Markov Logic Network for Metaphor Set Expansion

Jaya Pathak[a] and Pratik Shah[b]
Indian Institute of Information Technology Vadodara, Gandhinagar, India

Keywords: Metaphor Identification, Markov Logic Network (MLN), Information Completion.

Abstract: Metaphor is a figure of speech, that allow us to understand a concept of a domain in terms of the other. One of the sub-problems related to the metaphor recognition is of metaphor set expansion. This in turn is an instance of information completion problem. We, in this work, propose an MLN based approach to address the problem of metaphor set expansion. The rules for metaphor set expansion are represented in the first order logic formulas. The rules are either soft or hard depending on the nature of the rules according to which corresponding logic formulas are then assigned weights. Many a times new metaphors are created based on usages of Is-A pair knowledge base. We, in this work model this phenomena by introducing appropriate predicates and formulas in clausal form. For experiments, we have used dataset from Microsoft concept graph consisting Is-A patterns. The experiments show that the weights for the formulas can be learnt using the training dataset. Moreover the formulas and their weights are easy to interpret and in-turn explains the inference results adequately. We believe that this is a first effort reported which uses MLN for metaphor set expansion.

1 INTRODUCTION

Metaphor plays a vital role in expressing and communicating human emotions, ideas and concepts. Expressing unknown in terms of known is a key to learning. We use metaphors in our daily lives for effective communication. For example, *"John is a Shining Star"*, tells us about John's personality and achievements by comparing the properties of a star with John. Metaphor is a mapping of concepts from a source domain to a target domain (Lakoff and Johnson, 1980). The source concepts are mapped to target when they share some common traits. In, 'John is a Shining Star', the source concept is *Shining Star* and target concept is *John*.

To understand metaphoric figures, the knowledge of various concepts (a person, animal, etc) is useful. Psychologist Gregory Murphy stated that "Concepts are the glue that holds our mental world together". So for a metaphor to make sense, any of the two concepts should be explicitly defined, i.e. either the source or the target. In the above example, both the source (*Shining Star*) and target the (*John*) are clearly defined.

[a] https://orcid.org/0000-0001-9368-8753
[b] https://orcid.org/0000-0002-4558-6071

Before identifying a source-target pair, we should recognize a metaphoric sentence. There are several existing work for recognizing metaphors. For a Type-I metaphor, an ideal way is to use the concept of Is-A relation. The sentence having Is-A pattern can be a potential metaphoric sentence. The first noun in a Is-A sentence is the target, while the second noun is source. After pairing them as $(Target, Source)$, check whether there is any knowledge about that pair. To differentiate metaphoric and literal sentences, two sets symbolized as Γ_m (Metaphor set) and Γ_H (Hearst pair set) respectively are used.

We in our work, model the metaphor set expansion as an information completion problem. We propose to solve the same using Markov logic Network. The two key challenges of machine learning concepts are uncertainty and complexity of the rule base. MLN merges the logical and statistical models into a single representation. One of the reasons to use MLN for information completion problem of metaphor set expansion is to infer about a complex and uncertain problem in an explainable manner. We use MLN for metaphor set expansion by formulating the formulas of the rule base in the first order logic. We work with Type-I metaphors, where source and target concepts are derived from Is-A pattern. The proposed MLN based approach is validated based on the experiments

Pathak, J. and Shah, P.
Markov Logic Network for Metaphor Set Expansion.
DOI: 10.5220/0010205606210628
In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021) - Volume 2*, pages 621-628
ISBN: 978-989-758-484-8

conducted. For experiments, we have used Tuffy (Niu et al., 2011). It is an efficient Markov logic network inference engine. For experiments, we have used dataset from Microsoft concept graph for short text understanding (Wu et al., 2012). It provides the core version of Is-A data mined from billions of web pages (Cheng et al., 2015). With the available Is-A relations we derive the set Γ_H of Hearst pairs. Hearst pairs are instances and concepts pairs i.e. *(Instance,Concept)*. Assuming that a small set of metaphors Γ_m is available, the task of identifying new metaphors and include them in Γ_m is called metaphor set expansion.

