Integration Means of Communication Possible with the Use of Inference knowledge bases

Hooman kashanian[1], Mohsen Parhizgar[1*], Fatemeh Shobeiri[1], Elham Dehghan niri[1]

[1]Department of Electronics and Computer Engineering, Ferdows Branch, Islamic Azad University, Ferdows, Iran.

Summary

Extract knowledge from large sources of information has become a controversial issue and therefore large knowledge bases have been established. But to deduce from this knowledge base is very time consuming and difficult. And one of the ways to overcome this construct a graph of the knowledge base and the inference is based on probabilities. That can extract elements of knowledge of the resulting graph. Our two countries in the use of such facilities to strengthen the knowledge base we have derived. First, we propose a new method for the synthesis of the knowledge base and surface text presented in a graph charts we used in the previous work. Second, how similar operation in the space of possible inferences integration and diversity characteristic that is common in the cut surface .this allows us to be effective and innovative methods, like distributing more symbolic combined with logical deduction. The proposed approach is piped to large database selection and related algorithms were run on it. And the results showed improvements in knowledge extraction of resources.

Key words:

Semantic Communication, the Possible Inference, Graph Knowledge Base.

1. Introduction

Most of what in recent years has been the creation of large knowledge bases, or by collecting by different users, such as WordNet[1], Freebase[2] and DBPedia[3] or automatically from web content or resources The collection is, like NELL [1] and YAGO [4] The knowledge base includes millions of entities and the relationships between them are real. However, even though they are too big, they still are incomplete and lacking in large parts of potential relationships between entities are common[5]. So the task of inference in the knowledge base, knowledge base is forecast new connection with a review of the case is very important. A study [6] carried out a model to calculate the amount of communication that using this computational model could achieve semantic relationships between words contained in the knowledge base WordNet using extract meaning.

Since the possible inference models in the knowledge base can be formed on the basis of the human brain. In another study[7] was able to use the knowledge base clustering theory plausible reasoning texts and the communication means effective step for the future additional research on the Extraction of knowledge and its possible inferences.

A promising technique to derive the possible inference is nothing in the knowledge base first by [8] the proposal was presented. In this way, the Path Ranking Algorithm (PRA) is called, the knowledge base is encoded as a graph and to find routes that connect the nodes origin and purpose of those relationships is used. These routes as characteristics used in a logistic regression classifier that predicts the new samples a given relationship. Using the knowledge base connections as predicate, the route can be considered as an expression Horn, and so modified PRA in[6] that can be considered as a kind of logical inference differential. One of the main disadvantages inference random walk, joint knowledge base graph is, if there is no way to connect the two nodes in the graph, PRA[9] can't predict any of communication between them. Two new techniques to make better use of set texts to derive the knowledge base we introduce. Second, we use the same vector space in a possible inference to reduce the scattering surface forms. If a route requires an edge, such as "flow through," then we kind of edge (such as "run-through") with the possibility of a vector space similarity between the two types of edge, accept. This allows us the same theory of the distribution of logical deduction symbolic space with low dispersion characteristics of the results of [6] combined.

Then in the second part of the review and the third part of construction work on the fourth graph ranking algorithms and probabilistic inference in the fifth vector space experiments in the sixth and seventh in the results.

2. Related Work

In [10] proposed In recent years many knowledge bases (KBs) have been constructed, yet there is not yet a verb resource that maps to these growing KB resources. A resource that maps verbs in different languages to KB relations would be useful for extracting facts from text into the KBs, and to aid alignment and integration of knowledge across different KBs and lan-guages. Such a

Manuscript received September 5, 2016
Manuscript revised September 20, 2016

multi-lingual verb resource would also be useful for tasks such as machine translation and machine reading. Experiments show that it effectiveness of the mapping from verbs to knowledge base relations to extract more instances for the knowledge base relations [10]. Possible inference knowledge bases by [11] was the first time. This research was developed after the use of a large and complex as a graphic presentation that it was connected to the knowledge base. Relational inference is a crucial technique for knowledge base population. The central problem in the study of relational inference is to infer unknown relations between entities from the facts given in the knowledge bases [12]. later showed that replacing labels superficial relationship with a hidden view of the relationship leads to better predict performance [9]. This result is intuitive: the characteristic that by PRA is considered a great view of the scattered and superficial relationships. PRA is a method for performing link prediction in a graph with labeled edges by computing feature matrices over node pairs in the graph. The method has a strong connection to logical inference[13] Connections "[river] in the [province] flows" and "[river] in [province] runs" almost the same meaning, and both have to be predicting Its knowledge base. However, if one of these relationships can be seen in the training data, and the other only in the test set, then none will be useful to predict. [9] tried to hide it by finding a symbolic representation of shallow relationships (such as categories) solution and replace the labels on the graph that shows hidden. This makes it likely that the level of training data and also be viewed at the time of the test and improve their performance, higher.

