Tárgyas szerkezetek elemzése egy normalizált PMI tenzor felbontásával

Makrai Márton

2020. június 3. – MILab szeminárium

Pre-neurális téma, személyes motiváció

egy disszertáció záróköve

- a `4lang` szemantikus háló (2013, 2014)
  - fogalmak fontossága a definiálás rekurzív folyamatában
  - igei bővítményszerepek = mélyesetek
- szóvektorok (2013-2018)
  - magyar analógiás kérdések, fordítás,
    lexikai relációk (antonímia, okság, hipernímia)
- a kettőt összekötni
  - igei argumentumszerkezet vizsgálata lineáris algebrai eszközökkel
  - távlati: az igei többértelműség és a tenzoros embeddingvektorok

igék, bővítmények, szelekció, szintaxis és szemantika

- Levin (1993)
- verb clustering:
  - *mond állít elmond javasol nan ír szeret közöl válaszol*
  - *encourage, invite, know, welcome, see, advise, send*
- collocation extraction (Bouma, 2009)
  - idiosyncrasy in the linguistic distribution
    - reduced syntactic modifiability
  - reduced semantic compositionality
    - a sense of the combination is habitual or even fixed

Áttekintés

Association scores

- Linguistic frequencies form *sparse* arrays
- frequencies span many orders of magnitude, *Zipf* = power law
  - Manin (2008) and Gittens, Achlioptas, and Mahoney (2017)
- for scale, sparse tensors populated with sophisticated scores
- $\log(f + 1)$ (Pennington, Socher, and Manning, 2014; Sharan and Valiant, 2017)
- three-mode PMI (Cruys, 2009; Cruys, 2011; Cruys, Poibeau, and Korhonen, 2013; Bailey, Meyer, and Aeron, 2018)
- Positivity $\Leftarrow$ sparse
  - attribute higher scores to actual co-occurrences than unattested
  - replaces negative PMI entries with zero
- two different three-variable generalization of PPMI
  - the more standard total correlation (that we still denote `pmi`) and *interaction information*
- We generalize Log Dice (Rychlý, 2008) to three axes

$$\log \frac{3f(x, y, z)}{f(x) + f(y) + f(z)} + c$$

(Positive) Pointwise Mutual Information

$$\log \frac{p(xy)}{p(x)p(y)}$$

- sparse $\Rightarrow$ positive
- one of the standard association measures, in lexicography by Church and Hanks (1990)

Higher-order PMI Cruys (2011)

- point-wise measure: $\mathbb{E}(PMI) = MI$
- two multivariate generalizations of mutual information (Shannon and Weaver, 1949) $\Rightarrow$ two multivariate point-wise variants
- *Interaction information* (McGill, 1954) and
  - based on the notion of conditional mutual information
    - inclusion and exclusion principle (szitaformula), except it has the numerator and the denominator swapped to ensure a proper set-theoretic measure

$$\log \frac{p(x,y)p(x,z)p(y,z)}{p(x,y,z)p(x)p(y)p(z)}$$

- *Total correlation* (Watanabe, 1960)
  - quantifies the amount of information that is shared among the variables

$$\log \frac{p(x,y,z)}{p(x)p(y)p(z)}$$

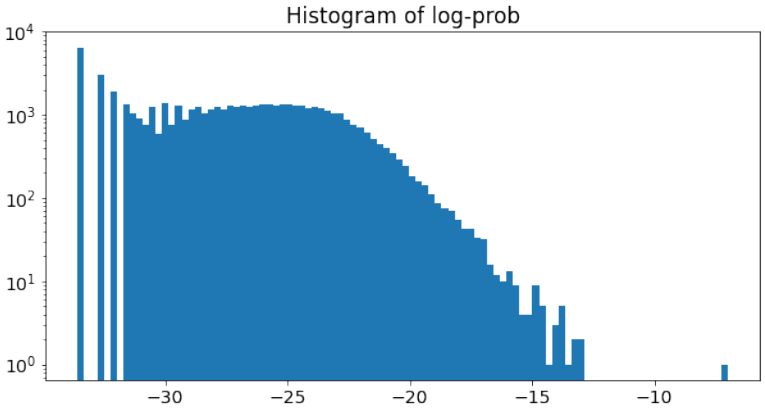
  - we just call it PMI
    - Following the literature (Villada Moirón 2005, Cruys (2009), Cruys, Poibeau, and Korhonen (2013), and Bailey, Meyer, and Aeron (2018))

