CEUR-WS.org/Vol-4196/paper_15.pdf

Biomedical Ontology Matching using Relational Graph Neural Networks and RDFs Meta-Path Rules

Shervin Mehryar[1,*], Julie Loesch[1], Emma Pinckers[1], Alan van den Akker[1], Yannick Damoiseaux[1], Athanasios Stavrinopoulos[1], Nikola Prianikov[1] and Remzi Celebi[1]

[1]*Institute of Data Science, Maastricht University, Paul-Henri Spaaklaan 1, 6229 GT, Maastricht, Netherlands*

Abstract

The growing complexity of relational data in knowledge graphs necessitates advanced models to capture intricate graph structures. In the domain of health and life sciences, the use of biomedical ontologies prevails in many applications from database management to retrieval and publication. Due to heterogeneity and lack of standardization to create local ontologies, the reusibility and interoperability for these resources become limited whereby often manual and time-consuming processes are put in place to match representations for cross-domain applications. In this paper we explore embedding-based methods as an alternative approach for entity matching among biomedical ontologies at different complexity and interoperability levels and propose a novel framework base on Relational Graph Convolutional Networks (R-GCN) in combination with symbolic meta-rule integration. We compare our results to the state-of-the-art baseline models using metrics such as Hits@k, F-scores and Mean Rank (MR) and demonstrate the effectiveness of the proposed model in improving ontology matching tasks across multiple complex biomedical datasets.

Keywords

Ontology Matching, Relational Graph Convolutional Networks, Neuro-symbolic AI, Knowledge Graph Embeddings

1. Introduction

Ontology Matching (OM) is an actively growing area of research in Semantic Web (SW) that focuses on efficiently aligning and producing a set of correspondences between related entities. The process of OM is a key enabler in ontology-based data integration and federated query tasks to provide standard semantics across multiple systems [1]. A wide range of applications in Bioinformatics and Biomedicine can benefit from interlinked ontologies towards evidence-based research, content browsing, and semantically rich query answering [2]. Aligning ontologies in these applications involves producing a series of connections between semantically associated elements from each respective source. The alignments lead to interoperable representations that can be directly merged or indirectly used by knowledge base reasoners to improve their expressive power.

Various organizations exist with mission to develop and maintain high quality and specialized ontological knowledge, including National Cancer Institute Thesaurus (NCIT), Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), Online Mendelian Inheritance in Man (OMIM), Orphanet Rare Disease Ontology (ORDO), and Foundational Model of Anatomy (FMA). These organizations follow independent conventions and rely on local design patterns to build ontologies. As new information is discovered, data elements and structures are also updated by each independent organization. These shortcomings related to developmental patterns and evolutions are rooted in the lack of coordination and cause challenges to the interoperability of heterogeneous sources. In Figure 1 an example concept is shown describing types and subtypes related to the same epithelial and neoplasmic tumor tissues by NCIT and SNOMED CT. In each ontology the presence of different conventional and structural patterns

---

SWAT4HCLS 2025: The 16th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences, February 24–27, 2025, Barcelona, Spain

[*] Corresponding author.

shervin.mehryar@maastrichtuniversity.nl (S. Mehryar); julie.loesch@maastrichtuniversity.nl (J. Loesch)
0000-0002-9062-6925 (S. Mehryar); 0009-0003-8319-5746 (Y. Damoiseaux); 0000-0001-7769-4272 (R. Celebi)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-04-22

are noticeable. For instance, the concept “id9436005” termed “Basophil adenoma” from SNOMED CT maps to concept “C2856” labeled “Pituitary Gland Basophil Adenoma” in NCIT. Structurally, these are instances of “Benign adenomatous neoplasm” and “Pituitary Gland Neoplasm” categories, each of which with a different number of subclasses and neighbouring nodes. This motivating example illustrates the manual effort and tedious process that is required to map such ontologies at large without efficient interlinking algorithms.