2 EXISTING WORK

A metaphor connects two concepts based on the content the resemblance between the concept domains. Resemblance describes the nature of the idea they share in common. A metaphor has two domains: the target (the one which is being targeted) and the source (the one whose characteristics are being used). The initial work on metaphor recognition known as *met*, was introduced by Fass in 1991 (Fass, 1991). It proposes a method for differentiating literals and metaphors. The problem with the approach is its hard decision rule base. Without a probabilistic framework, reasoning about metaphor explanation is not meaningful. Even when the probability assignment is done by one or the other approach, it is difficult to explain the inference drawn about given pair. The other approach use hand-coded knowledge which are still implausible and need technology boost. The previous works on recognizing and identifying metaphors uses contextual preferences, like the matching of similarity between predicate-object, object-object, etc which are difficult to recognize sometimes.

At present, there is a need for statistical methods to recognize and identify metaphors. The authors in (Shutova, 2010) put forward the need of creating a publicly available metaphor corpus consisting of known metaphoric sentences with the help of the existing dataset using statistical pattern matching.

Recently, in computational linguistics approach based on statistical inference have been applied to metaphor recognition. The authors in (Schulder and Hovy, 2014) use term relevance measure based on frequency of occurrence in target domains for metaphor detection. In (Tsvetkov et al., 2014), the authors uses lexical semantic features to distinguish between a metaphor and literal sentence by building a cross-lingual model.

In our work, we first identify the source-target mapping in sentences. The source domain is usually explicitly defined while the target is mostly unclear. Following are some of the existing methods for identification of metaphor set (Li et al., 2013).

2.1 Metaphor Identification

Detecting Is-A relation : There is a likelihood that sentences having Is-A pair are metaphors. But this is not true in general. For example, *"Apple is a fuel"* can be designated as a metaphor but *"Apple is a fruit"* has literal meaning and cannot be considered as a metaphor. Every Is-A relation need not always be a metaphor. Is-A pattern can be categorized in : the literal Is-A relation and metaphoric Is-A relation.

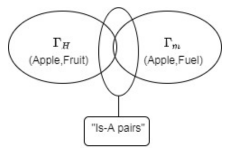

Figure 1: Relation between Γ_H, Γ_m and Is-A pairs (Li et al., 2013).

Γ_m **set of metaphors**, consists of metaphors in *(Source,Target)* pair form. For example, in *"Apple is a Fuel"*, $(Fuel, Apple)$ is the source-target pair, such that

$$(Apple, fuel) \in \Gamma_m$$

Hearst pattern data set Γ_H, is discussed in (Wu et al., 2012). This consists of literal Is-A relations in the form (x, h_x), a pair of *(hyponym,hypernym)* such that x be an instance of h_x. For example, *"Apple is a fruit"*, the Hearst pair

$$(x, h_x) \equiv (Apple, Fruit) \in \Gamma_H$$

2.2 Metaphor Set and Hearst Pair Set

Both the above sets do not overlap as the metaphors do not provide literal meaning in a sentence (Li et al., 2013). So, if any pair of Γ_m is also present in Γ_H then it is to be removed from Γ_m.

From the available data set from Microsoft concept graph of (Cheng et al., 2015; Wu et al., 2012) Is-A relations, we build the Hearst pairs Γ_H using pairs of concepts and instance as $(Instance, Concept)$. Given a sample data, "Apple is a Food" where Apple is an instance and Food is a concept, we get the pair $(Apple, Food)$.

The goal is to use the existing Γ_H and Γ_m pairs to expand Γ_m. Any pair is more likely to be a metaphor if it does not appear in the set extracted from the Hearst pattern.

To expand the set Γ_m set we use transitive properties on Is-A relations (Li et al., 2013). If $(x,y) \in \Gamma_m$ and $(x,h_x) \in \Gamma_H$, then by transitivity we can add (h_x,y) to Γ_m . For example, let $(Apple,Fuel) \in \Gamma_m$ and $(Apple,Food) \in \Gamma_H$ then, $(Food,Fuel) \in \Gamma_m$.

In (Li et al., 2013), the authors propose a probability assignment to a pair (x,y) as :

$$P(x,y) = \frac{occurrences\ of\ (x,y)\ in\ Is-A\ pattern}{occurrences\ of\ Is-A\ pattern} \quad (1)$$

We, use the transitivity as suggested above and include the same in the form of a first order formula representing the knowledge (rule base) about metaphor set expansion in MLN.

