

EXPLORING TRANSFORMERS FOR RANKING PORTUGUESE SEMANTIC RELATIONS

Hugo Gonalo Oliveira, hroliv@dei.uc.pt

CISUC, DEI, University of Coimbra, Portugal

Introduction

Knowledge Bases

- Well-defined relations, represented by triples ' arg_1 related-to arg_2 ', interpretable
- Manual creation \rightarrow time-consuming, coverage issues
- Automatic creation \rightarrow noisy
- Approaches for computing the confidence of extractions [1, 2]

Transformer Language Models (TLMs)

- e.g., GPT, BERT
- Unprecedented advances in NLP
- Encode much linguistic and world knowledge

Language models as Knowledge Bases! [3]

- Generate text after prompts
- Predict masked tokens
- Compute the likelihood of sequences of text**
 - Which may express semantic relations!
 - e.g. *A dog is a mammal*; *A knife is used for cutting*

Experimentation

TLMs for ranking lexico-semantic relations, in Portuguese

- Filter out very specific / incorrect relation instances

Instances from ten lexical resources [4]

- 862,693 instances, some extracted automatically
- 16 types: synonymy (4), hypernymy (2), part (2), member, purpose (2), causation, place, property (2)
- Variable utility
 - prototypical / widely accepted (*tree* hypernym-of *oak*, *to cook* purpose-of *oven*)
 - very specific (*cd store* place-of *elvis presley cd*; *give to girlfriend* purpose-of *kitty*)
 - underspecified / incomplete / incorrect arguments (*possessive said-about to make*, *various causes contest*)

- Weighted by the number of resources they are in (Res , 1–10)

Exploit two available TLMs

- BERTimbau [5] (base), BERT pre-trained for Portuguese
- GPoTuguese-2^a, GPT2-small fine-tuned for Portuguese.

Approach:

- handcraft textual templates that transmit / are compatible with relations of the target types (see Table below);
- instantiate the templates for each instance, according to its type;
- use a TLM for computing the likelihood of the resulting sequences.

More prototypical / reliable instances have higher scores?

Table: Lexical patterns indicating lexico-semantic relations.

Relation	Pattern	
SINONIMO.N.DE	A_1 é o mesmo que A_2	(A_1 is the same as A_2)
	A_1 é sinónimo de A_2	(A_1 is a synonym of A_2)
	A_1 é igual a A_2	(A_1 is equals to A_2)
SINONIMO.V.DE	A_1 é o mesmo que A_2	(A_1 is the same as A_2)
	A_1 é a mesma coisa que A_2	(A_1 is the same thing as A_2)
	A_1 é sinónimo de A_2	(A_1 is a synonym of A_2)
SINONIMO.ADJ.DE	A_1 é o mesmo que ser A_2	(A_1 is the same as being A_2)
	A_1 é o mesmo que A_2	(A_1 is the same as A_2)
	A_1 é sinónimo de A_2	(A_1 is a synonym of A_2)
SINONIMO.ADV.DE	A_1 é o mesmo que A_2	(A_1 is the same as A_2)
	A_1 é sinónimo de A_2	(A_1 is a synonym of A_2)
	fazer A_1 é o mesmo que fazer A_2	(to do A_1 is the same as to do A_2)
HIPERONIMO.DE	A_1 é hiperónimo de A_2	(A_1 is a hypernym of A_2)
	A_2 ou outro A_1	(A_2 or other A_1)
	A_2 é um tipo de A_1	(A_2 is a type of A_1)
HIPERONIMO.ACCAO.DE	A_2 e outros modos de A_1	(A_2 and other modes of A_1)
	A_2 e outras maneiras de A_1	(A_2 and other manners of A_1)
	A_2 ou outras maneiras de A_1	(A_2 or other manners of A_1)
PARTE.DE	A_2 tem A_1	(A_2 has A_1)
	A_2 possui A_1	(A_2 possesses A_1)
	A_1 do A_2	(A_1 's A_2)
PARTE.DE.ALGO.COM.PROP	A_2 porque tem A_1	(A_2 because it has A_1)
	A_2 tem A_1	(A_2 has A_1)
	A_1 do que é A_2	(A_1 of what is A_2)
MEMBRO.DE	A_1 é membro de A_2	(A_1 is a member of A_2)
	A_1 pertence a A_2	(A_1 belongs to A_2)
	A_1 faz parte de A_2	(A_1 is part of A_2)
FINALIDADE.DE	preciso de uma A_2 para A_1	(I need a A_2 for A_1)
	A_2 serve para A_1	(A_2 is for A_1)
	A_2 é usado para A_1	(A_2 is used for A_1)
FAZ.SE.COM	A_1 faz-se com A_2	(A_1 is made with A_2)
	A_2 para fazer A_1	(A_2 to make A_1)
	A_2 para A_1	(A_2 for A_1)
ACCAO.QUE.CAUSA	A_1 causa A_2	(A_1 causes A_2)
	A_2 resulta de A_1	(A_2 results from A_1)
	A_2 é um efeito de A_1	(A_2 is an effect of A_1)
LOCAL.ORIGEM.DE	A_2 vem de A_1	(A_2 comes from A_1)
	A_2 é de A_1	(A_2 is from A_1)
	A_2 de A_1	(A_2 of A_1)
DIZ.SE.DO.QUE	A_1 diz-se do que A_2	(A_1 is said of what A_2)
	A_1 diz-se daquele que A_2	(A_1 is said of the one that A_2)
	A_1 porque A_2	(A_1 because A_2)
DIZ.SE.SOBRE	A_1 diz-se sobre A_2	(A_1 is said about A_2)
	A_1 é relativo a A_2	(A_1 is relative to A_2)
	A_2 devido a A_1	(A_2 because A_1)
PROPRIEDADE.SEMELHANTE.A	A_1 é semelhante a A_2	(A_1 is similar to A_2)
	A_1 é parecido com A_2	(A_1 looks like A_2)
	A_1 parece A_2	(A_1 seems like A_2)

