.21	12	28	.01/.07/.22	16	21
Fusion, mBLM	.25/**.55**/**.84**	4	22	**.10**/.23/**.53**	14	26
OLD, mBLM	**.26**/.52/.78	4	23	.04/.17/.43	14	25
WN, BLM	.08/.27/.54	11	26	.06/.18/.41	14	26
MW, mBLM	.17/.39/.63	5	20	.05/.20/.43	15	25
WL, 80k	.03/.15/.36	18	26	.05/.11/.24	14	25
WL, *corr**	.07/.26/.52	10	25	.07/.18/.35	10	23

- Accuracy: Fraction of target words under the specified rank
- Performance: Fusion mBLM ~ Onelook (corrected for 3k lexicon)
- W2V, denoting naive additive composition, performs poorly
- All performances significantly above chance
- Our approach doesn't scale well (as seen through the use of 80k lexicon)

Concluding Remarks

- Shallow semantic extraction rivals the performance of sophisticated ML systems such as Onelook Reverse Dictionary and those using RNNs (Hill, F., et.al., 2015)
- Setting a new baseline to test any new RD systems against (as a sanity check for extracting deep, useful semantic features)
- Simple compositionally (Additive W2V) cannot explain phrasal semantics
  - We need automatic extraction of features - RNNs, etc. are the way forward towards automating phrasal semantic analysis.

Comments/Questions?

Contact us:
Sushrut Thorat - sushrut.thorat94@gmail.com
Varad Choudhari - varad.choudhari@gmail.com