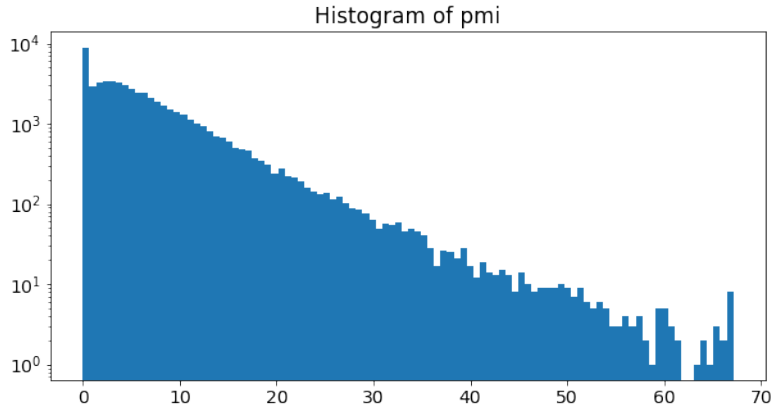
Salience and normalization

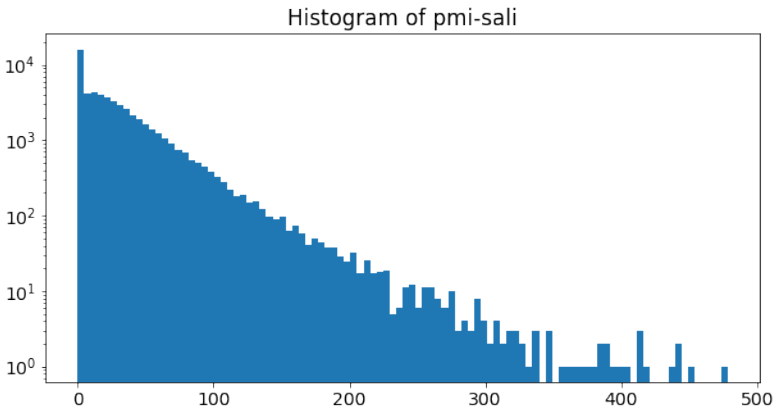
$$\log \frac{p(xy)}{p(x)p(y)}$$

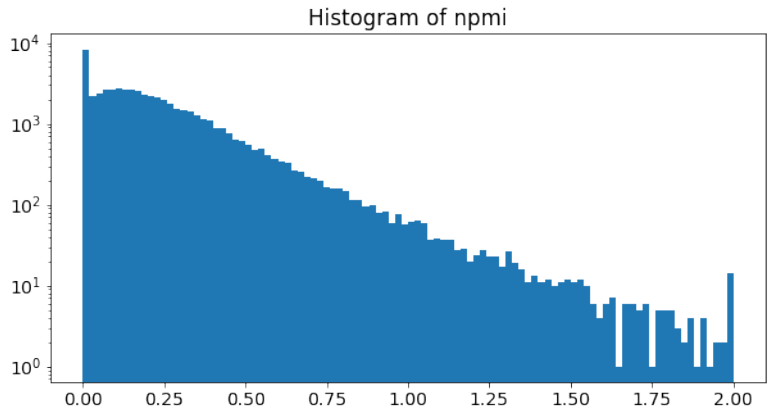
- biased towards rare events
  - Turney and Pantel (2010), Levy et al. (2015), and Zhuang et al. (2018)
- salience (Sketch Engine, Kilgarriff et al. (2004))
  - $*$=log f
- normalized variants (Bouma, 2009)
  - several ways of normalizing PMI, as the maximum value coincides with several other measures
  - $-\log p(x, y)$
  - $H(X, Y)$
  - we devide by $-\log(p(x, y, z))$

Histograms of the association score distributions I



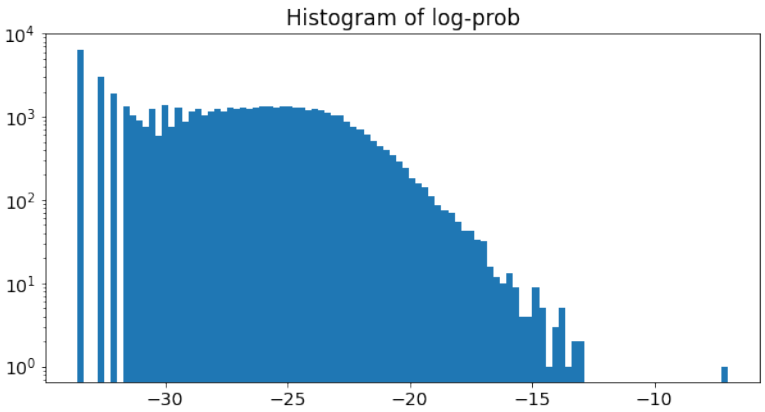


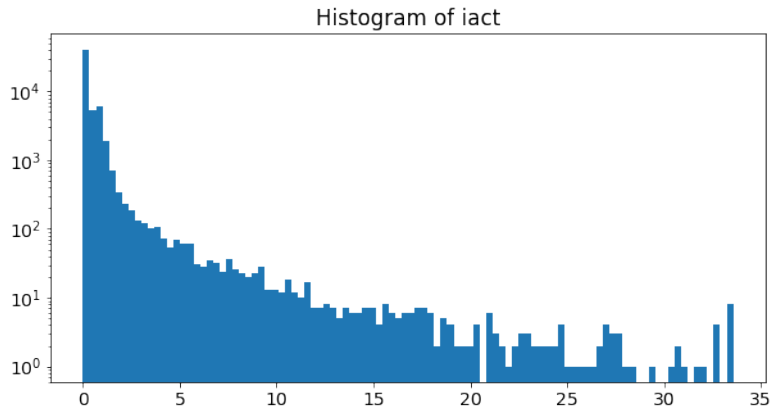


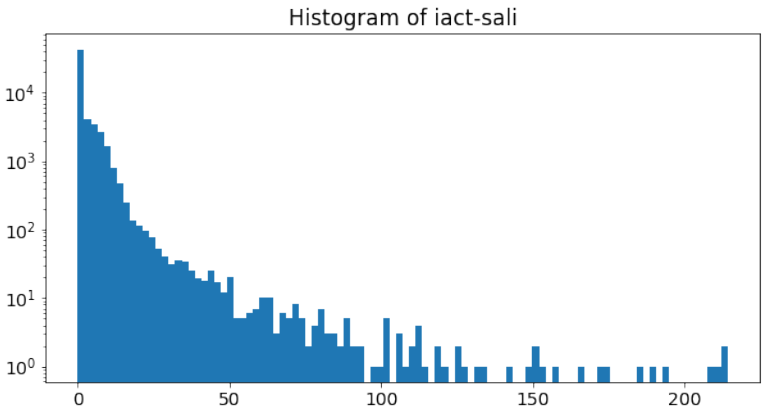


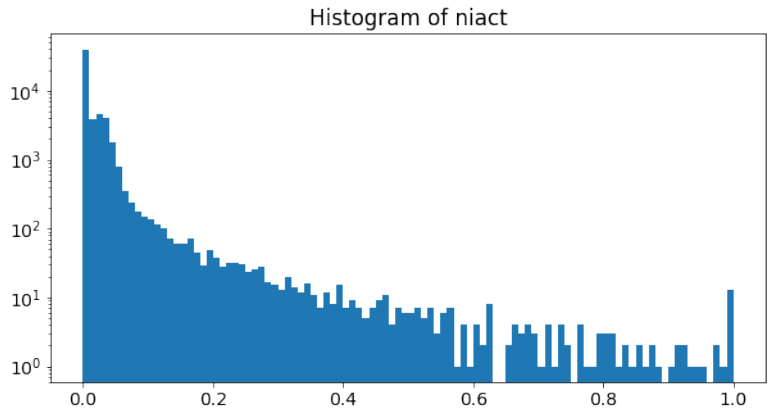
- Note the upper bound.

