Supplemental Material

Anonymous

1 Related Works

Our work is an intersection branch of ontology representation learning and neuro-symbolic computing. And we give a feasible way out for a symbol grounding problem — semantic image interpretation.

Our approach can be applied to various symbol grounding problems, including ontology population [Harnad, 1990, Cangelosi, 2011, Coradeschi et al., 2013], as well as tasks such as semantic image interpretation [Donadello et al., 2017, Arvor et al., 2019, Oltramari et al., 2020], named entity recognition [Seyler et al., 2018, Hu et al., 2016, Torisawa et al., 2007], and schema matching [Jiang et al., 2021, Liang et al., 2017, Unal et al., 2006]

1.1 OWL Ontology Embedding

The main motivation for embedding OWL ontology into vector space is to transfer the knowledge to vectors so it can be used directly in knowledge-requiring downstream tasks. Two kinds of methods exist with different input requirements. The first kind focuses on coupling the meta-data of an OWL ontology into an efficient graph, and then uses the generated corpus based on the graph as the input to the existing representation learning methods [Smaili et al., 2018, 2019, Chen et al., 2021]. The second kind focuses on modeling the logical semantics of an OWL ontology. EL2Vec [Kulmanov et al., 2019] approximates geometric models for $\mathcal{EL}$ ontologies and has achieved an interpretable embedding for Gene Ontology. E2R [Garg et al., 2019] can model the logical operators of intersection, union, negation, and universal quantifier, but fails to capture the distributive law. In all, there often comes a loss of the semantics of an OWL ontology in the transformation of most embedding methods [Smaili et al., 2018, 2019, Chen et al., 2021]. Though the geometric construction method [Kulmanov et al., 2019, Mondal et al., 2021, Sun et al., 2020, Jackermeier et al., 2023] can preserve the logical semantics well in $\mathcal{EL}$, the embedding may bring unexpected knowledge (unknown becomes true) because of the closed world assumption (CWA), and $\mathcal{EL}$ is not expressive as $\mathcal{ALC}$ studies in this work.

1.2 Neuro-symbolic Computing

Neural-symbolic computing aims at computing with both learning and reasoning abilities, to step towards the combination of symbolic and sub-symbolic systems. Current learning ability relies largely on differentiable programming to draw conclusions from observations and apply them, while current reasoning ability relies largely on logical programming to give conclusions inferred from premises and rules through deductive reasoning, give rules according to observations comprising premises and conclusions through inductive reasoning, and give premises that can interpret conclusions according to rules through abductive reasoning. So it comes with challenges in the integration and representation of these two kinds of programming paradigms. From the perspective of integration, research works differ in logical techniques that are mainly consumed. Neural-symbolic inductive logical programming [Wang et al., 2013, Bühmann et al., 2016, Yang et al., 2017, Evans and Grefenstette, 2018, Sen et al., 2022] and statistical relational learning (e.g. Markov logic network [Richardson and Domingos, 2006], probabilistic soft logic [Bach et al., 2017]) works seek to learn probabilistic logical rules from observations. This requires learning model parameters in a continuous space and the structure in a discrete space. SATNet [Wang et al., 2019] learns rules from labeled data by transforming the learning problem as SAT problem[1]. To combine the ability of deductive reasoning, the first line of research learn to reason by modeling the inference procedure using neural networks or replacing logical computations with differentiable functions [Towell and Shavlik, 1994, Hölldobler et al., 1999, Rocktäschel and Riedel, 2016, 2017, Diligenti et al., 2017, Ebrahimi et al., 2021]. But this neglects factual knowledge which bridges the physical world and the conceptual world, so the second line of research aims to find an interpretation (grounding) that satisfies theories which can be a mapping between these two worlds by encoding the satisfiability of theories in the loss function [Badreddine et al., 2022, Serafini and Garcez, 2016, Riegel et al., 2020, Topan et al., 2021, van Krieken et al., 2022]. The notable work Logical Tensor Network (LTN) [Badreddine et al., 2022] uses neural networks to represent the fuzzy function and predicates of theories, which is learned from labeled data. To solve the symbol grounding problem, LTN learns the interpretation with trained parameters that can maximize the satisfiability of theories. But these works cannot find explanations of observations according to theories, so abductive learning-based neural-symbolic works are proposed to use the explanations getting through abductive reasoning to promote the interpretability of the computing [Zhou, 2019, Huang et al., 2020, Tsamoura et al., 2021, Cai et al., 2021]. From the view of representation, some works are based on classical logic — propositional logic [Towell and Shavlik, 1994, Zhou, 2019, Tsamoura et al., 2021, Cai et al., 2021], description logic [Bühmann et al., 2016, Eberhart et al., 2019, Ebrahimi et al., 2021], or first-order logic [Hölldobler et al., 1999, Wang et al., 2013, Rocktäschel and Riedel, 2016, Serafini and Garcez, 2016, Rocktäschel and Riedel, 2017, Yang et al., 2017, Evans and Grefenstette, 2018, Sen et al., 2022], others are based on non-classical logic, such as fuzzy logic [Diligenti et al., 2017, Riegel et al., 2020, van Krieken et al., 2022], or probabilistic logic [Wang et al., 2013, Manhaeve et al., 2018]. Our work tries to combine the deductive reasoning ability in a novel way, which follows the work of LTN. As stated in the introduction, when using declarative knowledge, data bias should be ignored, so different from works that only maximize the satisfiability of the knowledge base, we give a compulsory way to inject knowledge.