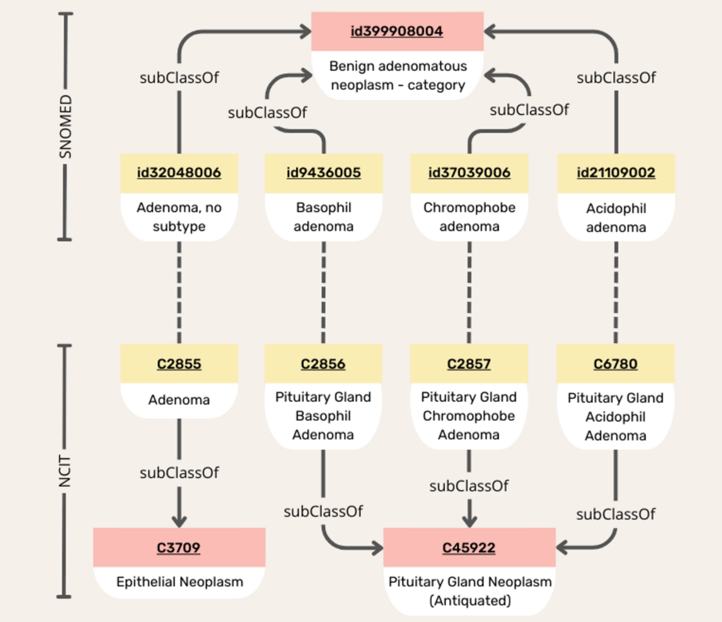

Figure 1: An example of Ontology Matching where the dashed lines show similar nodes in different ontologies.

In this paper we propose a novel methodology to predict the existence of links among respective entities among multiple ontologies based on Graph Neural Networks (GNN) as a powerful architecture for processing data over graphs. The algorithm learns representations on independent graphs and aligns them based on structural and semantic similarities. In particular, the Relational Graph Convolution Network (R-GCN) is utilized to support multiple relational types [3]. A new relation type is then introduced to propagate local information across graphs to perform ontology matching. In order to maintain structural integrity of learned representations, we further integrate Resource Description Framework Schema (RDFs) path rules. Our approach described in detail in section 3 can support varying structural complexities and is scalable to support large dataset sizes common in biomedical and bionformatic applications.

Through extensive experimentation described in section 4, we showcase the performance of the proposed model and compare against a baseline implementation. We assess the performance using various metrics including Hits@k, F-scores and Mean Rank (MR) on a number of biomedical datasets with varying complexity (summarized in Table 1). In order to measure the degree to which the algorithm scales with respect to the added complexities, we perform two additional sets of experiments including more class expressions and axioms. In doing so we demonstrate how well the algorithm adapts to evolving knowledge representations measured by the number of added facts and ontology rules. We further perform ablation testing to measure the generalization of the proposed model in terms of the percentage training data needed for optimal performance. We present and discuss the results, limitations, and future directions for the proposed framework in section 5.

2. Related Work

The survey in [4] provides a thorough overview of methods developed for ontology mapping. Overall, OM methods can be categorized into rule-based, embedding-based, and neuro-symbolic based. The latter combines the power of structural graphs with embedding representations, and is the primary focus in our work. Rule-based methods, including the work in [5], apply logical reasoning to extract rules that are in turn used to compute ontological concept similarities. The approach in [6] further incorporates external ontologies as background knowledge, together with lexical (label) and structural (relations) information. Lexical information can include names of all listed classes and synonym properties captured by “hasExactSynonym" and “hasRelatedSynonym", while structural information contain the “is a" and “part of" relationships. On snomed2fma and snomed2nci datasets (two tasks also evaluated in our work), they achieve f1-scores of 40.7 and 67.1 points, respectively. Rule-based approaches make strong assumptions about the underlying ontologies, require careful crafting of features, and in most cases do not transfer well as we will demonstrate (see section 5).

Embedding-based methods learn vector space representations for entities either at triplet level or at path level [7] in order to compute their similarity scores. To align concepts from these different spaces, it becomes essential to learn a transformation matrix M, with the aim of $Me_1 \approx e_2$ for each entity pair (e_1, e_2). The computation of matrix M can be done through supervised, semi-supervised [8, 9], and unsupervised as well as few-shot learning approaches [10, 11]. In particular, singular value decomposition is applied for computing M in order to translate and rotate embeddings from the source vector space to the target vector space in [12]. This approach termed Absolute Orientation (AO) is further described and used as the baseline model in our experiments (see section 4).

Graph neural networks that leverage and aggregate local neighborhood information to compute embeddings have gained significant attention as a neuro-symbolic approach in knowledge base and reasoning applications [13, 14]. Specific to the task of ontology matching, ancestry nodes are integrated into the training process of a transformer based architecture for equivalent matching as a step toward symbolic rule integration in [10]. The approach in [15] further combines RDFs meta-path rules with graph attention networks for link prediction in a single ontology/graph setting. While these methods address specific performance issues, broader objectives related to the scalability and robustness of neuro-symbolic approaches for matching and integration of heterogenuous data sources across health and life science applications remain open. In what follows we describe our approach for addressing the open issues.