It has been shown that there is value in using clusters of semantically similar surface forms rather than just individual words for improving performance; in tasks such as knowledge base inference [14]. This offer is still fragile, because it is a symbolic presentation of test data and training data is prone to non-compliance. If you use too coarse a classification algorithm, the feature would be useful is if it's too soft and elegant, non-compliance there will be more. We have to solve these issues with [6] direct views of the edge during the operation possible inference we use. The two previous techniques, the most common things that we do in this paper, and we do our own research, we compare with them. [9] have adopted an approach that is very similar to the method set out in Section 2, the preprocessing set for superficial relationships. However, instead of creating a graph with nodes that represent nominal terms, they added shallow edges of the communications that are directly related to the existence of nodes in the graph. Like us, they are using the alias relationship between each pair of lips potential added that the statements by the name of a contact sample surface is shown. (Like many nominal terms Freebase) could then quickly scalability[15] using a probabilistic logic vector of features such as tunnels to reduce the knowledge base. [16] using the Path Ranking Algorithm. Nell possible to derive appropriate knowledge base extract. [17] using the random walk method and language of neural networks to calculate semantic relatedness between texts have taken the appropriate steps.

Other related work. Also related to the present work is recent research on programming languages for probabilistic logic [18]. This work, called ProPPR, uses random walks to locally ground a query in a small graph before performing propositional inference over the grounded representation. In some sense this technique is like a recursive version of PRA, allowing for more complex inferences than a single iteration of PRA can make. However, this technique has not yet been extended to work with large text corpora, and it does not yet appear to be scalable enough to handle the large graphs that we use in this work. How best to incorporate the work presented in this paper with ProPPR is an open, and very interesting, question.

These commonsense resources could easily be incorporated into the graphs we use for performing random walk inference. Lines of research that seek to incorporate distributional semantics into traditional natural language processing tasks, such as parsing [19], named entity recognition [20], and sentiment analysis [21], are also related to what we present in this paper. While our task is quite different from these prior works, we also aim to combine distributional semantics with more traditional methods (in our case, symbolic logical inference), and we take inspiration from these methods.

3. Build a graph

Our method for deriving knowledge base, which is described in section 3, randomly walking in a graph provides the specifications for the logistic regression classifier achieved. Before you precisely explain this technique, we first describe how a graph $= (\mathrm{N}, \varepsilon, \mathrm{R})\,\mathcal{G}$ Knowledge Base and a set of samples from a set of samples extracted from a superficial relationship together. Create a graph of the knowledge base is simple and convenient. (n_1, r, n_2)connects .Something that is a little more difficult is how to make the graph of a graph and how it is connected to the graphic knowledge base. We describe methods to each of them.

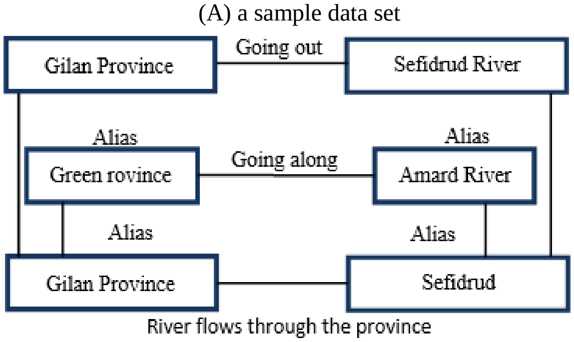
To create a graph of a set, the first set preprocessing we obtain a set of superficial relationships, such as those by extracting information systems such as OLLIE extracted [22]. noun phrase is indicative and there is no edge between the entities nominal terms. Our two graphs together with the use of pseudonyms in our knowledge base in which entities linked to potential references noun phrase [23]. Each phrase in the graph surface communication nodes connected entities, and the name can be mentioned in terms of the knowledge base [24]. As reported by [25] was carried out, the edges output connector not exist, but instead expressed the opinion that the phrase could refer to the existence of the knowledge base.