[1] this is the satisfiability (SAT) problem which aims to determine whether there exists an interpretation that satisfies a given formula

1.3 Semantic Image Interpretation

A symbol grounding application of our work is semantic image interpretation (SII) [Hudelot et al., 2005, Neumann and Möller, 2008, Krishna et al., 2017, Donadello et al., 2017], which aims to generate a structured and human-readable description of the content of images. Current successful SII researches [Donadello et al., 2017, Arvor et al., 2019] rely on background knowledge of the images. LTN [Donadello et al., 2017] models predicates and functions as neural networks and learns the representation through maximizing the satisfiability in a supervised way. The main struggle of these neural-symbolic works in leveraging logical knowledge to adapt to the symbol grounding problem is that the revision signal cannot be properly conveyed. Some early works [Hudelot et al., 2008, Dasiopoulou et al., 2009] revise image semantics solely with fuzzy description logic, the idea of which can be summarized as selecting plausible image scene descriptions (assertions) ignoring disjoint cases, and then handle the inconsistency by removing assertions using methods e.g. reversed tableaux expansion procedure. Though these fuzzy description logic based methods can remain as reliable parts of perceptual grounding as possible, these works achieve most of the expected revisions and are specific to the ontology designed for image scene classification, which lacks generalization ability.

SII was formerly known as multimedia information retrieval. Fuzzy logic is recognized for its ability to link linguistic categories with numerical data and to express users' preferences in a qualitative and gradual manner and is frequently used in multimedia information retrieval. Though fuzzy logic was solely used to model uncertain pieces of statistical information without giving a revision to the contradictory part (unreliable part) of statistical information Meghini et al. [2001], Krishnapuram et al. [2004], Straccia [2010]. Therefore, when the retrievals should not violate the context knowledge base, the retrievals generated by these models are not safe.

2 Properties of Zadeh-$\mathcal{ALC}$

Table 1 is a supplemental material for section 4.2.

3 Proof for Soundness

Proposition 1 *(Soundness of the semantics) Let $\sharp\mathcal{O}$ be an $\mathcal{ALC}$ ontology, and φ be a fuzzy assertion. $\mathcal{O} \approx\!\!\!\!\mid\;\; \varphi$ iff. $\sharp\mathcal{O} \models \sharp\varphi$ (i.e. fuzzy entailment is consistent with entailment in $\mathcal{ALC}$).*

Proof. 1.$\Rightarrow$ Consider any fuzzy interpretation $\mathcal{I}$ that is a model of $\mathcal{O}$. $\mathcal{I}$ can also equivalently transform into a crisp interpretation, which is shown in Table 2. $\sharp((\exists r.D)^{\mathcal{I}}(a))$ corresponds to the maximum value of a set of values in a column of $\sharp((C \sqcap D)^{\mathcal{I}}(a))$. $\sharp((\forall r.D)^{\mathcal{I}}(a))$ corresponds to the minimum value of a set of values in a column of $\sharp((C \sqcup D)^{\mathcal{I}}(a))$. By enumerating a 3-dimensional $r^{\mathcal{I}}(\cdot, a)$, we can also summarize that the semantics of existential quantifier or forall quantifier is not changed after the crispy transformation. Based on this result, inductively defined complex concept assertions'

Table 1: Properties of Zadeh-$\mathcal{ALC}$ with equality assertion used in our model and product real logic [van Krieken et al., 2022] used in LTN [Badreddine et al., 2022]. Cell with • is the meaning of 'has'.