3. Methodology

In this section the details on the proposed approach for ontology matching are given, which is based on a graph neural network architecture that supports multi-typed edges. We formulate the problem as link prediction over a multi-typed relational graph with a new type of relation, namely “sameAs”, inserted between nodes that match between a pair of ontologies. The “sameAs” relation exists by design exclusively between nodes from separate ontologies and is in addition to already existing edges within each respective ontology. The ontologies themselves include multiple relation (node) types respectively and may share common ones. Embeddings are learned for various node and edge types through the application of the R-GCN encoder algorithm. Given a pair of input embeddings, a binary classifier is trained in the decoder to determine if a “sameAs” edge exists between matching nodes.

The proposed R-GCN encoder [3] is similar to the graph convolution network operating on local graph neighborhoods but is designed to support large-scale relational data. The R-GCN further accounts for the various edge types that are present in the graph and learns structural representations with respective to different edge types. In effect, it applies the same updates as in the single-type relation graph convolution network approach, with a modification to the local neighbourhood sampling that depends on the edge-types. More formally, for a multi-type graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$, the neighborhood of type r for a node u can be denoted by $\mathcal{N}_r(u)$. The vector representation for the node $\vec{u}$ is learned by aggregating the incoming relation-dependent messages calculated through a summation over all nodes $v \in \mathcal{N}_r(u)$ that constitute the r-neighbourhood of u, i.e. the set of nodes $\{v \in \mathcal{V} | (v, r, u) \in \mathcal{G}\}$. The incoming messages in vector form are then aggregated together and passed through an element-wise activation function. This feed-forward operation is applied up to K times creating a K-layer R-GCN encoder.

The decoder on the other hand uses a DistMult function [16] providing a score to indicate the likelihood that an edge type r exists for the triple (h, r, t). The score for the edge in triple format is defined by a ternary dot-product as $score(h, r, t) = \sum_{i=1}^{d} h_i \cdot r_i \cdot t_i$, where d is the number of dimensions. This definition assumes a symmetric relation such that $score(h, r, t) = score(t, r, h)$, which holds true naturally in the context of ontology matching where we are interested in matching nodes. For the task at hand this aim is to predict whether an edge of type "sameAs" exists or not, irrespective of the directionality of r. To this end, a thresholding strategy is used as $\sigma(\text{score}(h, r, t)) \geq 0.5$, where $\sigma()$ is the Sigmoid function defined by $\sigma(x) = 1/(1 + e^{-x})$ and r is the unique node matching relation "sameAs". For a triple $(h, r = \textit{“sameAs”}, t)$ scoring above the set threshold, the head h and tail r nodes are matched.

To expand the receptive field and reasoning capabilities of the model we further apply true-path rules similar to the work in [15]. The authors present a scalable neural-symbolic method called *2-Hop GAT* for reasoning over large and complex ontologies. This method builds on the Graph Attention Network (GAT) by incorporating two-hop meta paths, which effectively capture transitivity. By extending GAT to include nodes two hops away, the proposed approach improves reasoning capabilities. Additionally, a variant called *Filtered 2-Hop GAT* is introduced, which applies a filtering mechanism to direct two-hop meta paths to capture two specific RDF rules: (1) subclass transitivity, where if A is a subclass of B and B is a subclass of C, then A is also a subclass of C; and (2) if A is of type B and B is a subclass of C, then A is also of type C.

To further stabilize the training of R-GCN given the above modifications, we further introduce two new negative sampling strategies with respect to the types of edges utilized during the training. For edges that are contained within a single ontology, since both end nodes come from the same source $\mathcal{G}$, a negative edge is created by corrupting either the head or tail node with another node from the same source. As for relations crossing between two different graphs, with one endpoint in the first graph $\mathcal{G}_1$ and the other end node in the second ontology $\mathcal{G}_2$, the head and tail corruptions are applied assuming the corrupted choices also are from each respective graph. Therefore, in the first strategy the negative edge $(h, \neg r, t)$ implies that $h, t \in \mathcal{G}$ and in the second strategy $h \in \mathcal{G}_1$ and $t \in \mathcal{G}_2$. The negative edges are generated at an equal rate to the size of positive edges, i.e. one negative edge for each positive edge. The code is accessible on our Github: https://github.com/epnck/RGCN-OM.

