TaxFree: a Visualization Tool for Candidate-free Taxonomy Enrichment

Irina Nikishina[1] , **Ivan Andrianov, Alsu Vakhitova**[2]**, and Alexander Panchenko**[2]

[1]Universität Hamburg

[2]Skolkovo Institute of Science and Technology

irina.nikishina@uni-hamburg.de, i.a.andrianov@gmail.com, alsu.vakhitova@gmail.com,

a.panchenko@skoltech.ru

Abstract

Taxonomies are widely used in a various number of downstream NLP tasks and, therefore, should be kept up-to-date. In this paper, we present **TaxFree**, an open source system for taxonomy visualisation and automatic Taxonomy Enrichment without pre-defined candidates on the example of WordNet-3.0. As opposed to the traditional task formulation (where the list of new words is provided beforehand), we provide an approach for automatic extension of a taxonomy using a large pre-trained language model. As an advantage to the existing visualisation tools of WordNet, **TaxFree** also integrates graphic representations of synsets from ImageNet. Such visualisation tool can be used for both updating taxonomies and inspecting them for the required modifications.

1 Introduction

In this paper, we focus on visualisation of taxonomic structures which are quite relevant for many Natural Language Processing (NLP) tasks, e.g. lexical entailment (Herrera et al., 2005) and entity linking (Moro and Navigli, 2015; Sevgili et al., 2022). Taxonomies are tree-like structures where words are considered as nodes (synsets) and the edges are the relations between them. Such kind of relationship is called a hypo-hypernym relationship. For instance, let us consider two words: "apple" and "fruit". The former word is hyponym ("child") to the latter and the latter is hypernym ("parent") to the former.

However, taxonomies are hard to maintain, while the manual taxonomy annotation process is very expensive and time-consuming. Moreover, it requires expertise in the field. The process of selection new words is even more challenging for the large existing taxonomies like WordNet (Miller, 1995), as most existing words already present there. We expect that the large pre-trained language models such as BERT (Devlin et al., 2019) and GPT

(Brown et al., 2020) could be useful in the task, as they are pre-trained on large-scale corpora. It has been proven that language models possess syntactic, semantic and word knowledge which could be applied for further language inference (Radford et al., 2019).

In this demo we demonstrate how the existing taxonomy can be enriched automatically without predefined candidates on WordNet-3.0. Figure 1 demonstrates the candidate-free task setting where the node "milk.n.01" ("n" stands for "noun", "01" denotes the ordinal number of word sense, the standard notation for the synset in WordNet) is extended with multiple synsets of different types of milk: "'low-fat milk", "chocolate milk", "dry milk", etc.
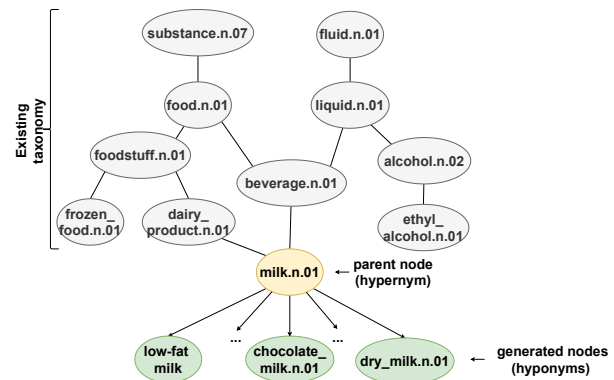

Figure 1: Candidate-free Taxonomy Enrichment task solved by **TaxFree**. The node "milk.n.01" is enriched with hyponyms (different types of milk).

TaxFree (see example of the visualisation page in Figure 2) is an open source, web-based visualisation and enrichment tool for taxonomies. We demonstrate the capacities of **TaxFree** on WordNet-3.0 with support for visual representation of WordNet synsets using ImageNet (Deng et al., 2009). The tool demonstrates an approach of automatic taxonomy graph extension by predicting new nodes (synsets) using BERT (Devlin et al., 2019) as a pretrained language model. It allows the user to search

in WordNet by words (lemma) or nodes (synset), to visualise the context of the query word (with $hop = 2$), to generate new leaf nodes or nodes between the two existing ones. To generate new nodes we apply the Cross-modal Contextualized Hidden State Projection Method (Nikishina et al., 2022b). The approach includes several stages: (i) learning embeddings of the WordNet taxonomy and new synsets at the required places that we want to predict (ii) projecting all graph embeddings into the hidden states space of BERT, and (iii) decoding them back to text candidates.