Property	Zadeh-$\mathcal{ALC}$	Product Real Logic
$C \sqcap \neg C \cong \bot$		
$C \sqcup \neg C \cong \top$		
$C \sqcap C \cong C$	•	
$C \sqcup C \cong C$	•	
$\neg\neg C \cong C$	•	•
$\neg(C \sqcap D) \cong \neg C \sqcup \neg D$	•	
$\neg(C \sqcup D) \cong \neg C \sqcap \neg D$	•	
$C \sqcap (D \sqcup E) \cong (C \sqcap D) \sqcup (C \sqcap E)$	•	
$C \sqcup (D \sqcap E) \cong (C \sqcup D) \sqcap (C \sqcup E)$	•	
$\forall r.C \cong \neg \exists r.\neg C$	•	

semantics also cannot be changed after crisp transformation. And the semantics of $\mathcal{O}$ will not be changes by the crisp transformation. And similarly for atomic role assertions. Therefore, the crisp transformation of $\mathcal{I}$ is also a model of $\sharp\mathcal{O}$. And for every model of $\sharp\mathcal{O}$, $\mathcal{I}$ satisfies $\sharp\varphi$, then its crisp transformation $\sharp\mathcal{I}$ is also a model of $\sharp\mathcal{O}$ and $\sharp\mathcal{I} \models \sharp\varphi$ holds. Therefore, $\sharp\mathcal{O} \models \sharp\varphi$. 2. $\Leftarrow$ If $\sharp\mathcal{O} \models \sharp\varphi$, consider the crisp interpretation

	$\sharp(\neg C^{\mathcal{I}}(a))$	$\sharp((C \sqcap D)^{\mathcal{I}}(a))$			$\sharp((C \sqcup D)^{\mathcal{I}}(a))$		
$D^{\mathcal{I}}(a)$ / $C^{\mathcal{I}}(a)$	/	$[0, 1-\alpha)$	$[1-\alpha, \alpha]$	$(1-\alpha, 1]$	$[0, 1-\alpha)$	$[1-\alpha, \alpha]$	$(1-\alpha, 1]$
$[0, 1-\alpha)$	T	F	F	F	F	U	T
$[1-\alpha, \alpha]$	U	F	U	U	U	U	T
$(\alpha, 1]$	F	F	U	T	T	T	T

Table 2: The truth value table after crisp transformation $(\cdot)$ on three types of concept assertions. F is short for False. T is short for True. And U is short for unknown.

$\mathcal{I}$ discussed above, it is similar to proof that the fuzzy extension of each model $\mathcal{I}$ of $\sharp\mathcal{O}$, is also the model of $\mathcal{O}$, and satisfies φ, so we have $\mathcal{O} \approx\!\!\!\!\models \varphi$. To sum up, this proposition is proven to be true.

Theorem 1. *For any $\mathcal{ALC}$ ontology $\mathcal{O}$, one can construct in polynomial time a normalized $\mathcal{ALC}$-ontology $\mathcal{O}$' of polynomial size in $|\mathcal{O}|$ using the normalization described above such that (i) for every model $\mathcal{I}$ of $\mathcal{O}$, there exists a model $\mathcal{J}$ of $\mathcal{O}'$ such that $\mathcal{I}$ is semantically equivalent to $\mathcal{J}$ in $sig(O)$, denoted as $\mathcal{I} \sim_{sig(O)} \mathcal{J}$, and (ii) for every model $\mathcal{J}$ of $\mathcal{O}'$ there exists a model $\mathcal{I}$ of $\mathcal{O}$ such that $\mathcal{I} \sim_{sig(O)} \mathcal{J}$.*