2.3 Probabilistic Model for Metaphor Set Expansion

Metaphor expansion problem is same as that of a information completion. We propose to use Markov logic network (MLN) for the same. MLN integrates of the logical and statistical models into a single representation. The knowledge representation is done using first order logic and probabilistic graphical model. Logic handles the casual relationship between rule base and probability deals with the uncertainty (Richardson and Domingos, 2006). We work with Type-I metaphors, and try to expand the existing metaphor set. With the closed world assumption, it is possible to extract metaphors from the existing knowledge base (Γ_m and Γ_H).

In the next section we give an introduction to Markov Logic Network and the notation that will be used for the subsequent sections. We will present the MLN with the help of a set of examples which will help formulate the metaphor set expansion problem in MLN framework.

3 MARKOV LOGIC NETWORK

Markov logic network is considered as a possibility for unified learning mechanism. It is a probabilistic graphical approach (Richardson and Domingos, 2006). MLN comprises of formulas in first-order logic form and to each of the formula a real valued weight is assigned. The nodes of the network graph are the terms from the ground formulas and the edges are the logical connectives among them. The nodes represents atomic variables and edges describe the probabilistic interaction between them. Consider a Markov network for formula h:

$$h: \ Instance(u) \rightarrow Concept(u)$$

constant set $C_1 = \{apple, orange, fruit, company\}$.

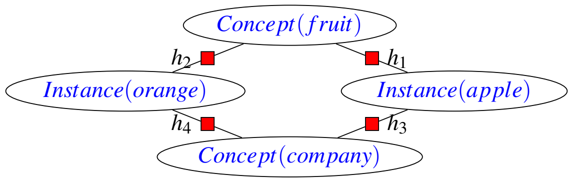

Figure 2: A formulated ground Markov logic Network.

Relation between variables are represented using First order logic in a Markov Network. First order logic makes representation easy as compared to other logical representations and are constructed using **variables, constants, functions and predicates**. From Figure 2, the variables $\{u,v\}$ range over the object domain and takes any value from the constant. Constants represent the objects in the domain of interest. A function maps from one object set to another e.g. *Instance(), Concept()*. Predicate represents the relationship between objects e.g. *metaphor(x,y)*. Predicate value can be either true or false. Using the above representations with logical connectives and quantifiers, **formulas** (clauses) are modeled.

A *ground atom* is a term taking a constant value from set and is the smallest unit of a formula. In figure 2, *Instance(apple)* is a ground atom with *'Instance'* as a function taking *'apple'* as input from the constant set C_1. An **atomic formula** is defined as a predicate term used with a constant set, $Instance(apple) \rightarrow Concept(fruit)$. Grounding refers to the method of removing variables with constant value. A **ground formula** is defined as set of predicates accompanied by connectives using constant values as inputs. In Figure 2, h_1,h_2,h_3,h_4 are the ground formulas for formula h. A **possible world** is described as the truth values $(0,1)$ assigned to each ground atom in a network. A world is defined as the set of truth values which can be assigned to all ground atoms (Domingos et al., 2008).

Together, Markov Network with First order logic are combined as Markov logic Network (MLN) (Richardson and Domingos, 2006).

3.1 Markov Logic Network

A Markov logic network (MLN) (Richardson and Domingos, 2006) is a set of weighted first order logic formulas (F,w), where F is the set of formulas in first order logic form and $w \in \mathbb{R}^{|F|}$. A ground network is formed w.r.t to the grounding of the formulas using a set of constants $C = \{c_1, c_2, ..., c_{|C|}\}$. With the constant set, the ground Markov network is built which contains sub-graph for each ground formula in the network. There are nodes for each ground predicate or clause in the network, if any ground predicate is TRUE, the value of the corresponding node will be 1.

For a set of constants, it can produce different networks. The probability distribution over a possible world ω stated by ground Markov network is given as

$$P(W = \omega) = \frac{1}{Z} exp\left(\sum_{i=1}^{F} w_i n_i(\omega)\right) = \frac{1}{Z} exp^{w_i * n_i} \quad (2)$$

where $n_i(\omega)$ is number of true groundings of F_i in ω, true grounding refers to the number of grounded formulas satisfied in specific world (Richardson and Domingos, 2006). The $\omega_i s$ are the truth values of the atoms in a specific world and Z is the partition function.