Thus, the contribution of this demonstration system is three-fold:

- Firstly, it allows users to search and visualize the query node within its context (on the example of the English WordNet-3.0);

- Secondly, it allows users to automatically extend the existing taxonomy without predefined (or manually defined) nodes using Cross-modal Contextualized Hidden State Projection Method (Nikishina et al., 2022b);

- Thirdly, it integrates ImageNet representations to the WordNet synset description card.

The link to the demo is as follows: **https://taxgen.ltdemos.informatik.uni-hamburg.de**.

The code link: **https://github.com/skoltech-nlp/taxgen-demo**. Link to the screencast video demonstrating the system: **https://youtu.be/GF2AVlnWGag**.

2 Related Work

In this section we review the existing approaches for Taxonomy Enrichment as well as the existing tools for taxonomy visualization.

2.1 Taxonomy Enrichment

To the best of our knowledge, there are no existing approaches for Candidate-free Taxonomy Enrichment. All methods require the list of pre-defined words to be added to the taxonomy. There exist several recent papers on Taxonomy Enrichment that make use of word vector representations and/or large pre-trained language models. For instance, (Nikishina et al., 2022a) present an approach applying numerous of text and graph embeddings as well as their combinations; (Takeoka et al., 2021) solves the same problem, but for the low-resource

scenario using BERT-based classifier and Hearst Patterns (Roller et al., 2018); (Cho et al., 2020) regard the taxonomy enrichment task as a sequence-to-sequence problem and train an LSTM model on the WordNet data. A detailed overview of other taxonomy-related tasks is presented in (Jurgens and Pilehvar, 2016; Nikishina et al., 2022a).

2.2 Taxonomy Visualisation

Plenty of tools are available for generic visualisation of networks like Gephi (Bastian et al., 2009), GraphX (Gonzalez et al., 2014), D3[1], GraphViz (Ellson et al., 2001). At the same time, there might be found some tools specific for the visualisation of wordnets, which are not available from the original interface of WordNet. For example, (Collins, 2006) is one of the earliest papers that present a design paradigm. Visualisation from (Kamps and Marx, 2002) demonstrates not only the relations between synsets, but also denotes lemmas as graph nodes. WordNet Atlas (Abrate and Bacciu, 2012) is designed for "users like computer scientists that are not familiar with computational linguistics and/or WordNet.

Another paper (Giabelli et al., 2020) presents NEO: a tool for Taxonomy Enrichment that allows to enhance the standard occupation and skill taxonomy. The authors collect the terms from the Online Job Vacancies corpus and add them to the taxonomy automatically. Another visualisation of lexical graphs based on WordNet[2] is very similar to the one we present in our paper, however, it is lemma-based and does not allow dynamic extension of the graph using taxonomy enrichment technology.

3 Candidate-free Taxonomy Enrichment

This section presents Candidate-free Taxonomy Enrichment — the problem of new word prediction in order to enhance the existing taxonomy. We briefly describe the Cross-modal Contextualized Hidden State Projection Method (CHSP) used in the demo and the results obtained on the dataset based on WordNet-3.0.

3.1 Task Formulation

Formally, the task of candidate-free taxonomy enrichment may be formulated as follows: given taxonomy T and the position of the synset $s_i \in S$