Proposition 2 ***(Soundness of learning to ground in DF-$\mathcal{ALC}$)*** *When the hierarchical loss converges to 0, the learned interpretation $\mathcal{I}''$ is the model of the given $\mathcal{ALC}$ ontology $\mathcal{O}$. For any model $\mathcal{J}$ of $\mathcal{O}$, $\mathcal{I}'' \sim_{sig(\mathcal{O})} \mathcal{J}$.*

Proof. when loss converges to 0, the learned $\mathcal{I}''$ satisfies any $C \sqsubseteq D$ in the normalized ontology $\mathcal{O}'$, so $\mathcal{I}''$ is the model of $\mathcal{O}'$. And according to Theorem 1, any model of $\mathcal{O}'$ is semantically equivalent to the model of $\mathcal{O}$. So this proposition is proved to be true.

4 Rule-based Learning: Example Analyse

Example 1. Given a perceptual grounding $\mathcal{I}$ in the domain $\{s_1, s_2\}$, $\mathtt{A}^{\mathcal{I}}(s_1) = 0, \mathtt{A}^{\mathcal{I}}(s_2) = 0, \mathtt{B}^{\mathcal{I}}(s_1) = 0.9, \mathtt{B}^{\mathcal{I}}(s_2) = 0, \mathtt{r}^{\mathcal{I}}(s_1, s_2) = 0.9, \mathtt{r}^{\mathcal{I}}(s_1, s_1) = \mathtt{r}^{\mathcal{I}}(s_2, s_1) = \mathtt{r}^{\mathcal{I}}(s_2, s_2) = 0$, notated as vectors $\mathtt{A}^{\mathcal{I}} = [0, 0]$, $\mathtt{B}^{\mathcal{I}} = [0.9, 0]$.

According to the semantics of Zadeh-$\mathcal{ALC}$, in $\mathcal{O}_1$, $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} = [0, 0]$, which satisfies $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} \leq \mathtt{B}^{\mathcal{I}}$, so hierarchical loss is 0, and no revision is executed. But this is not what we want. As we know that s_1 is likely to be $\mathtt{B}$, and $\mathtt{r}(s1, s2)$ is likely to be true, so s_2 is likely to be a membership of $\mathtt{A}$. In $\mathcal{O}_2$, $(\forall \mathtt{r}.\mathtt{A})^{\mathcal{I}} = [0.1, 1]$, which does not satisfy $(\forall \mathtt{r}.\mathtt{A})^{\mathcal{I}} \leq \mathtt{B}^{\mathcal{I}}$, and hierarchical loss is 1.1. Through gradient decent, until loss becomes 0, $\mathtt{A}^{\mathcal{I}} = [0.24, 0]$,$\mathtt{B}^{\mathcal{I}} = [0.4, 1]$, $\mathtt{r}^{\mathcal{I}}(s_2, s_1)$ will be 0.7 and $\mathtt{r}^{\mathcal{I}}(s_1, s_2)$ will be 1. In $\mathcal{O}_3$, with hierarchical loss, $\mathcal{A}^{\mathcal{I}}$ will be $[0.38, 0]$, $\mathcal{B}^{\mathcal{I}}$ will be $[0.35, 0]$and $\mathtt{r}^{\mathcal{I}}(s_1, s_1)$ will be $[0, 36]$. In $\mathcal{O}_4$, with hierarchical loss, $\mathcal{B}^{\mathcal{I}} = [0, 0]$ and $\mathtt{r}^{\mathcal{I}}(s_1, s_2) = 0$.

Example 2. Given a perceptual grounding $\mathcal{I}$ in the domain $\{s_1, s_2\}$, $\mathtt{A}^{\mathcal{I}} = [0, 0.9]$, $\mathtt{B}^{\mathcal{I}} = [0, 0]$, $\mathtt{r}^{\mathcal{I}}$ is the same as in Example. 1.