In the next section, we formulate the metaphor set expansion in the framework of MLN.

4 PROBLEM FORMULATION

4.1 Metaphor Set Expansion as Information Completion

We use MLN for metaphor set expansion by formulating the rules of expansion as formulas in first order logic. In this work we discuss Type-I metaphors only. We use this knowledge represented in the form of FOL formulas, to expand the existing metaphor set. By using the closed world assumption, we show how to extract metaphors from the existing knowledge base.

FOL formulas are derived below:

- If a pair (x, y) belongs in Γ_m, it is a metaphor. If the pair is in Is-A knowledge base Γ_H then that means it has literal meaning and is not a metaphor. If pair (x, y) belongs to Γ_m as well as in Γ_H , then it has to removed from Γ_m, *because a metaphor does not have a literal meaning.*

$$\infty \quad H(x,y) \implies \sim m(x,y) \quad (3)$$

  where $H(x, y)$ denotes (x,y) is a Hearst pair and $m(x, y)$ indicates (x,y) is a metaphor.

- If a new pair doesn't belong to any of the sets, there is a possibility that it is a new metaphor. We expand the Γ_m by obtaining metaphors derived from Γ_m and Γ_H. We use a known metaphor $m(x, y)$ and a compatible Hearst pattern $H(x, h_x)$ for deriving a new metaphor.

  Let $(x, y) \in \Gamma_m$ and $(x, h_x) \in \Gamma_H$, using transitive property we add this derived metaphor (h_x, y) to Γ_m.

$$w_1 \quad m(x,y) \wedge H(x,h_x) \implies m(h_x,y) \quad (4)$$

- Next, we use the new metaphor for further expansion.

$$w_2 \quad m(x,y) \wedge H(h_x,x) \implies m(h_x,y) \quad (5)$$

For each grounding of formula the weights are same. A formula with infinity (∞) as a weight indicates a hard rule.

4.2 MLN Constructed for Metaphor Set Expansion Problem

We present an example to illustrate the working of Markov logic network. Let H be the set of Hearst pairs and m be the metaphoric pairs, the two formulas (3 and 4) are re-written as

$$g_1 : H(x,y) \implies \sim m(x,y)$$

$$g_2 : m(x,y) \wedge H(x,z) \implies m(z,y)$$

Given a knowledge base with x, y and z as variables and constant set $C = \{apple, fuel, fruit\}$.The formula g_1 will have 9 instances, as shown in Table 1. Similarly, formula g_2 will have 27 instances shown in Table 2.

Combining all the ground predicates obtained from formula g_1 and g_2 together, we get the ground Markov network with 18 ground predicates. Figure 5 shows a section of the ground Markov network.

Table 1: Ground formulas for g_1 with all possible predicates Ⓝ indicates a ground predicate.

• $g_{1,1}$:	① $H(apple, fruit) \Rightarrow \sim m(apple, fruit)$ ②
• $g_{1,2}$:	③ $H(apple, fuel) \Rightarrow \sim m(apple, fuel)$ ④
• $g_{1,3}$:	⑤ $H(fruit, fuel) \Rightarrow \sim m(fruit, fuel)$ ⑥
• $g_{1,4}$:	⑦ $H(fruit, apple) \Rightarrow \sim m(fruit, apple)$ ⑧
• $g_{1,5}$:	⑨ $H(fuel, apple) \Rightarrow \sim m(fuel, apple)$ ⑩
• $g_{1,6}$:	⑪ $H(fuel, fruit) \Rightarrow \sim m(fuel, fruit)$ ⑫
• $g_{1,7}$:	⑬ $H(apple, apple) \Rightarrow \sim m(apple, apple)$ ⑭
• $g_{1,8}$:	⑮ $H(fruit, fruit) \Rightarrow \sim m(fruit, fruit)$ ⑯
• $g_{1,9}$:	⑰ $H(fuel, fuel) \Rightarrow \sim m(fuel, fuel)$ ⑱

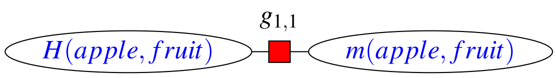

Figure 3: Predicates with factor nodes for formula g_1.