---

[1]https://github.com/d3/d3
[2]https://github.com/aliiae/lexical-graph

TaxFree: WordNet3.0 visualization for candidate-free taxonomy enrichment

Back to root   Reset graph   dog.n.01   Move to   Download image

dog.n.01

dog,domestic dog,Canis familiaris

a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds
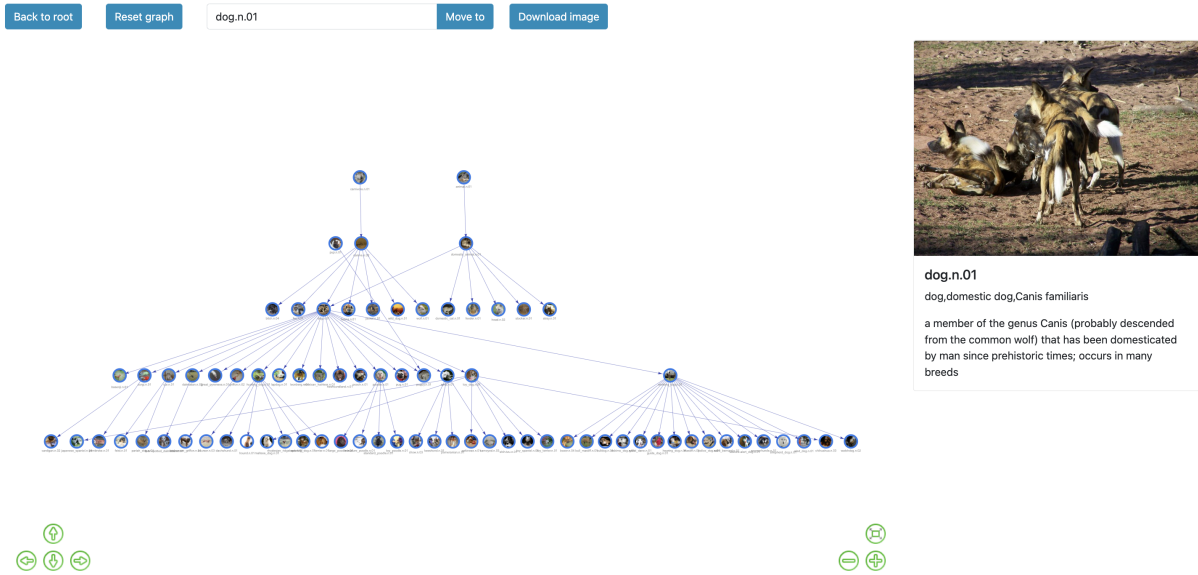
Figure 2: Visualisation example for the node "dog.n.01"

TaxFree: WordNet3.0 visualization for candidate-free taxonomy enrichment

Back to root   Reset graph   word   Move to   Download image

entity.n.01

entity

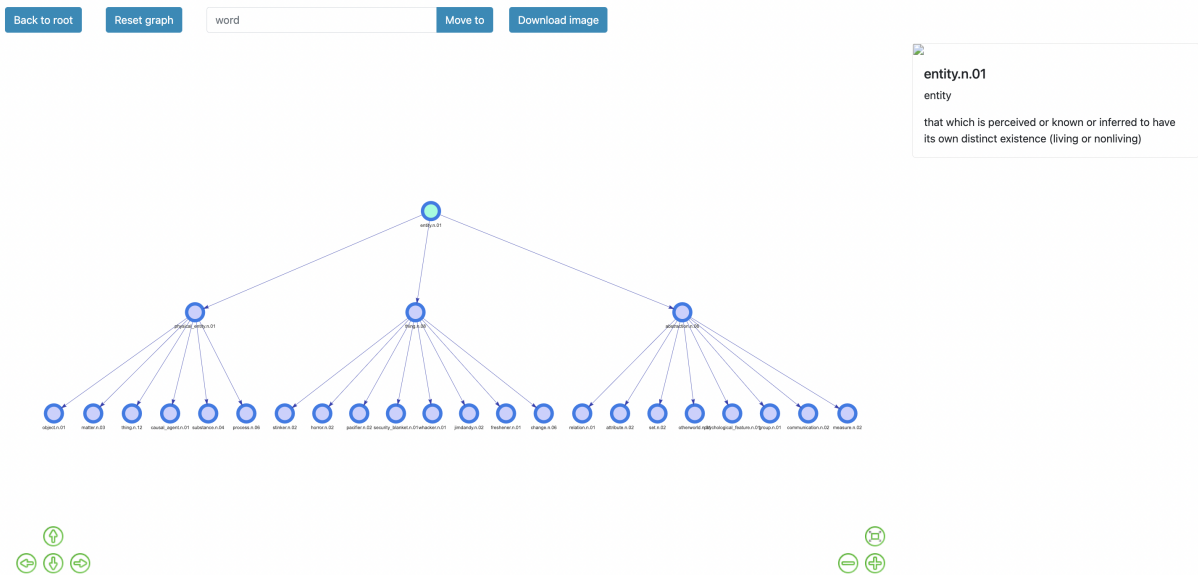that which is perceived or known or inferred to have its own distinct existence (living or nonliving)