According to the semantics of Zadeh-$\mathcal{ALC}$, in $\mathcal{O}_1$, $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} = [0.9, 0]$, which does not satisfy $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} \leq \mathtt{B}^{\mathcal{I}}$, so hierarchical loss is 0.9. Through gradient decent, until loss becomes 0, $\mathtt{A}^{\mathcal{I}}$ is decreased as $[0, 0]$, and $\mathtt{B}^{\mathcal{I}}$ is increased as $[0.9, 0]$. $\mathtt{A}^{\mathcal{I}}$ is not expected to be changed and $\mathtt{B}^{\mathcal{I}}$ is expected to be increased as $[0.9, 0]$. In $\mathcal{O}_2$, $(\forall \mathtt{r}.\mathtt{A})^{\mathcal{I}} = [0.9, 1]$, $\mathtt{A}^{\mathcal{I}}$ will be revised as $[0, 0]$, $\mathtt{B}^{\mathcal{I}}$ will be revised as $[0.5, 1]$, and $\mathtt{r}^{\mathcal{I}}(s_1, s_2) = 1$. In $\mathcal{O}_3$ and $\mathcal{O}_4$, there is no revision.

Example 3. Given a perceptual grounding $\mathcal{I}$ in the domain $\{s_1, s_2\}$, $\mathtt{A}^{\mathcal{I}} = [0, 0.9]$ and $\mathtt{B}^{\mathcal{I}} = [0.9, 0]$, $\mathtt{r}^{\mathcal{I}}(s_1, s_2) = \mathtt{r}^{\mathcal{I}}(s_1, s_1) = \mathtt{r}^{\mathcal{I}}(s_2, s_1) = \mathtt{r}^{\mathcal{I}}(s_2, s_2) = 0$.

According to the semantics of Zadeh-$\mathcal{ALC}$, $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} = (\forall \mathtt{r}.\mathtt{A})^{\mathcal{I}} = [0, 0]$, which satisfies $(\exists \mathtt{r}.\mathtt{A})^{\mathcal{I}} \leq \mathtt{B}^{\mathcal{I}}$, so no revision is executed. In $\mathcal{O}_3$, $\mathtt{B}^{\mathcal{I}}$ and $\mathtt{r}^{\mathcal{I}}$ is revised if there are other s_n that $\mathtt{A}^{\mathcal{I}}$ is not zero. In $\mathcal{O}_4$, there is no revision.

5 Experiment Setting

5.1 Performance Evaluation

Masked ABox Revision In the masked ABox revision task, we used 6 ontologies ("Ontodm" and "Nifdys" are not consistent), while in the conjunctive query answering task, we used 4 consistent ontologies.

The ontologies used for the experiments are taken from Bioportal[2], which, currently, includes more than 700 biomedical ontologies from different sources. We require the

[2] http://bioportal.bioontology.org/ontologies

ontologies to have at least the logical operator of negation, disjunction, or universal quantifier, as well as 100 ABox assertions. Five ontologies fall into this set, with two of them ("Ontodm" and "Nifdys") not consistent in some assertions; it remains to see whether DF-$\mathcal{ALC}$ would revise these errors. A taxonomy ontology (Sso) is also added for comparison. We also test a terseness ontology "Family", which contains multiple instantiated families but its knowledge is incomplete. Based on "Family", we augment it into "Family2" by adding some knowledge that can bridge with the instantiation. The information about these ontologies is shown in Table 3. Adam optimizer was used with a learning rate of 2e-4 to learn the grounding. Early stopping with 10 epochs tolerance was used to limit the running time.

	Family	Family2	GlycoRDF	Nifdys	Nihss	Ontodm	Sso
# TBox axioms	2032	2054	1453	6435	318	3476	2050
# ABox axioms	224	224	518	2920	146	1113	366
# Concepts	19	19	113	2751	18	838	176
# Roles	4	4	91	68	16	78	22
# Individuals	202	202	219	102	106	187	158
Expressivity	$\neg$	$\neg, \sqcap, \exists$	$\neg, \sqcup, \exists$	$\neg, \sqcup, \exists$	$\neg$	$\neg, \sqcap, \sqcup, \exists, \forall$	/

Table 3: Ontology information

The mask rate of ABox ranges from $\{20\%, 40\%, 60\%, 80\%\}$. We set the unknown region as $[0.2, 0.8]$. Meanwhile, the truth values greater (less) than $\alpha = 0.8$ ($1 - \alpha = 0.2$) were assumed to be true (false).