4. Experimental Setup

The Ontology Alignment Evaluation Initiative (OAEI)[1] organizes annual campaigns dedicated to assessing novel methods of ontology matching through extensive open-source datasets for testing purposes. Our models were tested on two dataset collections from the OAEI 2022 campaign. The first collection, named Bio-ML, is a machine learning friendly biomedical track designed for equivalence and subsumption matching tasks. This track presents a unified evaluation framework suitable for both ML-based and non-ML-based OM systems. The datasets of this track are based on Mondo and UMLS Metathesarus, combining seven ontologies in total. The second collection, Anatomy, consists of two ontologies with the aim to align the anatomical concepts between mice and humans. In total, we

---

[1] https://oaei.ontologymatching.org/2022/

defined six Ontology Matching (OM) tasks, each involving two ontologies. All of these OM tasks are summarized in the Table 1, including the list of ontology pairs used for the matching task, the number of nodes, the number of edges and the broad category for each task.

The first task is related to anatomy real world case to match the Adult Mouse Anatomy with 26, 145 entities form 2, 744 classes and 11 edge types, and the NCI Thesaurus with 34, 214 nodes from 3, 304 classes and 13 edge types, describing the human anatomy. From the Bio-ML datasets, the matching tasks are defined as OMIM with 2, 420, 585 entities and 148 edge types to ORDO with 2, 728, 720 entities and 163 edge types, NCIT with 2, 420, 585 entities and 140 edge types to DOID with 161, 682 entities and 13 edge types, SNOMED with 2, 856, 034 entities and 199 edge types to FMA with 3, 729, 227 entities and 207 edge types, to NCIT (neoplas) with 4, 990, 055 entities and 233 edge types, to NCIT (pharm) with 4, 990, 055 entities and 233 edge types. For alignment, we extracted matching pairs are integrated the pairs as a triple using the“sameAs" relation to denote equivalence. We further extract and run experiments on six relation types, namely “subClassOf", “hasExactSynonym", “exactMatch", “intersectionOf", “disjointWith" and “someValuesFrom", and maintain nodes that are connected as such.

Table 1
Ontology Matching tasks and their categories.

Matching Task	Ontology		Category	Nodes		Edges	
	Source	Target		Source	Target	Source	Target
human2mouse	Human	Mouse	Anatomy	26145	34214	11	13
omim2ordo	OMIM	ORDO	Disease (Bio-ML)	2420585	2728720	148	163
ncit2doid	NCIT	DOID	Disease (Bio-ML)	2105624	161682	140	13
snomed2fma body	SNOMED	FMA	Body (Bio-ML)	2856034	3729227	199	207
snomed2ncit neoplas	SNOMED	NCIT	Neoplastic Process (Bio-ML)	2856034	4990055	199	233
snomed2ncit pharm	SNOMED	NCIT	Pharmacologic Substances (Bio-ML)				

As a final pre-processing step, the entire available dataset is further divded into three subsets taking into account the relation types between entitties. The first subset titled DS1, includes entities and edges with "subClassOf" and "sameAs" relations. The "subClassOf" relations connect different entities within the same ontology, while the "sameAs" relations link entities across different ontologies. We include the "hasExactSynonym" and "exactMatch" relations in the second subset, named DS2, which introduce structural complexities for graph algorithms. In the third subset labeled DS3, the remaining relation types and entities including the "intersectionOf" and "someValuesFrom" relations capturing class expressions and axiomatic complexities are added, forming the most comprehensive subset used. By doing so, we are able to perform experiments and measure the degree to which the performance of R-GCN is maintained with increasing ontological complexity (see Seciton 5).

For our baseline method, we followed a recent approach from Portisch et al. [12] which performs *absolute orientation* on the embedding spaces. In this approach, node embeddings from the source and target ontologies are learned separately using RDF2Vec [17]. A rotate matrix is then derived that maps the target embedding space onto the source embedding space. Finally, the nodes from the source and target ontologies are matched by computing the Euclidean distance. Similar to the process of training the R-GCN-DistMul model, positive and negative examples are used to train the various parameters of the model. We used the default RDF2Vec settings, as they are generally considered to provide sufficient performance in most cases. Specifically, RDF2Vec was trained for 100 epochs with a window size of 6 and an embedding dimension of 200. The random walker was configured with settings of `max_depth=6` and `max_walks=150`.