Figure 3: Default search page with the root "entity.n.01"

in this taxonomy, the task aims at predicting hyponyms $H_s \subseteq \{h_{s1}, ..., h_{sn}\}$ such that $H_s \notin E$, where E are edges in the taxonomy. Such formulation allows us to avoid the need of pre-supplied candidates making the task more challenging yet realistic. It might be expected that the information about new words could be already present in the large pre-trained networks.

3.2 Method Description

The Contextualized Hidden State Projection Method (CHSP) is a graph-based BERT architecture introduced by Nikishina et al. (2022b) that makes use of both node and text embeddings. Figure 4 demonstrates the overall architecture of the approach that we use in our demonstration system.

First, we train a graph representation model to compute GraphBERT (Zhang et al., 2020) embed-

Table 1: CHSP predicton scores for single-token hyponyms generation for different source graph embeddings, replacement strategies and substitution layer (x100).

| Method	Context	Replaced	Layer	MRR@5	MRR@10	MRR@20	P@1	P@5	P@10
Pattern comparison (Hanna and Mareček, 2021)									
"[MASK] is a {parent}"	Yes	No	-	2.461	2.704	3.091	1.546	1.289	1.057
"My favourite {parent} is a [MASK]"	Yes	No	-	0.554	0.863	1.001	0.000	0.464	0.490
"A {parent} such as a [MASK]"	Yes	No	-	0.168	0.193	0.235	0.000	0.155	0.103
BERT (parent embedding on inference)	No	No	-	1.003	1.083	1.203	0.940	0.251	0.188
fastText (nearest neighbours)	No	No	-	2.400	3.500	4.000	0.130	1.839	2.100
CHSP (Graph-BERT)	Yes	Yes	1st	4.502	4.995	5.371	3.093	1.598	1.340
		Mix	1st	1.448	1.813	2.033	0.773	0.876	0.979
		Yes	6th	5.503	6.216	6.453	3.093	2.371	2.010
		Mix	6th	2.981	3.500	3.836	1.546	1.649	1.495
		Yes	12th	5.215	5.674	6.027	3.093	2.113	1.598
		Mix	12th	**7.229**	**8.037**	**8.624**	**3.608**	**3.247**	**2.474**
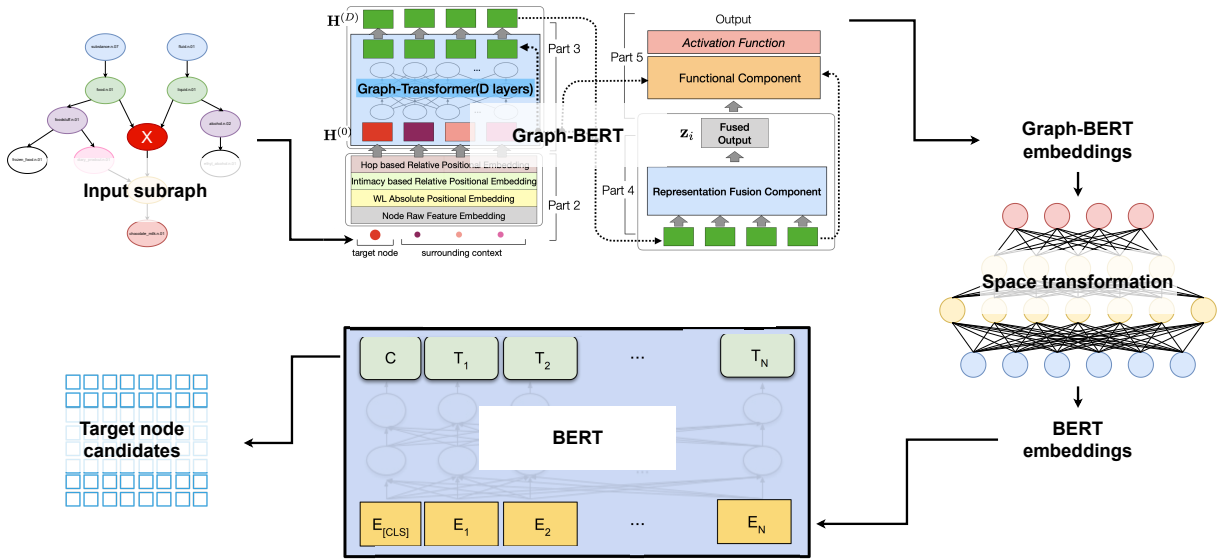
Figure 4: Cross-modal Contextualized Hidden State Projection Method (CHSP): graph-based BERT architecture that makes use of both node and text embeddings. Graph-BERT illustration source: (Zhang et al., 2020), BERT illustration source (Devlin et al., 2019). The computed GraphBERT embeddings from a taxonomy are projected from graph space to BERT space. Then BERT was used to predict candidates from the projected embeddings.