Upon further investigation, it was discovered that the cause of grounding failures in the "Family2" case was the presence of "unknown values". Specifically, the individual "F6M80" was asserted as an instance of "Male", but its parents were not identified, leading to unknown values for "Son(F6M80)" and "Child(F6M80)". Despite all the values falling within the unknown region [0.2, 0.8], the fact that $\mathrm{Son}^{\mathcal{I}''}(\mathrm{F6M80}) = 0.5490 > \mathrm{Child}^{\mathcal{I}''}(\mathrm{F6M80}) = 0.5489$ could still validate that $\mathrm{Son}^{\mathcal{I}''} \not\sqsubseteq \mathrm{Child}^{\mathcal{I}''}$ in the learned grounding of Γ.

Conjunctive Query We choose two forms of conjunctive queries. For example, a conjunctive query might ask for all individuals that are both male and have a PhD degree. This type of query can be expressed in the form of $C \sqcap D$, where C and D are atomic concepts that represent the two conditions. Another type of conjunctive query involves the use of role names, which represent relationships between individuals. For example, a query might ask for all individuals who are male and married to someone with a PhD degree. This type of query can be expressed in the form of $C \sqcap \exists r.D$, where r is a role name that represents the "married to" relationship. By testing conjunctive queries in the forms of $C \sqcap D$ and $C \sqcap \exists r.D$, we can evaluate the ability of a grounding learning model to retrieve individuals based on a revised grounding.

We generated 20 queries in each form, and the answer set of each query was not empty. Considering the time complexity of using a logical reasoner to get the true answer set, we only used two forms of conjunctive queries (CQs) in-depth 2 (the depth is determined by the conjunction amounts in the query). We chose all individuals with $Q^{\mathcal{I}}(a) \geq 0.8$ to be the answer for query Q. And use the answers generated by logical reasoner as ideal answers to evaluate the predicted answers with precision and recall as metrics.

Bibliography

Damien Arvor, Mariana Belgiu, Zoe Falomir, Isabelle Mougenot, and Laurent Durieux. Ontologies to interpret remote sensing images: why do we need them? *GIScience & remote sensing*, 56(6):911–939, 2019.

Stephen H Bach, Matthias Broecheler, Bert Huang, and Lise Getoor. Hinge-loss markov random fields and probabilistic soft logic. 2017.

Samy Badreddine, Artur d'Avila Garcez, Luciano Serafini, and Michael Spranger. Logic tensor networks. *Artificial Intelligence*, 303:103649, 2022.

Lorenz Bühmann, Jens Lehmann, and Patrick Westphal. Dl-learner—a framework for inductive learning on the semantic web. *Journal of Web Semantics*, 39:15–24, 2016.

Le-Wen Cai, Wang-Zhou Dai, Yu-Xuan Huang, Yu-Feng Li, Stephen Muggleton, and Yuan Jiang. Abductive learning with ground knowledge base. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1815–1821, 2021.

Angelo Cangelosi. Solutions and open challenges for the symbol grounding problem. *International Journal of Signs and Semiotic Systems (IJSSS)*, 1(1):49–54, 2011.

Jiaoyan Chen, Pan Hu, Ernesto Jimenez-Ruiz, Ole Magnus Holter, Denvar Antonyrajah, and Ian Horrocks. Owl2vec*: Embedding of owl ontologies. *Machine Learning*, pages 1–33, 2021.

Silvia Coradeschi, Amy Loutfi, and Britta Wrede. A short review of symbol grounding in robotic and intelligent systems. *KI-Künstliche Intelligenz*, 27(2):129–136, 2013.

Stamatia Dasiopoulou, Ioannis Kompatsiaris, and Michael G Strintzis. Applying fuzzy dls in the extraction of image semantics. In *Journal on data semantics XIV*, pages 105–132. Springer, 2009.

Michelangelo Diligenti, Marco Gori, and Claudio Sacca. Semantic-based regularization for learning and inference. *Artificial Intelligence*, 244:143–165, 2017.

Ivan Donadello, Luciano Serafini, and Artur D'Avila Garcez. Logic tensor networks for semantic image interpretation. *arXiv preprint arXiv:1705.08968*, 2017.