We represent the ground predicates in the form of graph as shown in Figure 3 and 4. Each ground predicate is represented as a node, if two nodes are related there is an edge between them and this relation among nodes is given by a factor node.

Using constant set C, we get 9 sub-graphs $g_{1,1} \ldots g_{1,9}$ shown in Figure 3, each with 2 nodes using formula g_1 and 27 sub-graph $g_{2,1} \ldots g_{2,27}$ each with 3 nodes using formula g_2 as shown in Figure 4.

4.3 Inference using MLN

The constructed MLN can now be used for inference. For the given evidence it can compute probability of the desired query. For the above example, we have $m(apple, fuel)$ and $H(apple, fruit)$ in the evidence set. And we wish to compute how likely the predicate $m(fruit, fuel)$ is. The corresponding conditional probability is,

$$P[m(fruit, fuel)|\ m(apple, fuel),\ H(apple, fruit)] = \frac{P[m(fruit,fuel)=1, m(apple,fuel)=1, H(apple,fruit)=1]}{P[m(apple,fuel)=1, H(apple,fruit)=1]} \tag{6}$$

To solve a network with 18 predicates (specified in Table I) there will be 2^{18} possible worlds. If a predicate is present in a specific world its value will be 1 else 0. Lets denote each predicate with a number as follows, $H(apple, fruit)$ numbered as ①, $m(apple, fuel)$ as ④,$m(fruit, fuel)$ as ⑥ and so on. If presence of predicate is unknown it is denoted as $*$ (value as 0 or 1). In equation 6 numerator, value of ①,④,⑥ = 1 and rest unknown $*$. Similarly, for denominator ① and ④ have value as 1. Substituting values for predicates in equation 6, we get

$$P[m(fruit, fuel)|\ m(apple, fuel),\ H(apple, fruit)] = \frac{\{1,*,*,1,*,1,*,*,*,*,*,*,*,*,*,*,*,*\}}{\{1,*,*,1,*,*,*,*,*,*,*,*,*,*,*,*,*,*\}}$$

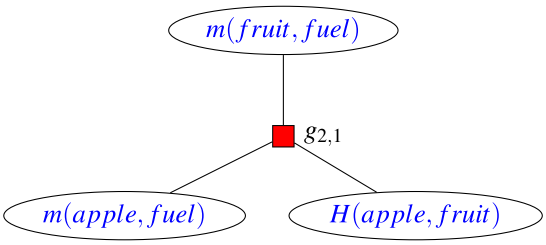

Figure 4: Factor node among related nodes.

Table 2: Ground formulas of g_2.

• $g_{2,1}$:	$m(apple,fuel) \wedge H(apple,fruit) \Rightarrow m(fruit,fuel)$
• $g_{2,2}$:	$m(apple,fruit) \wedge H(apple,fuel) \Rightarrow m(fruit,fuel)$
.	.
.	.
• $g_{2,26}$:	$m(fuel,fruit) \wedge H(fuel,apple) \Rightarrow m(apple,fruit)$
• $g_{2,27}$:	$m(fruit,Apple) \wedge H(fruit,fuel) \Rightarrow m(fuel,apple)$

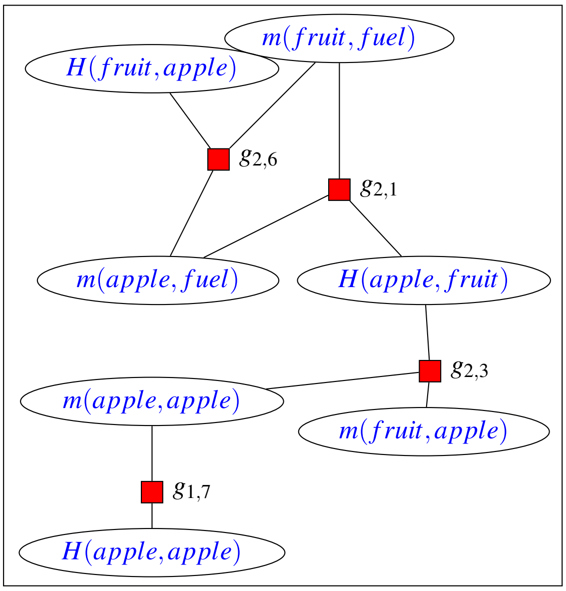

Figure 5: A section of ground Markov network.