In the encoder for the proposed R-GCN approach, 2 RGN operators are implemented, each followed by a ReLu and Dropout ($p = 0.2$) layer using the PyGeometric library[2]. Additionally, the hyperparameters were optimized using a validation split and set to `batch_size=1024`, `learning_rate=0.005`, `dropout=0.2`, `regularization=1e-2`, and `num_layers=3`. A more focused grid search for embedding dimensions to improve Hits@k metrics was performed and set as shown in Table 2. When running models with different hyperparameters on the same dataset, it was found that the embedding size affects performance. Some datasets benefit from larger embeddings, while others perform better with smaller ones. Larger embeddings do not always lead to better performance, as they increase complexity and the risk of overfitting on the training data, resulting in a model that generalizes poorly. In our experiments we create a separate train and test split for each task at a random 80-to-20 ratio with respect to "sameAs" edge type and train a new model in each case.

Table 2
Ontology Matching Tasks and their Categories.

Matching Task	**Embedding Dimension**	**Triples** (DS1)		**Triples** (DS2)		**Triples** (DS3)		**Triples**
		Source	Target	Source	Target	Source	Target	Match
human2mouse	200	3761	1807	9025	2151	10687	3788	2565
omim2ordo	400	4322	2551	29901	2551	-		3721
ncit2doid	100	5625	9373	-		5672	9373	4686
snomed2fma (body)	400	11903	64412	-		11903	153421	7256
snomed2ncit (neoplas)	400	1145	2619	-		1145	2730	3804
snomed2ncit (pharm)	100	899	16858	-		899	16882	5803

We will evaluate the produced alignments of concepts using the most common evaluation metrics in the domain of ontologies [11], Precision (P), Recall (R) and F-1 Score (F), which are defined as:

$$P = \frac{|M_{out} \cap M_{ref}|}{|M_{out}|}, \qquad R = \frac{|M_{out} \cap M_{ref}|}{|M_{ref}|}, \qquad F = \frac{2PR}{P + R} \tag{1}$$

where M_{ref} are the reference mappings and M_{out} are the mappings produced by the OM systems. In our case, M_{out} is constructed by taking the closest entities in the oriented embedding spaces. In addition, we will provide evaluations using $Hits@K$ and Mean Rank (MR).

$$Hits@K = \frac{|\{m \in M_{ref} | Rank(m) \leq K\}|}{|M_{ref}|}, \qquad MR = \frac{\sum_{m \in M_{ref}} Rank(m)}{|M_{ref}|} \tag{2}$$

where $Rank(m)$ is the ranking position of the matched concept by OM systems.

---

[2] https://pyg.org/

Dataset	Matching Task	Model	Metrics					
			P	R	F	Hits@1	Hits@10	Mean Rank
DS1	human-mouse	AO	0.8	0.77	0.76	0.27	0.67	14.9
		R-GCN	0.98	0.98	0.98	0.79	1	0.45
		R9-GCN	1	1	1	**0.88**	1	0.17
		R11-GCN	1	1	1	0.77	1	0.42
		R*-GCN	0.49	0.49	0.39	0.78	0.99	0.60
	ncit-doid	AO	0.76	0.74	0.73	0.39	0.75	13.2
		R-GCN	0.98	0.98	0.98	**0.92**	1	0.10
		R9-GCN	1	1	1	0.76	1	0.37
		R11-GCN	1	1	1	0.57	1	0.82
		R*-GCN	0.75	0.51	0.36	0.86	0.99	0.26
	omim-ordo	AO	0.6	0.59	0.59	0.0	0.02	55.8
		R-GCN	0.98	0.98	0.98	**0.91**	1	0.15
		R9-GCN	1	1	1	0.90	1	0.15
		R11-GCN	0.98	0.98	0.98	0.90	1	0.13
		R*-GCN	0.18	0.21	0.19	0.67	0.99	1.03
	snomed-fma.body	AO	0.52	0.52	0.52	0.13	0.32	37.05
		R-GCN	1	1	1	**0.99**	1	0.01
		R9-GCN	1	1	1	0.98	1	0.02
		R11-GCN	1	1	1	0.97	1	0.04
		R*-GCN	0.66	0.50	0.33	0.69	1.0	0.44
	snomed-ncit.neoplas	AO	0.4	0.4	0.4	0.17	0.33	35.2
		R-GCN	0.93	0.92	0.92	0.76	0.99	0.61
		R9-GCN	0.96	0.95	0.95	**0.87**	1	0.34
		R11-GCN	0.95	0.94	0.94	0.86	0.99	0.39
		R*-GCN	0.86	0.811	0.80	0.80	0.98	0.61
	snomed-ncit.pharm	AO	0.66	0.65	0.64	0.15	0.39	28.98
		R-GCN	0.95	0.95	0.95	0.58	0.98	1.15
		R9-GCN	0.99	0.99	0.99	0.48	1	1.11
		R11-GCN	0.96	0.96	0.96	0.46	1	1.23
		R*-GCN	0.79	0.76	0.758	**0.62**	0.99	0.93
DS2	human-mouse	R-GCN	0.45	0.50	0.34	0.85	1	0.25
		R*-GCN	0.88	0.85	0.84	**0.92**	1	0.12
	omim-ordo	R-GCN	0.76	0.74	0.74	0.83	1	0.35
		R*-GCN	0.83	0.78	0.77	**0.86**	1	0.18
DS3	human-mouse	R-GCN	0.77	0.61	0.54	0.36	0.99	1.75
		R*-GCN	0.97	0.97	0.97	**0.92**	1	0.09
	ncit-doid	R-GCN	0.50	0.50	0.43	0.12	0.84	5.34
		R*-GCN	0.90	0.89	0.88	**0.73**	1	0.65
	snomed-fma.body	R-GCN	0.79	0.64	0.59	0.42	0.66	10.19
		R*-GCN	0.86	0.82	0.82	**0.62**	0.98	0.95
	snomed-ncit.neoplas	R-GCN	0.26	0.46	0.31	0.08	0.71	6.85
		R*-GCN	0.84	0.77	0.75	**0.69**	0.98	0.97
	snomed-ncit.pharm	R-GCN	0.59	0.55	0.50	0.12	0.96	3.73
		R*-GCN	0.86	0.81	0.80	**0.56**	1	1.23