The existence of a connection between the nodes is certainly stronger connection name and attribute nodes is possible, but it requires more processing and display the graph is much larger. We offer a variety of edge that makes it possible to optionally have an associated vector that ideally what kind of meanings obtain edge.

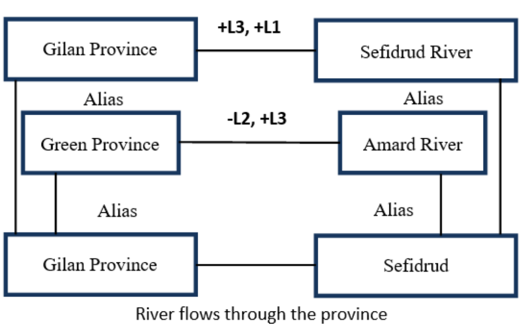
We have a relationship Persian Farsi to compare notes on experiments Word datasets we use. The graph on the sub-base structure used in our test shows the level of knowledge and communication. Note that Figure 1 and 1-A as analog-to irregular use of graphs shown in previous work.

«Sefidrud, River Flows Throughout the Province, Gilan »
«**Gilan**, Alias, Gilan»
«**Gilan**, Alias, Green Province»
« **Sefidrud**, Alias, Gilan River»
« **Sefidrud**, Alias, Amard River»
Shallow relationships :
«Amard River, Going along, Green Province»
«Gilan River, Going out, Gilan Province»
Integration :
«Going along» : [0.9 , **-0.1** , 0.2]
«Going out» : [0.8, **-0.3**, 0.1]

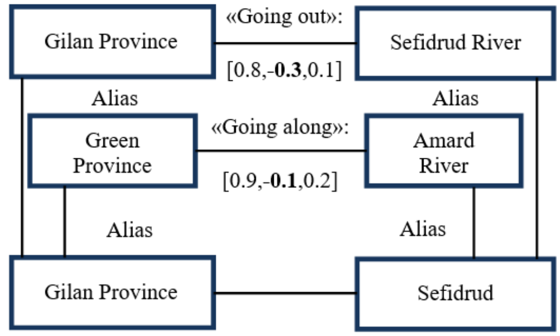
(A) a sample data set



(B) an example of a graph that combines a knowledge base and communications.



(C) a sample graph that alternative communication of the surface with a classification label, the sample was done by [9].



(D) a sample graph that as it is presented in this paper, the display vector space on the edge of the surface.

Fig 1. Structure of the graph so that the test sample used in this article.

KB uses a graph that only the edges of the graph is simply a collection that only includes edge "River flows through the province" and not shown.

4. The Path Ranking Algorithm

Rating long edge [6] did graph. This algorithm has been improved [7] with a brief overview on the start, and after its reform in the algorithm offering that allows us. In conclusion, we may integrate semantic connections. The algorithm can be used as a way to extract local graph structure to create non-linear combinations of characteristics to be considered a predictive model. A characteristic combination creates pairs of nodes in a graph, and then logistic regression analysis to classify the pairs of nodes as those used in the special relationship. Formally, to deliver a graph “g” with nodes N, edgesε, and tags edge R and a set of pairs of nodes$(s_i, t_i) \in D$, you can create a relationship matrix where the rows and columns corresponding to the node pairs tags are corresponding edge. Gives .because of the very large feature space (the set of all possible sequences of edge labels, with cardinality$\sum_{i=1}^{l} |R|^i$, which is assumed to be a maximum along the border l), first stage that must be done is select the parameters, using probabilistic inference is made in the graph. The second step is to calculate semantic relatedness that each cell in the matrix characterized by a possible deduction limit is calculated from the following attributes associated with it.