5 EXPERIMENTAL RESULTS

5.1 Dataset and Tool

We have used dataset from Microsoft concept graph for short text understanding (Wu et al., 2012), that provides the core version of Is-A data mined from billions of web pages. This data contains 5,376,526 unique concepts, 12,501,527 unique instances, and 85,101,174 Is-A relations (Cheng et al., 2015). From these Is-A relations, we build the Hearst pairs Γ_H using pairs of concepts and instance as $(Instance, Concepts)$ and include related metaphor pairs in Γ_m to expand the existing metaphor set Γ_m. Experiments are carried in 'Tuffy' (Niu et al., 2011). It takes 3 input files: (1) **program.mln** stores predicates with their definitions and formulas with their respective weights. (2)**evidence.db** consisting available ground terms and (3) a **query.db**. The output to the inference result is **test.txt**.

5.2 Expansion of Metaphor Set in Case of Unique Hearst Pairs in the Evidence Set

To know the confidence of the outcome, we calculate the marginal probabilities for derived metaphor. These results may vary as Tuffy takes different samples each time. An example of a program for metaphor expansion given in section 4.2, probability values of the resultant marginal probability of

Table 3: A program for metaphor set expansion with unique Hearst pairs.

Program.mln	Evidence.db	Query	Output.txt
$*obj(object)$	$obj(apple)$	$m(x,y)$	
$*H(obj,obj)$	$obj(fruit)$		
$m(obj,obj)$	$obj(fuel)$		
$!H(a1,a2) \vee !m(a1,a2).$	$H(apple,fruit)$		
$1 \quad !m(a1,a2) \vee !H(a1,a3) \vee m(a3,a2)$	$m(apple,fuel)$		
$1 \quad !m(a1,a2) \vee !H(a3,a1) \vee m(a3,a2)$			0.7110 $m(\text{“}fruit,fuel\text{”})$
$10 \quad !m(a1,a2) \vee !H(a1,a3) \vee m(a3,a2)$			
$5 \quad !m(a1,a2) \vee !H(a3,a1) \vee m(a3,a2)$			0.9960 $m(\text{“}fruit,fuel\text{”})$
A MLN program with different evidence set.			
$*obj(object)$	$obj(apple)$	$m(x,y)$	0.994 $m(\text{“}brand,money\text{”})$
$*H(obj,obj)$	$obj(brand)$		0.783 $m(\text{“}business,money\text{”})$
$*HighH(obj,obj)$	$obj(money)$		0.684 $m(\text{“}gadget,money\text{”})$
$m(obj,obj)$	$obj(business)$		
$!HighH(a1,a2) \vee !m(a1,a2).$	$HighH(apple,brand)$		
$10 \quad !m(a1,a2) \vee !HighH(a1,a3) \vee m(a3,a2)$	$H(apple,business)$		
$1 \quad !m(a1,a2) \vee !H(a1,a3) \vee m(a3,a2)$	$m(apple,money)$		

Table 4: A program for Metaphor set expansion with Multiple Hearst pairs.

Program.mln	Evidence.db	Query.db	Average marginal probabilities
$*obj(object)$	$obj(apple)$	$m(x,y)$	0.9960 $m(\text{“}fruit,fuel\text{”})$
$*H(obj,obj)$	$obj(fuel)$		0.6730 $m(\text{“}gadget,fuel\text{”})$
$*HighH(obj,obj)$	$obj(fruit)$		
$m(obj,obj)$	$obj(gadget)$		
$!HighH(a1,a2) \vee !m(a1,a2).$	$HighH(apple,fruit)$		
$10 \quad !m(a1,a2) \vee !HighH(a1,a3) \vee m(a3,a2)$	$H(apple,gadget)$		
$1 \quad !m(a1,a2) \vee !H(a1,a3) \vee m(a3,a2)$	$m(apple,fuel)$		

```
0.7400  m("art", "piece_of_cake")
0.7000  m("activity", "lifestyle")
0.5500  m("company", "teacher")
0.8100  m("fruit", "fuel")
0.7400  m("activity", "journey")
0.7100  m("material", "boon")
0.7700  m("material", "curse")
0.5500  m("place", "paradise")
0.7000  m("activity", "light")
0.7800  m("search_engine", "teacher")
0.7000  m("social_media", "teacher")
0.6700  m("social_media", "drug")
0.5300  m("newyork", "paradise")
0.5500  m("florida", "paradise")
0.5400  m("facebook", "teacher")
0.5000  m("apple", "teacher")
0.7600  m("banana", "fuel")
0.6500  m("chocolate", "medicine")
0.5100  m("california", "paradise")
0.6900  m("cycling", "journey")
0.6600  m("fish", "medicine")
```