Table 3

Model performance on matching tasks from different datasets. Baseline (AO) shown for DS1 with a single edge type. R9-GCN, R11-GCN, R*-GCN refer to variations to R-GCN with fist, second, and combined with guided-path rules. Evaluated by Precision (P), Recall (R), f-1 score (F), Hits@1, Hits@10 and Mean Rank. The best and second best Hits@1 scores are highlighted and underlined in each task.

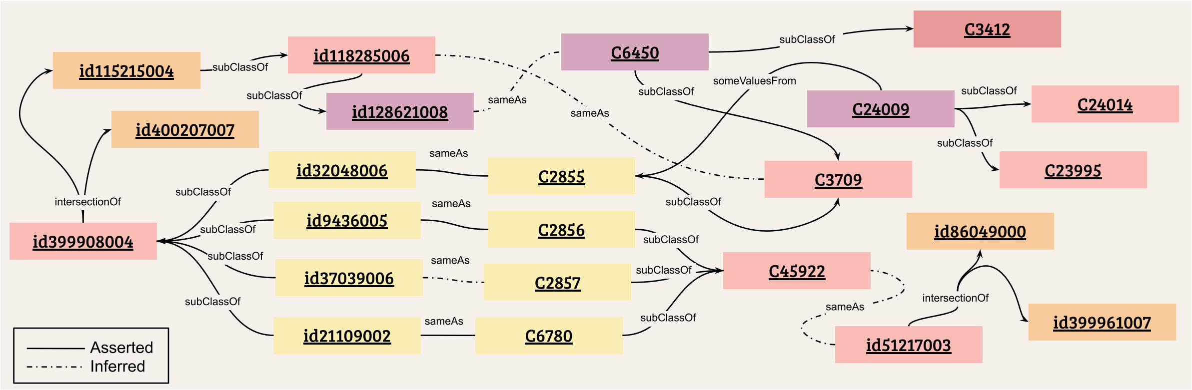

Figure 2: SNOMED to NCIT ontology matching. R*-GCN is applied to predict “sameAs" edges (dashed lines) based on multiple asserted edge types (solid lines). For instance, the SNOMED entity *Epithelial neoplasm* by the code *id118285006* is matched with similarly named concept having identifier *C3709* in NCIT via meta-path “intersectionOf” and “subClassOf”. The colors signify nodes at the same ontological hierarchy.

5. Results and Discussions

The performance results of the proposed model are shown in Table 3 in terms of three subsets (DS1, DS2, DS3) that differ in structure and complexity. More specificaly, in addition to the number of triples included as shown in Table 1, DS2 and DS3 include additional edge types described above. Each subset consists of six, two and five matching tasks respectively corresponding to different inter-ontology mapping tasks. We investigate the performance of the proposed model against the baseline model AO using R-GCN, as well as its three variations which include the first meta-path rule as R9-GCN and the second meta-path rule as R11-GCN. The last variation as R*-GCN is guided by both meta path rules and further uses the 2-hop filtering mechanism described in section 3. The asserted and inferred edges using the proposed method as applied to the subgraph shown in Figure 1, are shown in Figure 2.