Histograms of the association score distributions II



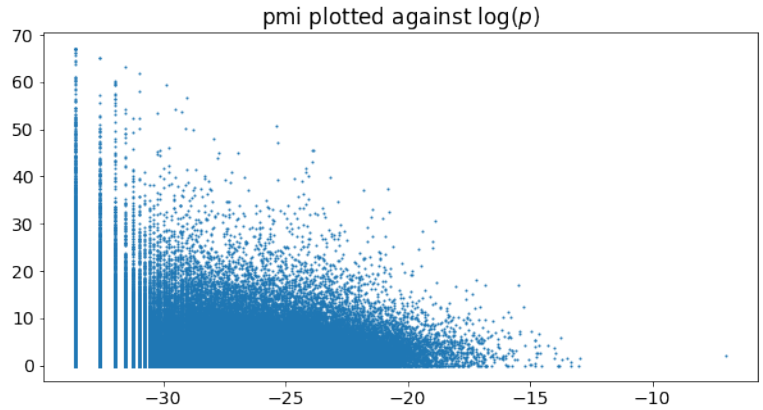


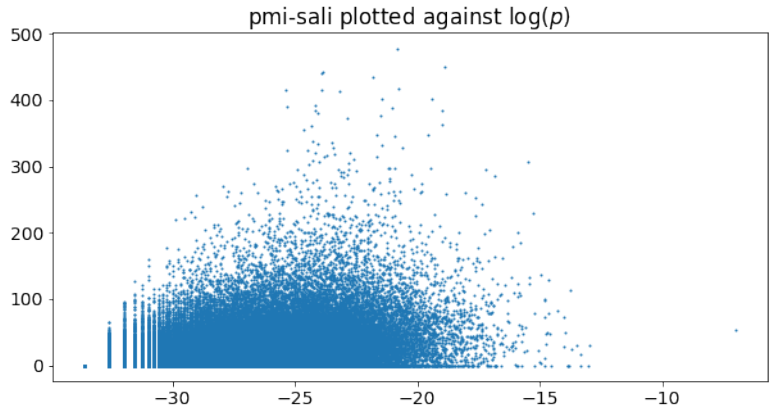


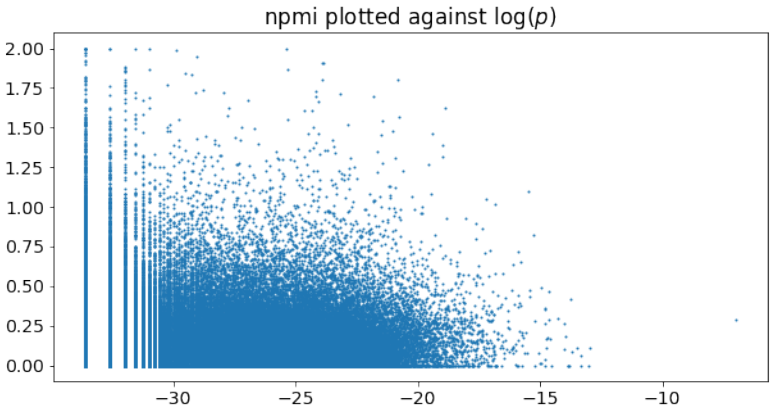


- Note the upper bound.

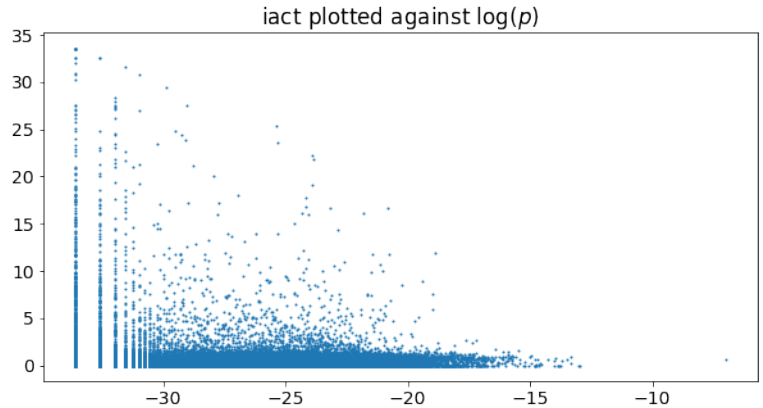
Biase towards rare events I

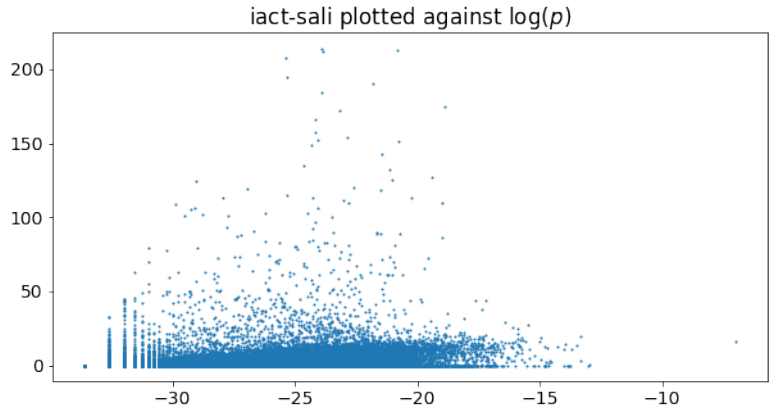


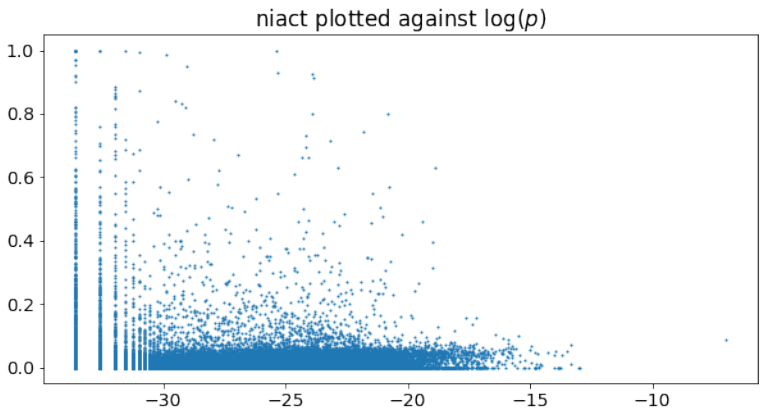




Biase towards rare events II







Áttekintés

1. Association scores
2. Low-rank decompositions and tensors
3. Experiments
   - Hiperparaméter-hangolás felügyelten: SVO szerkezetek hasonlósága
   - Látens dimenziók nézegetése
   - Az igevektorok klaszterezése
   - Szimmetria az alanyi és tárgyi vektorok között
   - Future and past work: Hungarian