5. The vector-based space of probable inference

Modifications months to fully calculate the characteristic is limited to that described above is a selection of properties, (Find potentially useful sequences of edge) is considered as normal from the edge of the symbolic use. In more formal, consider the type of route π .Keep in mind that π is a sequence of edges<e1, e2, …,el>where l along the way, we from πithe sequence will be used to show the type of edge. For calculating characteristic values, in some nodes and edges begin to comply with ithtype sequence is completed and to achieve a node. In particular, if a node n random walking out with a variety of edge is "m"$\{e_1, e_2, \dots, e_m\}$, then the type of edge that sets it in accordance with the πiselected, Then randomly from all external edges it selects the type. If there are no matches in the series, then the inference is possible from the start node to start. We'll modify any type of edge. When the random walking on the edge of the output node n and m are$\{e_1, e_2, \dots, e_m\}$، then instead of just the type of edge that is in accordance withπI to choose, it is possible that we are walking, your lips select nearlyπivector space. For any edge node n, we lips following are likely:

$$p(e_j|\pi_i) \propto \exp\left(\beta \times u(e_j).u(\pi_i)\right), \qquad \forall_j, 1 \le j \le m$$

That u (.) function that displays a lip operation returns and β is a parameter that determines the similarity spikiness how much weight to give the vector space. When β to be extremely close, then $\{e_1, e_2, \dots, e_m\}$normal view a delta function on the nearest edge of the πiestimates. Ifπiedges set out, then the algorithm method [6] converges. However, if πiset on the edge of the output is not in a knot and all kinds of differentπiedge, then this algorithm (where β is extremely close) resulted in a very uniform distribution of the edge of the node and no way to start walking again is no accident. To improve the restart behavior [6], we've added a parameter to restart the"α" and other value before normalization added to the total distribution: $p(restart|\pi_i) \propto \exp(\beta * \propto)$

When this type of restart is selected, random walking again be started from πI the source node, will follow. α is the set value of the maximum similarity between vectors of edge (other times) is larger, and β is set at infinity, the algorithm to accurately copy the sample [6] is All of the edges are showing a vector space: Its nickname can have a significant vector, and we do not use vectors to represent the knowledge base links, because doing so is not useful in practice (makes intuitive sense: the knowledge base shows hidden relationships is). Random nights when we do step in, if πI not taking any view, the algorithm [6] to come to a new edge to choose.

Here note that when you work with vector spaces, it is natural to try to classify the vectors to reduce the parameter space. πfeature in our model is any path, and if the two types of the way just in a different edge, then the outcome is not necessarily characteristic values for both directions are equal. It is reasonable that a simple classification algorithm during the course of implementation to reduce the number of duplicated features that improve performance. However, this was not practical, we try to use these vectors to classify the functions of when distinct categories is not used.

6. Tests

We selected the characteristic features of both the PRA calculations using "GraphChi" do a single graph-processing machine is effective [26]. We MALLET logistic regression with adjustment of L1 and L2 use. For negative evidence, from a closed world hypothesis that each pair (source, target) found during the characteristic calculation if not presented as a positive example, as a negative sample were treated. We use a manual network search to validate the training data described by the parameters we set ourselves. Regulatory parameters that we set the parameters L1 and L2, the random walking has to select the parameters and calculations for the PRA and spikiness parameters and restart the vector space for steps to be taken.

Table 1:Statistics of the data used in our experi-ments.

	NELL
Entities	1.2M
Relation instances	3.4M
Total relation types	520
Relation types tested	10
Avg. instances/relation	810
SVO triples used	404k

6.1 Data

NELL ran our tests in the knowledge base. View this database is shown in Table 1. Its ten NELL chose us to test my methods. Nell communications manually as relationships with the largest number of samples were selected that were reasonable accuracy The NELL relations were hand-selected as the relations with the largest number of known instances that had a reasonable precision (the NELL KB is automati-cally created, and some relations have low preci-sion). We split the known instances of these rela-tions into 75% training and 25% testing, giving on average about 650 training samples and 160 samples to test for any relationship we have provided. First, we filter our 4215 communication based on two criteria: the number of instances of the relationship should be between 1000 and 10000, and must not be in any medium. When you chose us, all the samples for each possible connection in the data were kept SVO. this remaining approximately 200 cases for every relationship that we again compared to 75% -25% to the training and test sets were divided.

6.2. Methods

Compare the ways we did that was associated with the graphs shown in Figure 1. Knowledge Base algorithm [6] method only in relation to the knowledge base .Knowledge Base + SVO Community level will add to the graph (Fig. 1-B). This means that almost by [8] presented, are comparable, although, as described in Section 5, there are significant differences in the performance graph. Base + SVO classified knowledge of the ways of [9] followed, but the X-ray structure used in this article is provided. Our approach is SVO knowledge base vector.