In the first set of evaluations DS1, each pair of tasks contains only class subsumption relations constituting a single relation type within the source and target ontologies, and a second relation type indicating the equivalent entities between the source and target ontologies - a sum total of two relation types. Over the baseline model AO, a considerable improvement in performance is achieved across all six matching tasks. Specifically, the precision, recall, f1-score, hits@1, hits@10, and Mearn Rank are improved by 0.53, 0.52, 0.48, 0.91, 0.98, and 37.04 points in the worst case and by 0.18, 0.21, 0.22, 0.53, 0.25, 12.16 points in the best case between AO and R-GCN. This demonstrates the capability of R-GCN as a neuro-symbolic approach that is capable of both learning intra-ontology representations fused with inter-ontology information via its graphical and message passing nature which allow the propogation of such information locally as well as across global graph structures. Including meta-path rules as in the case of R9-GCN and R11-GCN further improves the performance and robustness across matching tasks. The R*-GCN which combines the two rules, in comparison underperforms over mainly subsumption-based graphs of DS1.

In the second DS2 and third DS3 evaluation subsets, with two and four additional edge types within each ontology respectively, the aim is to investigate the scalability, expressiveness and robustness of the proposed model against increasing structural and axiomatic complexities. We spare with the baseline model which does not directly support multi-type edges. With R-GCN, the performance is maintained at 0.85 and 1.00 in Hits@1 and Hits@10 metrics for DS2. The R*-GCN further improves the Hits@1 result by 0.07 points. It also consistently outperforms R-GCN across DS3 data subsets. The drop in performance of R-GCN on DS3 is overall attributed to increasing dataset size, common in large scale knowledge graphs. The R*-GCN compensates for increased dataset size through the integration of meta-path rules [18].

Lastly, we evaluate the performance of the proposed model in scenarios where limited training data are available, summarized in Table 4. Particularly, we consider scenarios in which access to labeled edges for matching are limited and available at ratios 0% (zero-shot), 25%, 50%, 75% and up to 100% in the training set reflected by MR and F metrics. The MR in particular is chosen so as to best demonstrate the changes in performance in single entity matching scenarios. It can be observed that using more labeled edges can lead to a better generalization performance, with severe degradation results found at 0% as expected. Overall however, with access to 75% and in some cases even as low as 25% "sameAs" edges , optimal performance in terms of both MR and F metrics can be recovered.

Test Case	**0%***		**25%***		**50%***		**75%***		**100%***	
	MR	**F**	**MR**	**F**	**MR**	**F**	**MR**	**F**	**MR**	**F**
human-mouse	26.8	0.35	3.64	0.65	2.25	0.67	0.69	0.80	**0.45**	**1**
ncit-doid	23.1	0.31	5.35	0.33	2.36	0.75	1.43	**0.95**	**0.1**	0.92
omim-ordo	16.7	0.33	4.22	0.93	3.21	0.85	1.84	**0.98**	**0.15**	0.98
snomed-fma.body	39.6	0.32	1.11	0.33	0.35	0.89	0.13	0.88	**0.01**	**1**
snomed-ncit.neoplas	31.32	0.32	7.02	0.47	3.68	0.71	1.39	0.72	**0.52**	**0.92**
snomed-ncit.pharm	39.85	0.33	4.42	0.33	2.90	0.43	2.55	0.83	**1.15**	**0.95**

Table 4
Mean Rank (MR) and f1-score (F) metrics for each test case in DS1 when including a different percentage of the sameAs edges during training (* percentage of sameAs train edges in the training set). In each case, the best and second best scores are highlighted and underlined.

6. Conclusions

In this paper, we explored two embedding-based approaches for the Ontology Matching task. Our primary contributions include: a novel variation to the R-GCN-DistMult model in the context of OM. Through an expansive experimental setup for the evaluation of R-GCN-DistMult across different datasets and parameters, we establish the supriority of the proposed model for predicting equivalent entities on six different matching tasks for Anatomy and BioML subsets. By incorporating meta-path guided rules into the proposed R-GCN framework, our experiments show the robustness of the model to increasing structural and axiomatic complexities common across evolving ontologies. We further show that with access to between 25% to 50% of total labeled edges, optimal results can be achieved in most cases further attesting to the generalization capabilities of the proposed model.