Low-rank decompositions and tensors

- Turney and Pantel (2010): "four ways of looking at SVD" (i.e. LSA)
  - latent meaning
  - noise reduction
  - indirect aka. high-order co-occurrence
    - when two words appear in similar contexts
  - sparsity reduction
- two axes:
  - words and documents (LSA, Landauer and Dumais (1997))
  - words and dependency contexts (Levy and Goldberg, 2014a)
  - target and context words (standard words embeddings, Levy and Goldberg (2014b) and Pennington, Socher, and Manning (2014))
- tensor: ndim array, *mode* (data fusion), *axis*
- product, decomposition
- generalizations of low-rank matrix decomposition

Canonical Polyadic Decomposition (CPD)

Carroll and Chang (1970)

- no single generalization of the SVD concept
- CPD aka. CanDecomp, `Parallel Factor` model, rank decomp, Kruskal
- for latent parameter estimation

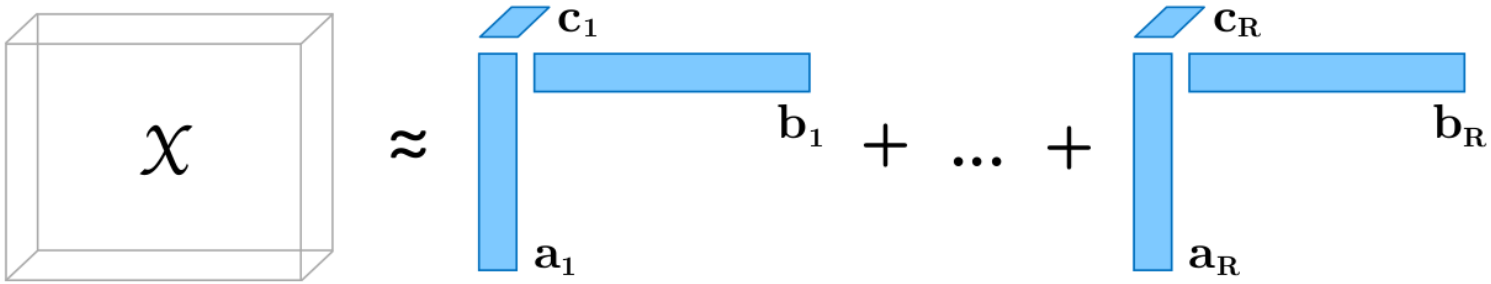


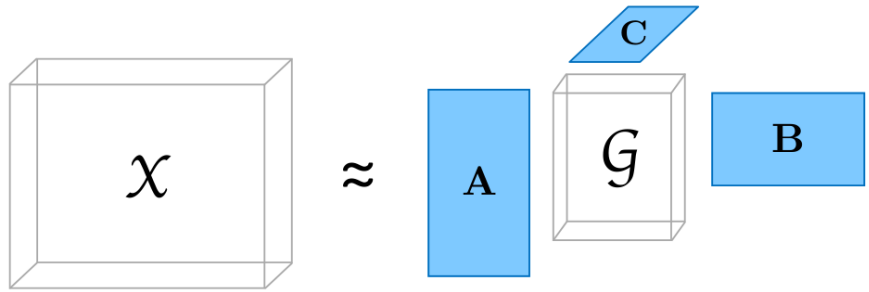
Figure: CDP expresses a tensor as a linear combination of rank-1 tensors. Figure from Rabanser, Shchur, and Günnemann (2017)

- alternating least squares, ALS
  (Carroll and Chang, 1970; Harshman, 1970)
- convergence: no guarantee, and cannot be detected in a trivial way
- Orth-ALS Sharan and Valiant (2017) improves on ALS

Tucker decomposition aka. Higher Order SVD

Tucker (1966))

- factorizes a tensor into
  - a core tensor G
  - multiplied by a matrix along each mode
- can be computed efficiently
  - In the case of



$$\text{subject} \times \text{verb} \times \text{object}$$

- rows of the three matrices contain embedding vectors of subjects, verbs, and objects (so far, the CDP and Tucker are thee same)
- entries of G determine the levels of interactions between s, v, o

Tucker decomposition and the "non-properties"

- not unique, because we can transform G without affecting the fit if we apply the inverse of that transformation to the factor matrices
- Uniqueness can be improved (Kolda and Bader, 2009) by imposing e.g. sparsity, making the elements small, or making the core "all-orthogonal", non-negativity (Zhou et al., 2015) and independence (Lahat, Adali, and Jutten, 2015)

Áttekintés

Corpus and implementational details

- occurrence counts of $\langle \text{subject}, \text{verb}, \text{object} \rangle$ triples
- in Universal Dependencies (Nivre et al., 2016) terms, `nsubj`, `ROOT`, and `dobj`, with the `upos` of the `ROOT` starting with `VB`)
- from the automatically dependency-parsed corpus DepCC (Panchenko et al., 2018),
- irrespectively of whether there were other arguments or adjuncts
- Empty fillers (subject or object) represented with a fixed symbol
- tensorly (Kossaifi et al., 2016)
- `https://github.com/makrai/verb-tensor`