dings. Then we learn a feed-forward neural network as a projection layer to transform target graph embeddings to the BERT vector space. We use the SemCor dataset (Langone et al., 2004) that maps WordNet entities with the corresponding words in the context and learn the projection from Graph-BERT space to BERT. The next step is to apply the projected embeddings as input to the masked language modelling part of BERT model. We have evaluated three different context constructions suggested in (Hanna and Mareček, 2021): 1. "[MASK] is a/an {parent}"; 2. "My favourite {parent} is a [MASK]"; 3. "{parent} such as a [MASK]" . The scores for the amount of true hyponyms in a list of predicted candidates are presented in the first three lines of Table 1.

Then we incorporate the result graph embed-

ding into the language model prediction using mixed (or contextualised) prediction: embedding of "[MASK]" token is averaged with projected graph embedding. The replacement can happen at three different stages: after first layer of BERT encoder, after sixth (middle) or after twelfth (last). Thus, the prediction head generates new lemmas that are treated as candidate hyponyms for the target nodes.

3.3 Experiments and Results

In this research, we perform experiments on Word-Net 3.0 (Miller, 134 1995) nouns (82,115 synsets, 117,798 lemmas). For each "parental" hypernym all its hyponyms (leaves) were replaced by a single "masked" node. This place in the taxonomy was then considered for extension and the candidates

predicted for the masked node could be compared against true hyponyms. All in all, we masked 4,376 leaves out of 65,422 noun leaves to 1000 "[MASK]" tokens. We limit our experiments to leaves only, replacing all children with one mask in order to be able to compare with a wide range of possible answers, as one synset might have several hyponyms.

We utilize Precision@k (P@k), Recall (R@k), and Mean Reciprocal Rank (MRR) for evaluation. The results are presented in Table 1. We observe that the patterns from (Hanna and Mareček, 2021) show results are mostly far from the top ones. This happened because the context encapsulated in the patterns in general contains little information. We also see that our method outperforms the BERT (parent embedding on inference) baseline (which is a simple prediction of encoded parent synset) and a simple approach on fastText nearest neighbours candidates.

4 System Design

TaxFree is designed to help lexicographers in their work on updating taxonomies and inspecting them for the required modifications. In the current section we discuss each part of the tool and its usage in detail.

4.1 Software Architecture

The system consists of a web-based user interface through which users can explore the WordNet-3.0 taxonomy. The front-end is implemented with JavaScript library **vis.js**[3] used to display networks consisting of nodes and edges. It supports the hierarchical layout and allows us to integrate with the network. Back-end is written in Python using **Flask**[4] framework. It has an API with several "GET" and "POST" queries that maintain functioning of the system: (i) searching for synsets, (ii) getting image by node id, (iii) getting the current node graph context, (iv) generating new nodes.

4.2 Main Page

As a start page 3 , you see the highest level of the taxonomy, a tree with the root node "entity.n.01" which is highlighted with green color. Normally, the target node is displayed within its two-hop neighbourhood. To the right of the graph visualization there is a card with the description of the

current node: its image from ImagNet (if any), definition and the list of lemmas. Above the graph visualization box there are two buttons: **"Reset graph"**, **"Back to root"** and a search box with a **"Move to"** button. **"Reset graph"** means that all generated nodes will be deleted from memory and only the initial WordNet-3.0 graph will be displayed. **"Back to root"** will return the user to the display of the root of the taxonomy, leaving all generated nodes untouched. Search bar allows to easily navigate through the taxonomy and display subgraphs for the queried node. More details for each box are provided in the corresponding subsections.