6.3. Evaluation

As the scales of assessment, the average precision (MAP) and mutual rank average (MRR), use and extraction of research relevant to assess the performance, conformance and adapted. We tried a couple significant permutation test. The results in Table 2 and Table 3 are shown. In Table 2 we mean accuracy for each communication NELL KB tested at the show.

Table 3: accuracy mean that the knowledge base is not connected to any knowledge base pipe is tested. The procedure is shown in bold on any relationship.

Relation	KB	KB + SVO	KB + Clustered SVO	KB + Vector SVO
Actor Starred In Movie	0.000	0.042	0.052	**0.051**
Athlete Plays For Team	0.120	0.269	**0.597**	0.592
City Located In Country	0.129	0.174	0.262	**0.388**
Journalist Writes For Publication	0.238	0.282	0.300	**0.329**
River Flows Through City	0.020	0.011	**0.162**	0.096
Sports Team Position For Sport	0.267	0.246	0.198	0.130
Stadium Located In City	0.110	0.340	0.288	**0.371**
State Has Lake	0.020	0.000	**0.026**	0.010
Team Plays In League	0.988	**0.989**	0.959	0.946
Writer Wrote Book	0.187	**0.234**	0.199	0.210

Table 2: Results on the NELL knowledge base.The bolded line is significantly better than all otherresults with p<0:025.

Method	MAP	MRP
KB	0.43	0.72
KB + SVO	0.256	0.810
KB + Clustered SVO	0.345	0.859
KB + Vector SVO	**0.430**	**0.910**

6.4. Topics

From this table we can see that the knowledge base + vector SVO (The method presented in this paper) significantly than previous methods in MAP and the MRR better.

We believe that this is due to a reduction in the dispersion characteristic that is possible using the vector space rather than symbolic representations (as that is theonly real difference between KB + Clustered SVO and KB + Vector SVO) to This [6] allows better use of the path is found in the training data. When we look at the results in Table 3 individual connections, see Knowledge Base + Vector SVO in most relationships is better than other methods.

The results can be seen that the average accuracy seems a little low for all test methods. This is because the MAP as possible accurate predictions accurately calculate the ranking list, which is not included in the list if the prediction is correct, the accuracy of the numbers is 0. In other words, examples of great relevance in our test suite that has been randomly drawn from the knowledge base and this calls weak to scale MAP damage. MAP is the high accuracy of prediction for each connection judgment, assures us that the main issue here is calling for MRR is reasonably high, especially in the knowledge base Nell.

As further evidence, if we accurately calculate the mean for each search node (instead of every relationship), and searches for they did not anticipate any system, we remove from the MAP domain (knowledge base) to 29 (Knowledge Base vector SVO) 45. In Nell (about 30% of the searches did not predict) and .40 (KB) is, (that 21% of the searches did not predict). So our method improves the performance of MAP in time are calculated this way.

But this is not an entirely fair scale, so the standard MAP is used to present the results of my core.

7. Conclusion

Using Graph construction of a more coherent knowledge base could create a graph in which the relationship between the semantic level and knowledge base to improve the text. In the next phase, we showed how probabilistic inference can be combined with the semantic relationship vector space, so that we can reduce the number and variety of surface features of texts. This allows us to derive similar distribution with innovative and efficient way to fuse symbolic logic. Tested on pipe connections with a lot of the NELL knowledge base, we have shown that our method significantly from previous studies about the inference knowledge base better.

References

[1] Carlson, A., Betteridge, J., Kisiel, B., Burr Settles, Estevam R Hruschka Jr. & Tom M Mitchell. 2010. Toward an architecture for never- ending language learning. In AAAI.

[2] Bollacker, K,. Evans, C,. Paritosh, P,. Sturge, T. & Taylor, J. 2008. Freebase: a col- laboratively created graph database for structuring human knowledge. In Proceedings of SIGMOD.

[3] N. Mendes, P,. Jakob, M,. Bizer C. 2012. Dbpedia for nlp: A multilingual cross-domain knowledge base. In Proceedings of the Eighth In- ternational Conference on Language Resources and Evaluation (LREC'12).

[4] Suchanek, F, M,. Kasneci, G. & Weikum, G,. 2007. Yago: a core of semantic knowl- edge. In Proceedings of WWW.

[5] West, R., Gabrilovich, E., Murphy, K. Sun, S. Gupta, R. & Dekang Lin. 2014. Knowledge base completion via search-based ques- tion answering. In WWW.