Outline

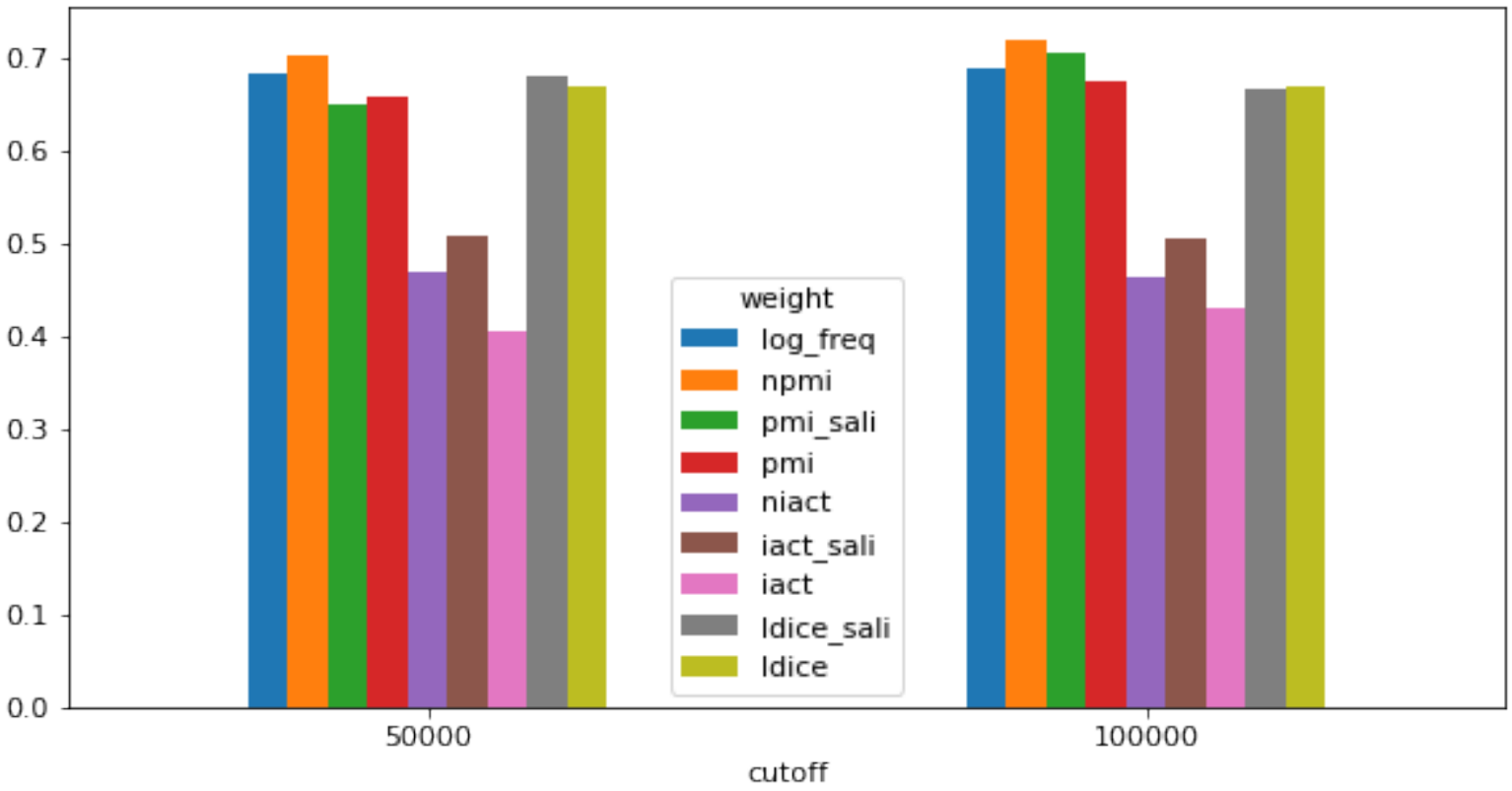
Similarity of transitive structures I

- for hyperparameter tuning (association score, rank, cutoff)
- the similarity task of SVO triples by Kartsaklis and Sadrzadeh (2014)
- SOTA = 0.760 by PAS-CLBLM (Hashimoto et al., 2014), using BNC

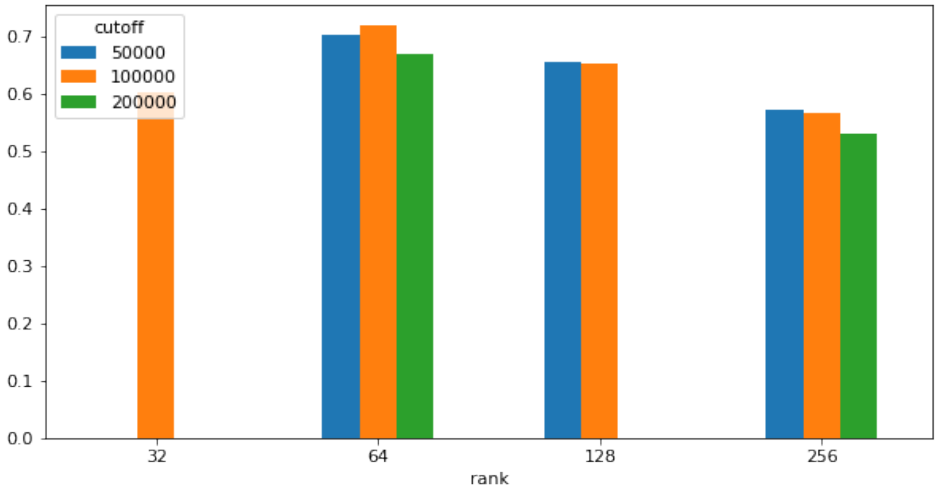
$$\tanh(w_s \odot s(j) + w_v \odot v(i) + w_o \odot o(k))$$

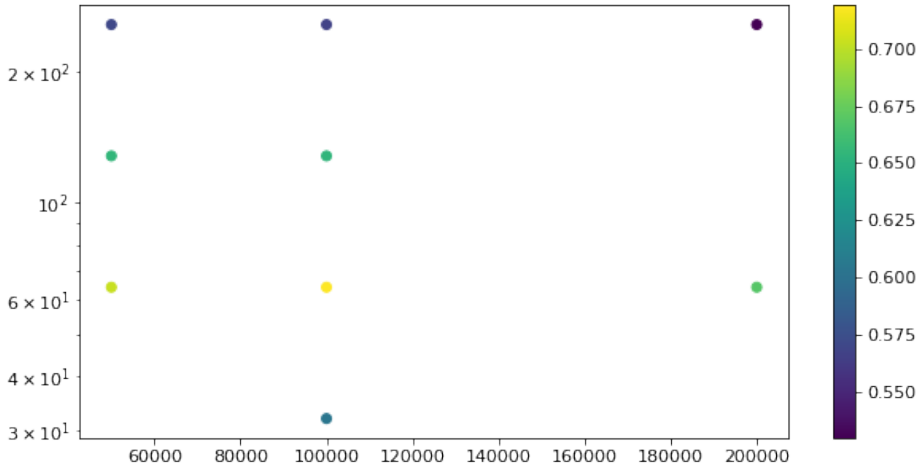
- we concatenate the embedding vectors
- (akár-negatív) Tucker-felbontásnál a legjobb kominációk: normalizált PPMI, rang = 64, cutoff = 100 k → Spearman = 0.71918
  - a konkatenált vektor dimenziója a bűvös párszázas tartományban
- Parafac (cutoff=50 k): 0.527507
- nem-negatív Tucker (pmi_sali, cutoff=300 k): 0.336537