4.3 Synset Search

The search bar accepts both synset names and lemmas and helps to disambiguate unclear queries to WordNet-3.0. The user can enter a word or a phrase separated with spaces or underscores. Moreover, noun synsets are also accepted, e.g. "cat.n.01" or "standard_poodle.n.01". If the synset name is not recognized there will be no error displayed, but the search bar will be empty again. For the entered lemma(s) there is a special pipeline that the query word goes through:

1. If there is only one synset corresponding to the query lemma, then this synsets will be displayed.

2. If there are more than one synset, therefore, the user is forwarded to the subgraph of the most common synset, displaying other disambiguation options under the synset description card (see Figure 6 as an example). Each disambiguated synset is presented with its synset name and definition.

After the query synset has been identified (either manually or automatically), the tool opens the required page with the query synset as the target node in context. Subgraph display and synset description card are described in the following subsections.

4.4 Subgraph Display

Central (query) synset is displayed with the closest "relatives" two hops away from the query (it might be less if there are no neighbours at the certain step away of the target node) in the central box of the page. It has green borders that highlight that the current image is the target one. However, if the image from the ImageNet is not presented, then the
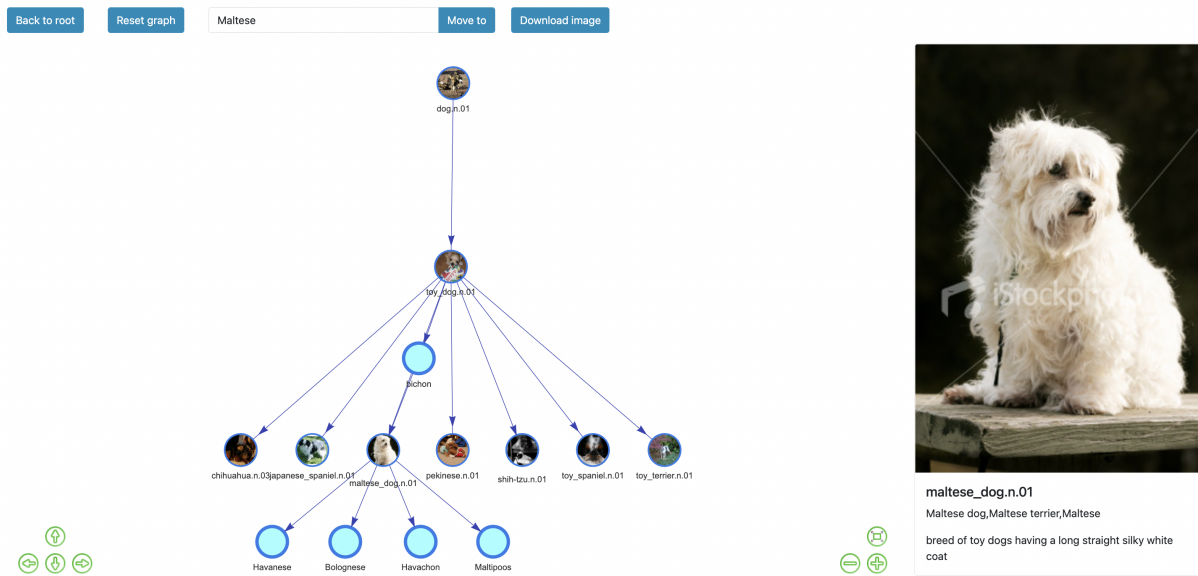

Figure 5: Generation of a new node for the leaf node "maltese_dog.n.01"

whole node is colored in green. Other nodes are not highlighted and colored in blue (in case there is no image to display).

