Integrating Knowledge Graphs With Logic Tensor Networks

Master’s Degree in Mathematical Engineering

Nicola Di Salvatore (318236)

November 25, 2024

Politecnico di Torino

Table of Contents

Introduction

- **Neuro-Symbolic Artificial Intelligence:** fields of artificial intelligence which integrates neural (data-driven) and symbolic (logic-driven) approaches;
- **Scene Graph Generation:** generates triples $\langle subject, object, relationship\rangle$ that describe an image (i.e., $\langle dog, surfboard, on\rangle$);
- **Knowledge bases:** repositories of structured information (i.e. Wordnet, ConceptNet, Visual Genome);
- **Knowledge embeddings:** representation of knowledge base entities and relationships as vectors (i.e. ConceptNet Numberbatch);
- **Logic Tensor Networks:** framework to inject prior knowledge in neural networks;
- **Entity Alignment:** association of corresponding elements among different knowledge bases;
- **First-Order Logic Statements:** input for logic tensor networks framework;
- **Objective of the thesis:** improve Visual Genome scene graphs by injecting prior knowledge from ConceptNet through logic tensor networks framework.

Table of Contents

2 Background

- **WordNet:** lexical knowledge base, where the core elements are synsets, characterized by:
  - a **concept**, with its **definition** and **part-of-speech** (noun, verb, preposition, adverb) ;
  - set of **lemmas**: individual words with the same meaning as the synset;
  - **relations** with other synsets;
- **ConceptNet:** common-sense knowledge base structured as a graph (nodes are concepts, relationships are edges. The main relationships considered are:
  - **Antonym/Synonym:** connects concepts with opposite/identical meanings;
  - **CapableOf/NotCapableOf:** indicates that a concept can/can't do the linked concept;
  - **InstanceOf** and **IsA:** connect a concept to its category;
- **Numberbatch:** associates each ConceptNet concept with a 300-dimensional array;
- **Visual Genome:** dataset made of 108,077 images annotated with:
  - **Objects:** main entities present in an image;
  - **Attributes:** entities that describe properties or qualities of objects;
  - **Relationships:** connect two objects with verbs or prepositions;
  - **Synsets:** represent categories of objects, attributes and relationships;

Logic Tensor Networks

- Real-Logic, the first-order logic language, has a signature (that is, the list of non-logical symbols) composed of the following elements:
  - $\mathcal{C}$: set of constants symbols c (e.g. img_1, img_2, ..., img_n);
  - $\mathcal{F}$: set of functions symbols, which transform one or more elements into another;
  - $\mathcal{P}$: set of predicates symbols, representing relationships or properties;
  - $\mathcal{X}$: set of variables symbols, that can assume different values.
- Grounding: associates each signature element to a numerical value:
  - $\mathcal{G}(c) \in \mathbb{R}^n$: grounding of variables $c \in \mathcal{C}$ by a real-valued vector in $\mathbb{R}^n$;
  - $\mathcal{G}(f) : \mathbb{R}^{n \cdot \alpha(f)} \to \mathbb{R}^n$: grounding of functions $f \in \mathcal{F}$;
  - $\mathcal{G}(p) : \mathbb{R}^{n \cdot \alpha(p)} \to [0, 1]$: grounding of predicates $p \in \mathcal{P}$.
- Let Θ be the set of parameters for the grounding functions. Logic Tensor Networks objective can be stated as an optimization problem:

$$\Theta^* = \arg\max_{\Theta} \mathcal{G}\left(\bigwedge_{\phi \in K} \phi \middle| \Theta\right) - \lambda \|\Theta\|_2^2,$$

Table of Contents

Alignment

3 Methodologies

The **alignment process** links each Visual Genome **object, attribute, relationship** (and, in particular, their **list of synsets**) to the most similar ConceptNet **concept**;

Load Data (Object/Attribute/Relationship)

↓

Retrieve WordNet synsets

↓

Extract Lemmas Embeddings and Part-Of-Speech

↓

Calculate Mean Embedding

↓

Find 5 Similar Concepts

↓

Select Aligned ConceptNet Concept

Object 1060294 → Synset calendar.n.01 → ... → Concept /c/en/calendar

FOL Statements Generation I

The following legend is used:

- $\mathcal{O}, \mathcal{P}$: sets of objects/attributes and relationships in Visual Genome;
- $\mathcal{O}', \mathcal{P}'$: set of objects/attributes and relationships classes;
- $\mathcal{O}_x^+$: set of hypernyms of the object/attribute x (found using the IsA relation in Conceptnet);
- $\mathcal{L}_\mathcal{O}, \mathcal{L}_\mathcal{P}, \mathcal{L}_{\mathcal{O}'}, \mathcal{L}_{\mathcal{P}'}$: labels sets of $\mathcal{O}, \mathcal{P}, \mathcal{O}', \mathcal{P}'$;
- $\mathcal{E}_{l_g}, \mathcal{N}_{l_g}$: set of relationships labels $l_{g'}$, where g' is a synonym, antonym of g (through the Synonym, Antonym relation in ConceptNet and further filtering using the similarity among embeddings);
- $\mathcal{PD}_{z'}, \mathcal{PR}_{z'}, \mathcal{ND}_{z'}$: the subset of $l_{\mathcal{O}'}$ containing objects/attributes labels that compose the positive, positive range, negative domain of the relationship z', obtained by using Visual Genome objects/attributes and Conceptnet relations $CapableOf$, IsA, $NotCapableOf$.

FOL Statements Generation II

Using the notation previously introduced, the first categories of axioms that have been built are the following:

- **Ontological Axioms:**
  - For Visual Genome relationships:
    If $z' \in \mathcal{O}_z^+ \cap \mathcal{P}'$, $l_z \in \mathcal{L}_\mathcal{P}$, $l_{z'} \in \mathcal{L}_{\mathcal{P}'}$, $\forall x, y \in \mathcal{O}$, $(l_z(x,y) \rightarrow l_{z'}(x,y))$
    Example: $\forall x, y \in \mathcal{O}$, (/c/en/inside(x,y) $\rightarrow$ /c/en/near(x,y));
- **Equivalence Axioms:**
  - For Visual Genome relationships:
    If $z \in \mathcal{P}$, $l_z \in \mathcal{L}_\mathcal{P}$, $l_{z'} \in \mathcal{E}_{l_z}$, $\forall x, y \in \mathcal{O}$, $(l_z(x,y) \leftrightarrow l_{z'}(x,y))$
    Example: $\forall x, y \in \mathcal{O}$, (fight(x, y) $\leftrightarrow$ battle(x, y));
- **Negative (or Mutual Exclusivity) Axioms:**
  - For Visual Genome relationships:
    If $z, z' \in \mathcal{P}$, $l_z \in \mathcal{N}_{l_{z'}}$, $\forall x, y \in \mathcal{O}$, $(\neg l_z(x,y) \vee \neg l_{z'}(x,y))$
    Example: $\forall x, y \in \mathcal{O}$, (¬/c/en/sit(x,y) $\vee$ ¬/c/en/walk(x,y));

- **Positive Domain Axioms:**
  If $z' \in \mathcal{P}', \ \forall x, y \in \mathcal{O}, \ \left(l_{z'}(x,y) \to \bigvee_{l_{x'} \in \mathcal{PD}_{z'}} \ l_{x'}(x) \right)$
  Example: $\forall x, y \in \mathcal{O}$, (/c/en/wear(x,y) $\to$ /c/en/biped(x) $\vee$ /c/en/person(x) $\vee$ /c/en/being(x) $\vee$ /c/en/animal(x));
- **Positive Range Axioms:**
  If $z' \in \mathcal{P}', \ \forall x, y \in \mathcal{O}, \ \left(l_{z'}(x,y) \to \bigvee_{l_{y'} \in \mathcal{PR}_{z'}} \ l_{y'}(y) \right)$
  Example: $\forall x, y \in \mathcal{O}$, (/c/en/wear(x,y) $\to$ /c/en/surface(y) $\vee$ /c/en/tool(y) $\vee$ /c/en/physical_object(y) $\vee$ /c/en/substance(y));
- **Negative Domain Axioms:**
  If $z' \in \mathcal{P}', \forall x, y \in \mathcal{O}, \ \left(l_{z'}(x,y) \to \bigwedge_{l_{x'} \in \mathcal{ND}_{z'}} \neg l_{x'}(x) \right)$
  Example: $\forall x, y \in \mathcal{O}$, (/c/en/wear(x,y) $\to$ $\neg$/c/en/heavier_than_air(x)).