Aaron Eberhart, Monireh Ebrahimi, Lu Zhou, Cogan Shimizu, and Pascal Hitzler. Completion reasoning emulation for the description logic el+. *arXiv preprint arXiv:1912.05063*, 2019.

Monireh Ebrahimi, Aaron Eberhart, and Pascal Hitzler. On the capabilities of pointer networks for deep deductive reasoning. *arXiv preprint arXiv:2106.09225*, 2021.

Richard Evans and Edward Grefenstette. Learning explanatory rules from noisy data. *Journal of Artificial Intelligence Research*, 61:1–64, 2018.

Dinesh Garg, Shajith Ikbal, Santosh K Srivastava, Harit Vishwakarma, Hima Karanam, and L Venkata Subramaniam. Quantum embedding of knowledge for reasoning. *Advances in Neural Information Processing Systems*, 32:5594–5604, 2019.

Stevan Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42 (1-3):335–346, 1990.

Steffen Hölldobler, Yvonne Kalinke, and Hans-Peter Störr. Approximating the semantics of logic programs by recurrent neural networks. *Applied Intelligence*, 11(1): 45–58, 1999.

Zhiting Hu, Xuezhe Ma, Zhengzhong Liu, Eduard Hovy, and Eric Xing. Harnessing deep neural networks with logic rules. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2410–2420, 2016.

Yu-Xuan Huang, Wang-Zhou Dai, Jian Yang, Le-Wen Cai, Shaofen Cheng, Ruizhang Huang, Yu-Feng Li, and Zhi-Hua Zhou. Semi-supervised abductive learning and its application to theft judicial sentencing. In *2020 IEEE international conference on data mining (ICDM)*, pages 1070–1075. IEEE, 2020.

Celine Hudelot, Nicolas Maillot, and Monique Thonnat. Symbol grounding for semantic image interpretation: from image data to semantics. In *Tenth IEEE International Conference on Computer Vision Workshops (ICCVW'05)*, pages 1875–1875. IEEE, 2005.

Céline Hudelot, Jamal Atif, and Isabelle Bloch. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, 159(15):1929–1951, 2008.

Mathias Jackermeier, Jiaoyan Chen, and Ian Horrocks. Box^{2}el: Concept and role box embeddings for the description logic el++. *arXiv preprint arXiv:2301.11118*, 2023.

Hang Jiang, Sairam Gurajada, Qiuhao Lu, Sumit Neelam, Lucian Popa, Prithviraj Sen, Yunyao Li, and Alexander Gray. Lnn-el: A neuro-symbolic approach to short-text entity linking. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 775–787, 2021.

Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A Shamma, et al. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *International journal of computer vision*, 123(1):32–73, 2017.

Raghu Krishnapuram, Swarup Medasani, Sung-Hwan Jung, Young-Sik Choi, and Rajesh Balasubramaniam. Content-based image retrieval based on a fuzzy approach. *IEEE transactions on knowledge and data engineering*, 16(10):1185–1199, 2004.

Maxat Kulmanov, Wang Liu-Wei, Yuan Yan, and Robert Hoehndorf. El embeddings: geometric construction of models for the description logic el++. *arXiv preprint arXiv:1902.10499*, 2019.

Chen Liang, Jonathan Berant, Quoc Le, Kenneth Forbus, and Ni Lao. Neural symbolic machines: Learning semantic parsers on freebase with weak supervision. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 23–33, 2017.

Robin Manhaeve, Sebastijan Dumancic, Angelika Kimmig, Thomas Demeester, and Luc De Raedt. Deepproblog: Neural probabilistic logic programming. *Advances in Neural Information Processing Systems*, 31, 2018.

Carlo Meghini, Fabrizio Sebastiani, and Umberto Straccia. A model of multimedia information retrieval. *Journal of the ACM*, 48(5):909–970, 2001.

Sutapa Mondal, Sumit Bhatia, and Raghava Mutharaju. Emel++: Embeddings for εl++ description logic. In *AAAI Spring Symposium on Combining Machine Learning and Knowledge Engineering*. CEUR-WS, 2021.