Similarity of transitive structures II



Similarity of transitive structures III





Outline

A nem-negatív Tucker-felbontás interakciói

- 19, [-rsb-, -lsb-, disease, result, -lsb-_NNS, -RSB-, %]
  22, [-lsb-_VBD, -lsb-_VBP, -lsb-_VBZ, -lsb-_VB, -lsb-_VBN, quote]
  18, [-rsb-, -rsb-_NNS, -RSB-, -lsb-, 6, 1, 4]
- 0, [, you, we, that, he, they, who]
  0, [control, utilize, modify, use, manipulate, construct, create]
  0, [system, structure, network, object, component, function, element]
- 0, [, you, we, that, he, they, who]
  4, [educate, assist, convince, persuade, kill, inform, ask]
  3, [people, he, I, she, other, you, we]
- 0, [, you, we, that, he, they, who]
  3, [grab, pull, put, throw, touch, knock, shove]
  2, [head, ball, ass, hand, cock, face, foot]
- 8, [FIG., figure, which, -rrb-, we, table, chart]
  15, [illustrate, show, take, depict, make, use, receive]
  0, [system, structure, network, object, component, function, element]

A nem-negatív Tucker-felbontás látens dimenziói

igék

0, [control, utilize, modify, use, manipulate, construct, create]
1, [have, include, provide, offer, give, contain, require]
2, [move, regard, monitor, evaluate, control, optimize, measure]
3, [grab, pull, put, throw, touch, knock, shove]
4, [educate, assist, convince, persuade, kill, inform, ask]
5, [include, have, affect, contain, take, represent, make]
6, [increase, decrease, reduce, measure, lower, boost, calculate]
7, [provide, give, receive, offer, increase, have, require]
8, [build, expand, promote, improve, develop, enhance, maintain]
9, [provide, offer, facilitate, initiate, prevent, involve, complete]

A nem-negatív Tucker-felbontás látens dimenziói

tárgyak

0, [system, structure, network, object, component, function, element]
1, [, what, that, way, 2, [head, ball, ass, hand, cock, face, foot]
3, [people, he, I, she, other, you, we]
4, [amount, system, income, component, rate, chance, revenue]
5, [frequency, concentration, density, probability, resistance, intensity]
6, [infrastructure, production, supply, business, operation, efficiency, sale]
7, [value, datum, information, number, amount, project, development]
8, [experience, ability, understanding, capability, opportunity, strength]
9, [transmission, exposure, production, spread, transfer, flow, position]

A nem-negatív Tucker-felbontás látens dimenziói

alanyok

6, [you, , that, we, I, they, company]
7, [, system, -rrb-, we, device, report, study]
8, [FIG., figure, which, -rrb-, we, table, chart]
9, [stock, action, condition, development, sale, Jack, government]
10, [play, value, day, round, war, number, conversation]
11, [he, who, she, you, that, we, they]
12, [I, you, , he, who, they, study]
13, [area, property, system, site, factor, device, room]
14, [it, Council, Committee, board, Board, House, Commission]
15, [we, , that, system, study, site, book]

A Parafac látens dimenziói

1. ∅, that, which, this, it, what, they
   cause, trigger, create, prevent, produce, induce, mean
   industry, life, ability, productivity, health, landscape, conduct
2. activity, work, report, page, volume, book, these
   install, give, build, show, remove, pull, need
   study, work, evaluation, assessment, development, review
3. area, milk, water, region, parameter, liquid, flavor
   change, practice, learn, advance, embrace, spend, prevent
   life, water, power, moment, energy, beauty, culture
4. car, kitchen, store, door, town, i, park
   have, take, give, offer, produce, use, get
   ∅, effect, signal, value, behavior, what, characteristic
5. i, everyone, pace, browser, car, mom, anyone
   have, facilitate, promote, enhance, ensure, improve, enable
   development, growth, activity, production, movement, recovery

Outline

A recept

- `UMAP(n_components=32, metric='cosine')`
- `HDBSCAN(min_cluster_size=10, min_samples=5)`
- tapasztalat: KPI Környezeti Adaptáció, Antarktiszi áttelelő csoportok naplóiban mondatklaszterezés

cluster	verbs
-1	be, go, take, come, give, look, work, provide, help, show, inclu...
45	meet, understand, drive, carry, perform, complete, finish, ident...
5	kill, catch, shoot, feed, email, marry, wake, date, judge, bless...
20	care, listen, gon, complain, pray, dream, wan, subscribe, swear,...
27	break, push, lay, stick, roll, touch, press, suck, kick, shake, ...
1	commit, expose, separate, heal, distinguish, kid, free, devote, ...
4	run, leave, open, enter, visit, fill, close, reserve, clean, cro...
6	tell, ask, call, thank, please, join, contact, draw, become, ass...
42	add, eat, prepare, drink, spread, cook, burn, taste, smell, pour...
29	check, view, click, display, generate, update, access, search, s...
3	do, make, think, know, see, want, find, feel, love, like, hear, ...
7	remind, strike, worry, blow, inspire, bother, surprise, confuse,...
11	improve, cover, protect, represent, maintain, achieve, ensure, a...
19	live, laugh, sing, cry, smile, relax, lean, dance, breathe, star...
12	have, get, need, receive, win, lose, seek, assume, earn, gain, o...
18	start, happen, begin, continue, lead, end, occur, prove, result,...

Outline

Szimmetria az alanyi és tárgyi vektorok között

- Parafac-felbontás (a számszerű feladatban legjobb paraméterekkel)
- a háromféle vektor ugyanabban a térben van
- cos-hasonlóság a kutya alanyi és tárgyi vektora között
- különböző: ∅, *he*, *she*, *they*, *that*, *I*, *you*, *we*, *it*, *which*
- hasonló: *$*, *device*, *apparatus*, *assembly*, *portion*, *plurality*, *role*
- élő/élettelen
- e.g. *portion say*

Outline

Preverbs

- prevlex (Kalivoda, 2019)
- preverbs (igekötők) interfere with argument structure and ambiguity

preverb	verb		args				
∅	bíz	trust sth	NOM		-bAn 'in'		
(rá) 'onto'	bíz	entrust sg to sy	NOM	ACC		-rA 'onto'	
meg PERFECT	bíz	trust sy	NOM		-bAn 'in'		
meg PERFECT	bíz	entrust sy with sg	NOM	ACC			INS
el 'away'	bíz(za)	get conceited	NOM	self-ACC			

Table: Argument structure variants of the Hungarian verb *bíz(ik)* based on Szécsényi (2019).