FOL Statements Generation IV

1. The set $\mathcal{S}_{l_{z'}}$ of the semantically similar relationships to z' is built;
2. For each $l_z \in \mathcal{S}_{l_{z'}}$ all the couples subject-relationships $\langle x, z \rangle$ are searched in Visual Genome, and the labels l_x of those subjects are added to the set $\mathcal{PD}_{z'}$;
3. For each $l_z \in \mathcal{S}_{l_{z'}}$ all the triples of the form l_x "CapableOf" l_z are extracted from ConceptNet and added to the set $\mathcal{PD}_{z'}$;
4. For each $l_z \in \mathcal{PD}_{z'}$, all the hypernyms of l_z (obtained through the ConceptNet relationship IsA) are added to $\mathcal{PD}_{z'}$;
5. The frequency of each $l_z \in \mathcal{PD}_{z'}$ is calculated and the label representing the 0.9-quantile is stored ;
6. The set $\mathcal{PD}_{z'}$ is filtered maintaining only the elements in $\mathcal{O}'$ and with a frequency higher than the threshold (the 0.9-quantile) calculated in step 5.

Table of Contents

4 Experimental Results

Alignment

4 Experimental Results

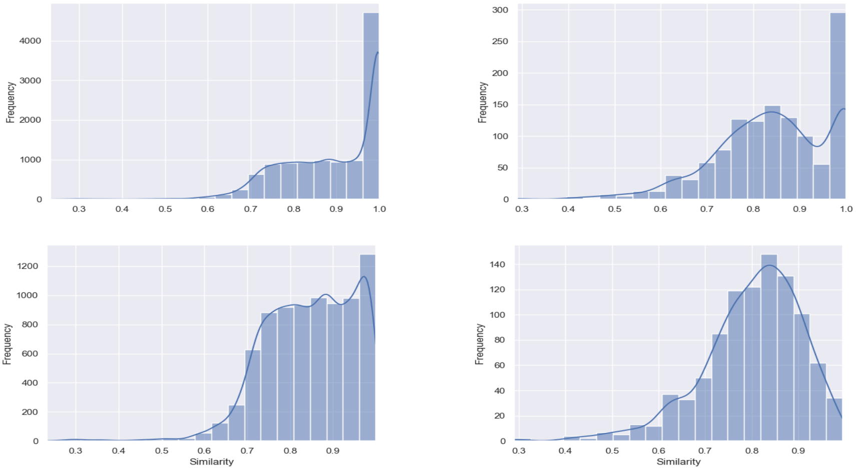
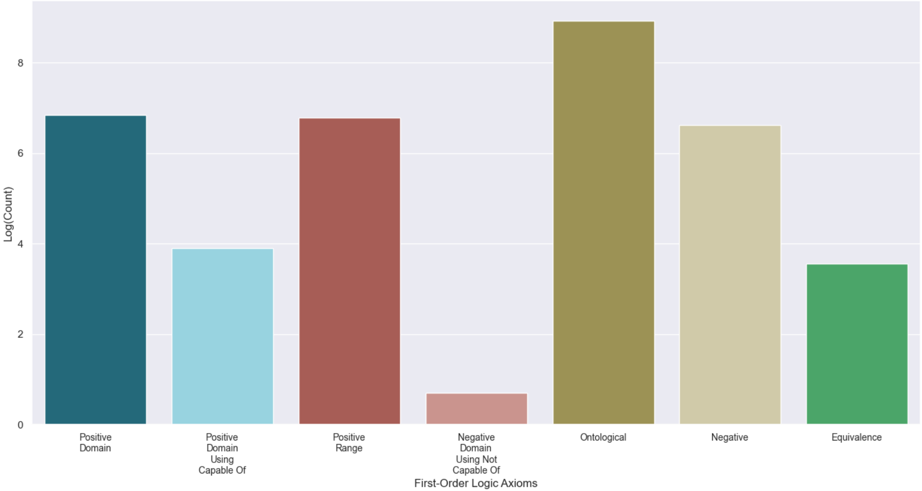


Figure: Best similarity for objects/attributes and relationships

FOL Statements Generation I

4 Experimental Results



FOL Statements Generation II

Table: First-Order Logic Axioms Count For Images

	Mean	Standard Deviation	25th Perc.	75th Perc.
PD	572.42	243.30	351.00	774.00
PR	518.96	181.20	417.00	660.00
PD - *CapableOf*	24.57	10.48	18.00	34.00
ND - *NotCapableOf*	0.97	0.65	1.00	1.00
Negative	355.64	172.27	233.00	494.00
Ontological	35.54	24.65	16.00	50.00
Equivalence	0.08	0.28	0.00	0.00
Total	1507.38	590.76	1154.00	1976.00