Bernd Neumann and Ralf Möller. On scene interpretation with description logics. *Image and Vision Computing*, 26(1):82–101, 2008.

Alessandro Oltramari, Jonathan Francis, Cory Henson, Kaixin Ma, and Ruwan Wickramarachchi. Neuro-symbolic architectures for context understanding. *arXiv preprint arXiv:2003.04707*, 2020.

Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine learning*, 62(1):107–136, 2006.

Ryan Riegel, Alexander Gray, Francois Luus, Naweed Khan, Ndivhuwo Makondo, Ismail Yunus Akhalwaya, Haifeng Qian, Ronald Fagin, Francisco Barahona, Udit Sharma, et al. Logical neural netwroks. *arXiv preprint arXiv:2006.13155*, 2020.

Tim Rocktäschel and Sebastian Riedel. Learning knowledge base inference with neural theorem provers. In *Proceedings of the 5th workshop on automated knowledge base construction*, pages 45–50, 2016.

Tim Rocktäschel and Sebastian Riedel. End-to-end differentiable proving. *Advances in neural information processing systems*, 30, 2017.

Prithviraj Sen, Breno WSR de Carvalho, Ryan Riegel, and Alexander Gray. Neuro-symbolic inductive logic programming with logical neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 8212–8219, 2022.

Luciano Serafini and Artur d'Avila Garcez. Logic tensor networks: Deep learning and logical reasoning from data and knowledge. *arXiv preprint arXiv:1606.04422*, 2016.

Dominic Seyler, Tatiana Dembelova, Luciano Del Corro, Johannes Hoffart, and Gerhard Weikum. A study of the importance of external knowledge in the named entity recognition task. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 241–246, 2018.

Fatima Zohra Smaili, Xin Gao, and Robert Hoehndorf. Onto2vec: joint vector-based representation of biological entities and their ontology-based annotations. *Bioinformatics*, 34(13):i52–i60, 2018.

Fatima Zohra Smaili, Xin Gao, and Robert Hoehndorf. Opa2vec: combining formal and informal content of biomedical ontologies to improve similarity-based prediction. *Bioinformatics*, 35(12):2133–2140, 2019.

Umberto Straccia. An ontology mediated multimedia information retrieval system. In *2010 40th IEEE International Symposium on Multiple-Valued Logic*, pages 319–324. IEEE, 2010.

Haitian Sun, Andrew Arnold, Tania Bedrax Weiss, Fernando Pereira, and William W Cohen. Faithful embeddings for knowledge base queries. *Advances in Neural Information Processing Systems*, 33:22505–22516, 2020.

Sever Topan, David Rolnick, and Xujie Si. Techniques for symbol grounding with satnet. *Advances in Neural Information Processing Systems*, 34:20733–20744, 2021.

Kentaro Torisawa et al. Exploiting wikipedia as external knowledge for named entity recognition. In *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*, pages 698–707, 2007.

Geoffrey G Towell and Jude W Shavlik. Knowledge-based artificial neural networks. *Artificial intelligence*, 70(1-2):119–165, 1994.

Efthymia Tsamoura, Timothy Hospedales, and Loizos Michael. Neural-symbolic integration: A compositional perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 5051–5060, 2021.

Ozgul Unal, Hamideh Afsarmanesh, et al. Using linguistic techniques for schema matching. In *ICSOFT (2)*, pages 115–120, 2006.

Emile van Krieken, Erman Acar, and Frank van Harmelen. Analyzing differentiable fuzzy logic operators. *Artificial Intelligence*, 302:103602, 2022.

Po-Wei Wang, Priya Donti, Bryan Wilder, and Zico Kolter. Satnet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver. In *International Conference on Machine Learning*, pages 6545–6554. PMLR, 2019.

William Yang Wang, Kathryn Mazaitis, and William W Cohen. Programming with personalized pagerank: a locally groundable first-order probabilistic logic. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 2129–2138, 2013.

Fan Yang, Zhilin Yang, and William W Cohen. Differentiable learning of logical rules for knowledge base reasoning. *Advances in neural information processing systems*, 30, 2017.

Zhi-Hua Zhou. Abductive learning: towards bridging machine learning and logical reasoning. *Science China Information Sciences*, 62(7):1–3, 2019.