$weight = e^{loss}$ normalised to the 1-10 interval

- For each model, BERT (B) and GPT (G), and pattern (1-3)
- Plus maximum (Mx) and average (Av) weight

Analysis

Table: Weights for three selected instances.

Instance	Res	B1	B2	B3	G1	G2	G3	Mx(B)	Av(B)	Mx(G)	Av(G)
feito SINONIMO.ADJ.DE concluído	3	5.28	5.56	6.48	5.37	5.59	6.34	6.12	5.86	6.34	5.99
cozer FINALIDADE.DE panela	2	4.11	4.75	4.62	4.54	5.18	4.74	4.75	4.64	5.18	4.93
sonar PARTE.DE submarino	1	4.69	4.50	4.56	5.68	4.25	5.84	4.54	4.95	5.63	5.57

Table: Pearson correlation between average weights and Res.

Relation	Mx(B)	Av(B)	Mx(G)	Av(G)
SINONIMO.N.DE	0.99	0.99	-0.33	-0.64
SINONIMO.V.DE	0.98	0.98	-0.86	-0.92
SINONIMO.ADJ.DE	0.94	0.91	-0.76	-0.65
SINONIMO.ADV.DE	0.82	0.82	-0.23	-0.28
HIPERONIMO.DE	0.99	0.99	-0.69	-0.75
HIPERONIMO.ACCAO.DE	0.98	0.98	-0.90	-0.95
PARTE.DE	0.56	0.64	0.01	0.06
PARTE.DE.ALGO.COM.PROP	0.72	0.70	0.97	0.98
MEMBRO.DE	0.82	0.86	-0.67	-0.82
FINALIDADE.DE	0.97	0.96	0.47	0.16
FAZ.SE.COM	0.88	0.82	0.99	0.98
ACCAO.QUE.CAUSA	0.75	0.73	0.81	0.80
LOCAL.ORIGEM.DE	-0.42	-0.39	0.08	0.01
DIZ.SE.DO.QUE	0.93	0.92	0.94	0.93
DIZ.SE.SOBRE	0.90	0.88	0.37	-0.02
PROPRIEDADE.SEMELHANTE.A	0.94	0.94	1.00	1.00
All	0.95	0.96	-0.88	-0.91

Figure: Average weight according to Res.

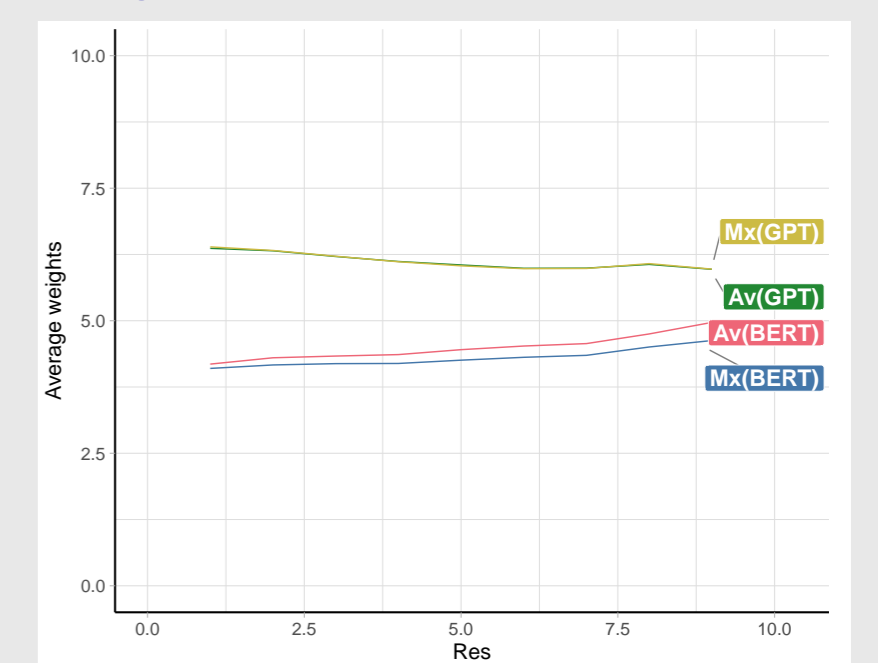


Table: Top and bottom-weighted instances with Res = 2.