- verb constructions from the Mazsola DB (Sass, 2015)
- *subject* ($\times$ *preverb*) $\times$ *verb* $\times$ *object* tensor
- UMAP to 10 dim (McInnes et al., 2018),
- clustering with HDBScan (McInnes, Healy, and Astels, 2017)
  - following the recommendations at readthedocs

The smallest clusters, preverbs separated

- vigyáz mosolyog csodálkozik bólint legyint bólogat
- kiált bámul szül les ordít könyörög
- halad mozog fejlődik erősöd versenyez küszködik
- jelent elérik követel tervez kíván tapasztal termel
- fekszik csap zárul telepedik fékez ereszkedik játszódik
- csatlakozik tiltakozik érdeklődik viselkedik reménykedik csalódik kételkedik
- létez létezik bizonyul érvényesül vál különbözik megerősöd
- érkezik találkozik utazik utaz távozik távoz tartózkodik játsz
- születik szerepel utal örül lakik emlékszik emlékezik haragszik gratulál
- nő csökken változik emelkedik növekedik növekszik bővül mérséklődik drágul
- rendelkezik vonatkozik gondoskodik irányul alapul minősül intézkedik módosul
- kerül lép számol következik mutatkozik kapcsolódik fejeződik avatkozik
- kinull elnull meglesz benull mar visszanull felnull lenull idenull kilesz odanull

The smallest clusters, preverbs not separated

- kínál szállít előállít behoz bevet
- szán szab kitölt meghosszabbít igazít
- megfelel számol kitér eltér törekedik kiindul
- megjelenik keletkezik hat hangzik terjed szaporodik
- születik létez létezik elkészül megszületik módosul
- oszt süt old fúj kiválik gyújt
- elfog eltávolít aláz elbocsát kivégez kihallgat
- benyújt kidolgoz terjeszt bead elvet előterjeszt beterjeszt
- menekül hozzájut megismerkedik bukkan bekapcsolódik elválik megválik
- megvásárol értékesít eljuttat beszerez megtekint továbbít lefoglal tárol
- tesz elfogad támogat használ megtesz elutasít visszautasít nehezít
- folytat kezd befejez lezár megnyit elhalaszt lebonyolít lefolytat
- magyaráz ismertet közzétesz megismétel összefoglal kommentál összegez tár
- mesél cserél megbocsát üzen szokik felhoz mer megiszik örököl kavar

Ackonwldgements

- köszönöm!
- Adalı, Berend, Borbély, Indig, Kalivoda, Kornai, Sass, Simon, Szécsényi
- anonymous reviewers
- szólj be az előadónak: mit írjak a cikkbe?

Bibliography I

Bailey, Eric, Charles Meyer, and Shuchin Aeron (2018). "Learning semantic word respresentations via tensor factorization". arXiv:1704.02686. URL: https://openreview.net/forum?id=B1kIr-WRb (cit. on pp. 5, 7).

Bouma, Gosse (2009). "Normalized (pointwise) mutual information in collocation extraction". In: *GSCL 2009: International Conference of the German Society for Computational Linguistics and Language Technology* (cit. on pp. 3, 8).

Carroll, J. D. and J. J. Chang (1970). "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition". In: *Psychometrika* 35, pp. 283–319 (cit. on p. 15).

Church, Kenneth W. and Patrick Hanks (1990). "Word association norms, mutual information, and lexicography". In: *Computational Linguistics* 16.1, pp. 22–29 (cit. on p. 6).

Cruys, Tim Van de (Mar. 2009). "A Non-negative Tensor Factorization Model for Selectional Preference Induction". In: *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics*. Athens, Greece: Association for Computational Linguistics, pp. 83–90. URL: https://www.aclweb.org/anthology/W09-0211 (cit. on pp. 5, 7).

Bibliography II

Cruys, Tim Van de (June 2011). "Two Multivariate Generalizations of Pointwise Mutual Information". In: *Proceedings of the Workshop on Distributional Semantics and Compositionality*. Portland, Oregon, USA: Association for Computational Linguistics, pp. 16–20. URL: https://www.aclweb.org/anthology/W11-1303 (cit. on pp. 5, 7).

Cruys, Tim Van de, Thierry Poibeau, and Anna Korhonen (June 2013). "A Tensor-based Factorization Model of Semantic Compositionality". In: *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Atlanta, Georgia: Association for Computational Linguistics, pp. 1142–1151. URL: https://www.aclweb.org/anthology/N13-1134 (cit. on pp. 5, 7).

Gittens, Alex, Dimitris Achlioptas, and Michael W. Mahoney (2017). "Skip-Gram – Zipf + Uniform = Vector Additivity". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada: Association for Computational Linguistics, pp. 69–76. DOI: 10.18653/v1/P17-1007. URL: http://aclweb.org/anthology/P17-1007 (cit. on p. 5).

Bibliography III

Harshman, R. A. (1970). "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis". In: *UCLA Working Papers in Phonetics* 16, pp. 1–84. URL: http://publish.uwo.ca/~harshman/wpppfac0.pdf (cit. on p. 15).

Hashimoto, Kazuma et al. (2014). "Jointly Learning Word Representations and Composition Functions Using Predicate-Argument Structures". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 1544–1555 (cit. on p. 21).

Kalivoda, Ágnes (Jan. 2019). "Véges erőforrás végtelen sok igekötős igére [A finite resource for infinitely many Hungarian particle verbs]". In: *XV. Magyar Számítógépes Nyelvészeti Konferencia*. Ed. by Gábor Berend, Gábor Gosztolya, and Veronika Vincze, pp. 331–344 (cit. on p. 36).

Kartsaklis, Dimitri and Mehrnoosh Sadrzadeh (June 2014). "A Study of Entanglement in a Categorical Framework of Natural Language". In: *The 11th workshop on Quantum Physics and Logic*. arXiv:1412.8102 (cit. on p. 21).