[6] Kashanian, H., Mahmoudi, F. & Masoomy, B. Introduced a measure to improve the semantic relationships between vocabulary. Eighteenth Conference of Computer Society of Iran, Sharif University of Technology,1391.

[7] Kashanian, H. Improved clustering texts using plausible Argument. The first conference of computer and information technology, Dorood, 1391

[8] Lao, N, & Yang Wang, W., Cohen. 2010. Relational retrieval using a combination of path-constrained random walks. Machine Learning.

[9] Gardner, M., Pratim, P., Kisiel, B. & Mitchell, T. 2013. Improving learning and inference in a large knowledge-base using latent syntactic cues. In Proceedings of EMNLP. Associa- tion for Computational Linguistics.

[10] D. T. Wijaya and T. M. Mitchell, "Mapping Verbs In Different Languages to Knowledge Base Relations using Web Text as Interlingua," 2016.

[11] Lao, N,. & Cohen, W,. 2010. Relational re- trieval using a combination of path-constrained random walks. Machine learning, 81(1):53–67.

[12] Q. Liu, L. Jiang, M. Han, Y. Liu, and Z. Qin, "Hierarchical Random Walk Inference in Knowledge Graphs," in Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, 2016, pp. 445-454.

[13] M. Gardner and T. Mitchell, "Efficient and expressive knowledge base completion using subgraph feature extraction," in Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 1488-1498.

[14] W. Ammar, G. Mulcaire, Y. Tsvetkov, G. Lample, C. Dyer, and N. A. Smith, "Massively multilingual word embeddings," arXiv preprint arXiv:1602.01925, 2016.

[15] Miao. F., Qiang, Z. & Thomas, F, Zh . 27 Mar 2015. Learning Embedding Representations for Knowledge Inference on Imperfect and Incomplete Repositories, arXiv:1503.08155v1 [cs.AI].

[16] Lao, N., Mitchell,T,. Yang Wang. & W. Cohen.2015. RandomWalk Inference and Learning in A Large Scale Knowledge Base, Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213.

[17] Goikoetxea, J., Soroa A. & Agirre, E. 2015. RandomWalks and Neural Network Language Models on Knowledge Bases, IXA NLP Group University of the Basque Country Donostia, Basque Country, Human Language Technologies: The Annual Conference of the North American Chapter of the ACL, pages 1434–1439, Denver, Colorado,

[18] William Yang Wang, Kathryn Mazaitis, and William W. Cohen. 2013. Programming with personalized pagerank: A locally groundable first-order probabilistic logic. In Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management, CIKM '13, pages 2129–2138, New York, NY, USA. ACM.

[19] Socher, R., Bauer, J., Christopher D Manning, & Andrew Y Ng. 2013a. Parsing with compo- sitional vector grammars. In In Proceedings of the ACL conference. Citeseer.

[20] Alexandre Passos, Vineet Kumar, and Andrew Mc-Callum. 2014. Lexicon infused phrase embed-dings for named entity resolution. arXiv preprint arXiv:1404.5367.

[21] Socher, R., Perelygin, A., Jean Y Wu,. & Chuang, j,. Christopher D Manning, Andrew Y Ng, and Christopher Potts. 2013b. Recursive deep mod- els for semantic compositionality over a sentiment treebank. In Proceedings of EMNLP.

[22] Kachites McCallum, A., 2002. Mallet: A ma- chine learning for language toolkit.

[23] Pratim Talukdar, P., Wijaya, D., & Mitchell, T. 2012. Acquiring temporal constraints be- tween relations. In Proceedings of the 21st ACM international conference on Information and knowl- edge management, pages 992–1001. ACM.

[24] Riedel, S,. Yao, L,. McCallum, A. & Marlin, M. 2013. Relation extraction with matrix factorization and universal schemas. In Pro- ceedings of NAACL-HLT.

[25] Lao, N., Subramanya, M., Pereira, f. & WCohen, W. 2012. Reading the web with learned syntactic-semantic inference rules. In Pro- ceedings of EMNLP-CoNLL.

[26] Kyrola, A., Blelloch, G., Guestrin, & C. 2012. Graphchi: Large-scale graph computation on just a pc. In Proceedings of the 10th USENIX Symposium on Operating Systems Design and Im- plementation (OSDI), pages 31–46.