Relation	Top(Av(B))	Bottom(Av(B))
SINONIMO.N.DE	beira-cairel	mesa-de-cabeceira-mesinha-de-cabeceira
SINONIMO.V.DE	cortar-foicar	empeçonhar-empeçonhar
SINONIMO.ADJ.DE	instável-lábil	infortunado-mal-afortunado
SINONIMO.ADV.DE	individualmente-particularmente	nem que a vaca tussa-nem a pau búfalo-búfalo asiático
HIPERONIMO.DE	senhora-condessa	metamorfosear-acostumar
HIPERONIMO.ACCAO.DE	mover-depor	pára-brisas-aeroplano
PARTE.DE	governo-navegação	pé.de.cor.de. açafrao-crociopede
PARTE.DE.ALGO.COM.PROP	saúde-válido	cebola-albarrá-liliáceas
MEMBRO.DE	esforço-campanha	tirar.rolha.da.garrafa-saca-rolhas
FINALIDADE.DE	tratar-procurador	galvanopuntura-agulha
FAZ.SE.COM	líquido-taleiga	empeçonhar-empeçonhamento
ACCAO.QUE.CAUSA	livrar-livre	freixo-de-espada-à-cinta-freixonita
LOCAL.ORIGEM.DE	estado-catarinense	pectinibrânquio-ter.brânquia.em.forma.de.pente
DIZ.SE.DO.QUE	consecutivo-seguir	mnemotécnico-mnemotecnia
DIZ.SE.SOBRE	austriaco-áustria	inaceitável-inadmissível
PROPRIEDADE.SEMELHANTE.A	essencial-principal	

Weights are sensitive to: relation + frequency of words

- Sequences with words that the TLM has seen more times get higher weights, even if the sequence is not semantically-coherent.
- $W(\text{cachorro ou outro animal}) > W(\text{carro ou outro animal}) > W(\text{esquilo ou outro animal})$

Answering Similarity Tests

node2vec [6] embeddings of six networks

- No weights, weighted with Res, Mx(B), Av(B), Mx(G), Av(G)
- $dimension = 64$, $window = 3$, $walks = \{10, 100, 200\}$, $walk\ length = \{10, 30\}$

Used to answer Portuguese adaptations of similarity tests

- SimLex-999, WordSim-353 [7], PT-65 [8]

No noticeable differences...

Wrap up

TLMs can be used for filtering out noisier extractions, e.g., in the automatic creation of knowledge bases

- Trend is that weights are lower for instances with long and specific arguments

No impact when computing semantic similarity

- Better-suited tasks?

Weighted PT-LKB instances available

- https://github.com/NLP-CISUC/PT-LexicalSemantics/blob/master/Relations/triplos_pesados_norm.tsv.zip

Acknowledgements

This work was financially supported by national funds through FCT, within the scope of project CISUC (UID/CEC/00326/2020) and by European Social Fund, through the Regional Operational Program Centro 2020. It is also developed in the scope of COST Action CA18209 Nexus Linguarum, supported by COST (European Cooperation in Science and Technology). <http://www.cost.eu/>.

References

- D. Downey, O. Etzioni, and S. Soderland, "A probabilistic model of redundancy in information extraction," in *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI'05*, (San Francisco, CA, USA), pp. 1034–1041, Morgan Kaufmann Publishers Inc., 2005.
- P. Cimiano and J. Wenderoth, "Automatic acquisition of ranked qualia structures from the web," in *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 888–895, 2007.
- F. Petroni, T. Rocktäschel, S. Riedel, P. Lewis, A. Bakhtin, Y. Wu, and A. Miller, "Language models as knowledge bases?," in *Proceedings of 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 2463–2473, ACL, 2019.
- H. Gonalo Oliveira, "A survey on Portuguese lexical knowledge bases: Contents, comparison and combination," *Information*, vol. 9, no. 2, 2018.
- F. Souza, R. Nogueira, and R. Lotufo, "BERTimbau: Pretrained BERT models for Brazilian Portuguese," in *Proceedings of Brazilian Conference on Intelligent Systems (BRACIS 2020)*, vol. 12319 of LNCS, pp. 403–417, Springer, 2020.
- A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 855–864, 2016.
- A. Querido, R. Carvalho, J. Rodrigues, M. Garcia, J. Silva, C. Correia, N. Rendeiro, R. V. Pereira, M. Campos, and A. Branco, "LX-LR4DistSemEval: A collection of language resources for the evaluation of distributional semantic models of Portuguese," *Revista da Associação Portuguesa de Linguística*, no. 3, pp. 265–283, 2017.
- R. Granada, C. Trojahn, and R. Vieira, "Comparing semantic relatedness between word pairs in portuguese using wikipedia," in *Computational Processing of the Portuguese Language - 11th International Conference, PROPOR 2014*, LNCS, (Cham), pp. 170–175, Springer, 2014.

^a<https://huggingface.co/pierreguillou/gpt2-small-portuguese>