Kilgarriff, Adam et al. (July 2004). "Sketch engine". In: *Proceedings of Euralex*. Ed. by Geoffrey Williams and Sandra Vessier. Lorient, Université de Bretagne-Sud, Faculté des lettres et des sciences humaines, pp. 105–116 (cit. on p. 8).

Bibliography IV

Kolda, Tamara G and Brett W Bader (2009). "Tensor decompositions and applications". In: *SIAM review* 51.3, pp. 455–500 (cit. on p. 17).

Kossaifi, Jean et al. (2016). "Tensorly: Tensor learning in python". In: *Journal of Machine Learning Research (JMLR)* 20. arXiv preprint arXiv:1610.09555, pp. 1–6 (cit. on p. 19).

Lahat, Dana, Tülay Adali, and Christian Jutten (2015). "Multimodal data fusion: an overview of methods, challenges, and prospects". In: *Proceedings of the IEEE* 103.9, pp. 1449–1477 (cit. on p. 17).

Landauer, Thomas K and Susan T Dumais (1997). "A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge.". In: *Psychological review* 104.2, p. 211 (cit. on p. 14).

Levin, Beth (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press (cit. on p. 3).

Levy, Omer and Yoav Goldberg (June 2014a). "Dependency-Based Word Embeddings". In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Baltimore, Maryland: Association for Computational Linguistics, pp. 302–308. URL: http://www.aclweb.org/anthology/P14-2050 (cit. on p. 14).

Bibliography V

Levy, Omer and Yoav Goldberg (2014b). "Neural Word Embedding as Implicit Matrix Factorization". In: *Advances in Neural Information Processing Systems 27*. Ed. by Z. Ghahramani et al., pp. 2177–2185 (cit. on p. 14).

Levy, Omer et al. (2015). "Do Supervised Distributional Methods Really Learn Lexical Inference Relations?" In: *NAACL* (cit. on p. 8).

Manin, Dmitrii Y. (2008). "Zipf's Law and Avoidance of Excessive Synonymy". In: *Cognitive Science* 32 (7), pp. 1075–1098 (cit. on p. 5).

McGill, William (1954). "Multivariate information transmission". In: *Transactions of the IRE Professional Group on Information Theory* 4.4, pp. 93–111 (cit. on p. 7).

McInnes, Leland, John Healy, and Steve Astels (Mar. 2017). "hdbscan: Hierarchical density based clustering". In: *The Journal of Open Source Software* 2.11. DOI: 10.21105/joss.00205. URL: https://doi.org/10.21105%2Fjoss.00205 (cit. on p. 36).

McInnes, Leland et al. (2018). "UMAP: Uniform Manifold Approximation and Projection". In: *The Journal of Open Source Software* 3.29, p. 861 (cit. on p. 36).

Nivre, Joakim et al. (May 2016). "Universal Dependencies v1: A Multilingual Treebank Collection". In: *Proc. LREC 2016*, pp. 1659–1666 (cit. on p. 19).

Bibliography VI

Panchenko, A et al. (2018). "Building a Web-Scale Dependency-Parsed Corpus from Common Crawl". In: *Proceedings of LREC 2018*. ELRA (cit. on p. 19).

Pennington, Jeffrey, Richard Socher, and Christopher Manning (2014). "Glove: Global Vectors for Word Representation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, pp. 1532–1543. DOI: 10.3115/v1/D14-1162. URL: http://www.aclweb.org/anthology/D14-1162 (cit. on pp. 5, 14).

Rabanser, Stephan, Oleksandr Shchur, and Stephan Günnemann (Nov. 2017). "Introduction to Tensor Decompositions and their Applications in Machine Learning". arXiv:1711.10781 [stat.ML]. URL: http://arxiv.org/abs/1711.10781v1 (cit. on p. 15).

Rychlý, Pavel (2008). "A Lexicographer-Friendly Association Score". In: *Proceedings of Recent Advances in Slavonic Natural Language Processing*, pp. 6–9 (cit. on p. 5).

Bibliography VII

Sass, Bálint (2015). "28 millió szintaktikailag elemzett mondat és 500000 igei szerkezet [28 million syntactically analyzed sentences and 500 000 verb constructions in Hungarian]". In: *XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)*. Ed. by Tanács Attila, Varga Viktor, and Vincze Veronika. Szegedi Tudományegyetem Informatikai Tanszékcsoport, pp. 303–308. ISBN: 978-963-306-359-0 (cit. on p. 36).

Shannon, Claude E. and Warren W. Weaver (1949). *The Mathematical Theory of Communication*. Urbana: University of Illinois Press (cit. on p. 7).

Sharan, Vatsal and Gregory Valiant (Aug. 2017). "Orthogonalized ALS: A Theoretically Principled Tensor Decomposition Algorithm for Practical Use". In: *Proceedings of the 34th International Conference on Machine Learning, ICML 2017*, pp. 3095–3104. URL: http://proceedings.mlr.press/v70/sharan17a.html (cit. on pp. 5, 15).

Szécsényi, Tibor (Jan. 2019). "Argumentumszerkezet-variánsok korpusz alapú meghatározása [Corpus-based identification of Hungarian argument structure variants]". In: *XV. Magyar Számítógépes Nyelvészeti Konferencia*. Ed. by Gábor Berend, Gábor Gosztolya, and Veronika Vincze. Szegedi Tudományegyetem TTIK, Informatikai Intézet, pp. 315–331 (cit. on p. 36).

Tucker, Ledyard R (1966). "Some mathematical notes on three-mode factor analysis". In: *Psychometrika* 31.3, pp. 279–311 (cit. on p. 16).

Bibliography VIII

Turney, Peter D. and Patrick Pantel (2010). "From Frequency to Meaning: Vector Space Models of Semantics". In: *Journal of Artificial Intelligence Research* 37, pp. 141–188 (cit. on pp. 8, 14).

Watanabe, Satosi (1960). "Information theoretical analysis of multivariate correlation". In: *IBM Journal of research and development* 4.1, pp. 66–82 (cit. on p. 7).

Zhou, Guoxu et al. (2015). "Efficient nonnegative tucker decompositions: Algorithms and uniqueness". In: *IEEE Transactions on Image Processing* 24.12, pp. 4990–5003 (cit. on p. 17).

Zhuang, Yimeng et al. (2018). "Quantifying Context Overlap for Training Word Embeddings". In: *EMNLP* (cit. on p. 8).