Figure 5 shows the result page for the synset "maltese_dog.n.01" as an example to demonstrate synsets with images. Here we can see that all nodes have their representations from WordNet-3.0. The node "maltese_dog.n.01" is a leaf node, therefore, it is placed in the bottom of the graph and has only co-hyponyms at the same level and one hypernym "toy_dog.n.01" and one hypo-hypernym "dog.n.01". The arrows always have the same direction: from abstract words to more concrete. By clicking on a node twice it will open you a subgraph for this node, as it would consider it as a next query word. Therefore, you might be able to navigate through the graph even without queries. The graph might be downloaded when pressing the **"Download graph" button**.

In the bottom of the visualization box there are buttons to zoom in/out and centering. You can also move the graph using "left", "right", "up" and "down" buttons on the screen or simply use mouse.

4.5 Synset Description Card

To the right of the subgraph display, there is a card with the summary of the query node. It consists of a definition, image from the ImageNet (if any), synset name, list of lemmas. If any information about the node is missing, then the row is skipped.

Image for the node is selected randomly, normally the first one from the ImageNet dataset. According to the statistics, only 19,167 synsets have their images.

4.6 New Synsets Prediction

Figure 5 demonstrates the process of adding new nodes to the taxonomy using the algorithm described in Section 3.2. First, you can generate a new node starting from a leaf. By clicking twice on it, we can generate children for the "maltese_dog" synset, which does not possess hyponyms. Otherwise, they would be displayed, as "maltese_dog" is the central node. Therefore, by clicking twice on the target (query) node we can predict a candidate child for it. Figure 5 shows that there were generated new nods - the names of dog breeds. Another option for new synset generation is to predict a new node which is placed between two nodes (in this case it means that on of them is hypernym to the other). To generate this node, the user should click twice on the edge that connects them. This option has been added in case there are unaccounted words that should be placed in the middle of the graph. Figure 5 also depicts the "dog.n.01" and "toy_dog" nodes. By clicking twice on the edge between them, a new node is generated which is supposed to be more general then its hyponym "maltese_dog" and narrower than the word "toy_dog". However, we have not evaluated the performance of this specific

44

dog.n.01

dog,domestic dog,Canis familiaris

a member of the genus Canis (probably descended from the common wolf) that has been domesticated by man since prehistoric times; occurs in many breeds

Maybe you meant:

frump.n.01	a dull unattractive unpleasant girl or woman
dog.n.03	informal term for a man
cad.n.01	someone who is morally reprehensible
frank.n.02	a smooth-textured sausage of minced beef or pork usually smoked; often served on a bread roll
pawl.n.01	a hinged catch that fits into a notch of a ratchet to move a wheel forward or prevent it from moving backward
andiron.n.01	metal supports for logs in a fireplace

Figure 6: Disambiguation blog for the "dog" lemma

type of node insertion, so we leave the application of our method for this subtask for further research.

5   Conlusion

The growing popularity of implementing taxonomies in different research and industry tasks has created the need for a platform for visualization of tree-like taxonomic subgraphs (query node in context). **TaxFree** provides such a platform for the visualisation and analysis of hypo-hypernymy subgraphs. The tool allows users to explore wordnet synsets in context and predict new synsets for the leaf nodes. Our work aims to bring taxonomies

to a broader audience, by making WordNet interface user-friendly in comparison to the standard WordNet[5] visualization.

Limitations

Despite multiple advantages of the presented system it still has several limitations we list below:

1. Firstly, currently our system cannot predict new hyponyms for synsets that are not leaves. Yet, methodologically it's possible.

2. Secondly, we demonstrate only the first image of the synset, while there might be several images for one concept.

Ethics Statement

In general, we do not see any ethical issues or negative consequences within the current work. At the same time, as we apply the pre-trained language model we may inherit social bias learned from the Web corpora.

References

Matteo Abrate and Clara Bacciu. 2012. Visualizing word senses in WordNet atlas. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, pages 2648–2652, Istanbul, Turkey. European Language Resources Association (ELRA).

Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. 2009. Gephi: an open source software for exploring and manipulating networks. In *Proceedings of the international AAAI conference on web and social media*, volume 3, pages 361–362.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.