Acknowledgments

The authors thank the organizers for their timely and helpful response to inquiries, and the reviewers for their valuable comments. This work is supported by the Horizon Europe Framework Program (Grant agreement No: 101057062), project AIDAVA (AI powered Data Curation & Publishing Virtual Assistant).

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati, M. Zakharyaschev, Ontology-based data access: A survey, International Joint Conferences on Artificial Intelligence, 2018.

[2] J. Hao, C. Lei, V. Efthymiou, A. Quamar, F. Özcan, Y. Sun, W. Wang, Medto: Medical data to ontology matching using hybrid graph neural networks, in: Proceedings of the 27th ACM

SIGKDD Conference on Knowledge Discovery & Data Mining, KDD '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 2946–2954. URL: https://doi.org/10.1145/3447548.3467138. doi:`10.1145/3447548.3467138`.

[3] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. Van Den Berg, I. Titov, M. Welling, Modeling relational data with graph convolutional networks, in: The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15, Springer, 2018, pp. 593–607.

[4] Z. Sun, Q. Zhang, W. Hu, C. Wang, M. Chen, F. Akrami, C. Li, A benchmarking study of embedding-based entity alignment for knowledge graphs, arXiv preprint arXiv:2003.07743 (2020).

[5] E. Jiménez-Ruiz, B. Cuenca Grau, Logmap: Logic-based and scalable ontology matching, in: The Semantic Web–ISWC 2011: 10th International Semantic Web Conference, Bonn, Germany, October 23-27, 2011, Proceedings, Part I 10, Springer, 2011, pp. 273–288.

[6] D. Faria, C. Pesquita, E. Santos, M. Palmonari, I. F. Cruz, F. M. Couto, The agreementmakerlight ontology matching system, in: On the Move to Meaningful Internet Systems: OTM 2013 Conferences: Confederated International Conferences: CoopIS, DOA-Trusted Cloud, and ODBASE 2013, Graz, Austria, September 9-13, 2013. Proceedings, Springer, 2013, pp. 527–541.

[7] H. Zhu, R. Xie, Z. Liu, M. Sun, Iterative entity alignment via joint knowledge embeddings., in: IJCAI, volume 17, 2017, pp. 4258–4264.

[8] Y. He, J. Chen, D. Antonyrajah, I. Horrocks, Bertmap: a bert-based ontology alignment system, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, 2022, pp. 5684–5691.

[9] S. Neutel, M. H. de Boer, Towards automatic ontology alignment using bert., in: AAAI Spring Symposium: Combining Machine Learning with Knowledge Engineering, 2021.

[10] Z. Wang, Contextualized structural self-supervised learning for ontology matching, arXiv preprint arXiv:2310.03840 (2023).

[11] M. Amir, M. Baruah, M. Eslamialishah, S. Ehsani, A. Bahramali, S. Naddaf-Sh, S. Zarandioon, Truveta mapper: A zero-shot ontology alignment framework, arXiv preprint arXiv:2301.09767 (2023).

[12] J. Portisch, G. Costa, K. Stefani, K. Kreplin, M. Hladik, H. Paulheim, Ontology matching through absolute orientation of embedding spaces, in: European Semantic Web Conference, Springer, 2022, pp. 153–157.

[13] J. Hao, C. Lei, V. Efthymiou, A. Quamar, F. Özcan, Y. Sun, W. Wang, Medto: Medical data to ontology matching using hybrid graph neural networks, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 2946–2954.

[14] S. Mehryar, R. Celebi, Semantic annotation of tabular data for machine-to-machine interoperability via neuro-symbolic anchoring., in: SemTab@ ISWC, 2023, pp. 61–71.

[15] J. Loesch, M. Dumontier, R. Celebi, Enhanced gat: Expanding receptive field with meta path-guided rdf rules for two-hop connectivity., in: QALD/SemREC@ ISWC, 2023.

[16] B. Yang, W. Yih, X. He, J. Gao, L. Deng, Embedding entities and relations for learning and inference in knowledge bases, in: Y. Bengio, Y. LeCun (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, 2015. URL: http://arxiv.org/abs/1412.6575.

[17] J. Portisch, M. Hladik, H. Paulheim, Rdf2vec light–a lightweight approach for knowledge graph embeddings, arXiv preprint arXiv:2009.07659 (2020).

[18] S. Mehryar, R. Celebi, Improving transitive embeddings in neural reasoning tasks via knowledge-based policy networks, in: Joint 2nd Semantic Reasoning Evaluation Challenge and 3rd SeMantic Answer Type, Relation and Entity Prediction Tasks Challenge, 2022, pp. 16–27.