FOL Statements Generation III

Figure: Image From Visual Genome

FOL Statements Generation IV

- **Positive Domain Axioms:**
  - /c/en/wear(x, y) $\rightarrow$ /c/en/biped(x) $\vee$ /c/en/person(x) $\vee$ /c/en/being(x) $\vee$ /c/en/animal(x),
- **Positive Range Axioms:**
  - /c/en/next(x, y) $\rightarrow$ /c/en/surface(y) $\vee$ /c/en/tool(y) $\vee$ /c/en/physical_object(y) $\vee$ /c/en/substance(y)
- **Positive Domain Using** $CapableOf$ **Axioms:**
  - /c/en/transport(x, y) $\rightarrow$ /c/en/artifact(x) $\vee$ /c/en/machine(x) $\vee$ /c/en/physical_object(x) $\vee$ /c/en/vehicle(x),
- **Negative Domain using** $NotCapableOf$**:**
  - /c/en/laugh(x, y) $\rightarrow$ $\neg$/c/en/substance(x) $\wedge$ $\neg$/c/en/food(x) $\wedge$ $\neg$/c/en/fuel(x) $\wedge$ $\neg$/c/en/good(x),
- **Ontological Axioms:**
  - /c/en/man(z) $\rightarrow$ /c/en/being(z),
- **Negative Axioms:**
  - $\neg$/c/en/preditor(x) $\vee$ $\neg$/c/en/physical_object(x),

FOL Statements Generation V

Axioms obtained by integrating the domain/range axioms with the ontological ones. They allow to visualize better the results:

- **Positive domain axioms**:
  - /c/en/wear(x, y) $\rightarrow$ /c/en/man(x),
  - /c/en/have(x, y) $\rightarrow$ /c/en/car(x) $\vee$ /c/en/van(x) $\vee$ /c/en/building(x),
  - /c/en/behind(x, y) $\rightarrow$ /c/en/van(x) $\vee$ /c/en/car(x),
  - /c/en/along(x, y) $\rightarrow$ /c/en/bicycle(x) $\vee$ /c/en/van(x) $\vee$ /c/en/car(x),
- **Positive range axioms**:
  - c/en/wear(x, y) $\rightarrow$ /c/en/car(y),
  - /c/en/have(x, y) $\rightarrow$ /c/en/building(y) $\vee$ /c/en/window(y),
  - /c/en/transport(x, y) $\rightarrow$ /c/en/car(y) $\vee$ /c/en/building(y),
  - /c/en/next(x, y) $\rightarrow$ /c/en/road(y) $\vee$ /c/en/car(y) $\vee$ /c/en/building(y).
- **Positive domain using $CapableOf$ axioms:**
  - /c/en/transport(x, y) $\rightarrow$ /c/en/car(x).

Table of Contents

5 Conclusion

Automatic generation of first-order logic statements:

- **Positive Domain** and **Positive Range** Axioms: high cardinality, high density in images (572 and 519 for the image on average);
- **Positive Domain Using** *CapableOf* Axioms: medium cardinality, medium density in images (25 for image on average);
- **Negative Domain**: low cardinality, using the ConceptNet relationship *NotCapableOf*, medium density in images (1 for image on average);
- **Ontological** axioms: high cardinality, low density in images, used to link objects/attributes with their hypernyms;
- **Negative** (or **Mutual Exclusivity**) Axioms: high cardinality, high density in images, high specificity
- **Equivalence** Axioms: medium cardinality, low density in images, help to associate one concept with another.

Future Works

- **Inserting** axioms into **neural networks** through logic tensor networks;
- Generating **range** and **domain** sets **for couples** $\langle subject, relationship \rangle$ and $\langle object, relationship \rangle$ (for example, /c/en/tool is in the set $\mathcal{PR}_{\langle /c/en/being, /c/en/wear \rangle}$);
- inserting **other categories** of first-order logic axioms, such as **inverse axioms** (i.e. $\forall x, y \in \mathcal{O},\ ($/c/en/inside(x,y) $\leftrightarrow$ /c/en/outside(y,x)$)$);
- Building the sets $\mathcal{PD},\ \mathcal{PR},\ \mathcal{ND},\ \mathcal{NR}$ without a deterministic algorithm, but with **first-order logic axioms containing new LTN predicates** such as:

$$\forall x \in \mathcal{O},\ \forall x' \in \mathcal{O}',\ \forall z \in \mathcal{P},\ \forall z' \in \mathcal{P}' :$$
$$\text{HasKnowledgeBaseLinkForDomain}(x, z)\ \wedge\ \text{SemanticallySimilar}(z', z)$$
$$\wedge\ \text{InHierarchy}(x, x')\ \implies\ \text{InDomain}(x', z')$$

Integrating Knowledge Graphs With Logic Tensor Networks

Thank you for listening!

Any questions?