Yejin Cho, Juan Diego Rodriguez, Yifan Gao, and Katrin Erk. 2020. Leveraging WordNet paths for neural hypernym prediction. In *Proceedings of the 28th*

---

[5] http://wordnet.princeton.edu

International Conference on Computational Linguistics, pages 3007–3018, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Christopher Collins. 2006. Wordnet explorer: applying visualization principles to lexical semantics. *Computational Linguistics Group, Department of Computer Science, University of Toronto, Toronto, Ontario, Canada.*

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

John Ellson, Emden Gansner, Lefteris Koutsofios, Stephen C North, and Gordon Woodhull. 2001. Graphviz—open source graph drawing tools. In *International Symposium on Graph Drawing*, pages 483–484. Springer.

Anna Giabelli, Lorenzo Malandri, Fabio Mercorio, Mario Mezzanzanica, and Andrea Seveso. 2020. Neo: a tool for taxonomy enrichment with new emerging occupations. In *International Semantic Web Conference*, pages 568–584. Springer.

Joseph E Gonzalez, Reynold S Xin, Ankur Dave, Daniel Crankshaw, Michael J Franklin, and Ion Stoica. 2014. {GraphX}: Graph processing in a distributed dataflow framework. In *11th USENIX symposium on operating systems design and implementation (OSDI 14)*, pages 599–613.

Michael Hanna and David Mareček. 2021. Analyzing BERT's knowledge of hypernymy via prompting. In *Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, pages 275–282, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Jesús Herrera, Anselmo Peñas, and Felisa Verdejo. 2005. Textual entailment recognition based on dependency analysis and *WordNet*. In *Machine Learning Challenges, Evaluating Predictive Uncertainty, Visual Object Classification and Recognizing Textual Entailment, First PASCAL Machine Learning Challenges Workshop, MLCW 2005, Southampton, UK, April 11-13, 2005, Revised Selected Papers*, volume 3944 of *Lecture Notes in Computer Science*, pages 231–239. Springer.

David Jurgens and Mohammad Taher Pilehvar. 2016. SemEval-2016 task 14: Semantic taxonomy enrichment. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 1092–1102, San Diego, California. Association for Computational Linguistics.

Jaap Kamps and Maarten Marx. 2002. Visualizing wordnet structure. In *Proc. of the 1st International Conference on Global WordNet*, pages 182–186.

Helen Langone, Benjamin R. Haskell, and George A. Miller. 2004. Annotating WordNet. In *Proceedings of the Workshop Frontiers in Corpus Annotation at HLT-NAACL 2004*, pages 63–69, Boston, Massachusetts, USA. Association for Computational Linguistics.

George A. Miller. 1995. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41.

Andrea Moro and Roberto Navigli. 2015. SemEval-2015 task 13: Multilingual all-words sense disambiguation and entity linking. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 288–297, Denver, Colorado. Association for Computational Linguistics.

Irina Nikishina, Mikhail Tikhomirov, Varvara Logacheva, Yuriy Nazarov, Alexander Panchenko, and Natalia V. Loukachevitch. 2022a. Taxonomy enrichment with text and graph vector representations. *Semantic Web*, 13(3):441–475.

Irina Nikishina, Alsu Vakhitova, Elena Tutubalina, and Alexander Panchenko. 2022b. Cross-modal contextualized hidden state projection method for expanding of taxonomic graphs. In *Proceedings of TextGraphs-16: Graph-based Methods for Natural Language Processing*, pages 11–24, Gyeongju, Republic of Korea. Association for Computational Linguistics.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Stephen Roller, Douwe Kiela, and Maximilian Nickel. 2018. Hearst patterns revisited: Automatic hypernym detection from large text corpora. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 358–363, Melbourne, Australia. Association for Computational Linguistics.

Özge Sevgili, Artem Shelmanov, Mikhail Y. Arkhipov, Alexander Panchenko, and Chris Biemann. 2022. Neural entity linking: A survey of models based on deep learning. *Semantic Web*, 13(3):527–570.

Kunihiro Takeoka, Kosuke Akimoto, and Masafumi Oyamada. 2021. Low-resource taxonomy enrichment with pretrained language models. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2747–2758, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

46

Jiawei Zhang, Haopeng Zhang, Congying Xia, and Li Sun. 2020. Graph-bert: Only attention is needed for learning graph representations. *CoRR*, abs/2001.05140.