(a) Marginal probability values with unique Hearst pairs.

```
0.9970  m("art", "piece_of_cake")
0.943   m("activity", "lifestyle")
0.9600  m("company", "teacher")
0.9910  m("fruit", "fuel")
0.9870  m("activity", "journey")
0.9940  m("material", "boon")
0.9980  m("material", "curse")
0.9470  m("place", "paradise")
0.9200  m("activity", "light")
0.9820  m("search_engine", "teacher")
0.8800  m("social_media", "teacher")
0.9070  m("social_media", "drug")
0.7400  m("newyork", "paradise")
0.6480  m("florida", "paradise")
0.5660  m("facebook", "teacher")
0.4080  m("apple", "teacher")
0.7000  m("banana", "fuel")
0.7400  m("chocolate", "medicine")
0.8520  m("california", "paradise")
0.6900  m("cycling", "journey")
0.8200  m("fish", "medicine")
```

(b) Marginal probability value with multiple Hearst pairs.

Figure 6: Average marginal probabilities of pairs to be included in the expanded metaphor set.

$m(Fruit, Fuel) = 0.7110$ shown in Table 3.

Additionally, variation in weights of the formulas provides a way to control the contribution of formulas in computation of the probability value of the query. For instance, in Table 3, with higher weights, $w_1 = 10$ and $w_2 = 5$ results in higher probability values compared with the previous.

```
H(fever,symptom)
H(pain,symptom)
H(anxiety,symptom)
H(depression,symptom)
m(pain,gift)
m(pain, part_of_life)
m(soccer,religion)
H(football,sport)
H(tennis,sport)
H(soccer,sport)
m(apple,fuel)
H(banana,fruit)
H(peach,fruit)
High_H(apple,fruit)
High_H(google,search_engine)
m(google,teacher)
H(wedding,event)
m(wedding,journey)
m(wedding,blessing)
m(life,race)
m(life,journey)
High_H(life,process)
```

```
H(name,information)
H(address,information)
H(date,information)
High_H(facebook,social_media)
H(twitter,social_media)
m(twitter,trend)
H(facebook,site)
m(facebook,drug)
H(facebook,social_network)
High_H(plastic,material)
High_H(paper,material)
m(plastic,boon)
m(plastic,curse)
H(lead,metal)
H(aluminum,metal)
H(gold,metal)
High_H(google,company)
High_H(apple,company)
H(amazon,company)
H(facebook,company)
m(amazon,titan)
```

Figure 7: Knowledge base used in weight learning for metaphor expansion.

5.3 Expansion of Metaphor in Case of Multiple Occurrences of Hearst Pairs in Evidence Set

If a particular pair is occurring more than once in the dataset, it means the usage of that Is-A pair is

```
///////////////AVERAGE WEIGHT OF ALL THE ITERATIONS///////////////
!H(v0, v1)  v !m(v0, v1) . //1.0hardfixed
2.2585        !m(v0, v1)  v !High_H(v0, v2)  v  m(v2, v1)  //2.0
8.1846        !m(v0, v1)  v !H(v2, v0)  v  m(v2, v1)  //3.0

///////////////WEIGHT OF LAST ITERATION///////////////
!H(v0, v1)  v !m(v0, v1) . //1.0hardfixed
2.2585        !m(v0, v1)  v !High_H(v0, v2)  v  m(v2, v1)  //2.0
8.1846        !m(v0, v1)  v !H(v2, v0)  v  m(v2, v1)  //3.0
```

Figure 8: Weights learned using evidence given in Figure 6 as knowledge base.

Table 5: A small dataset for comparing both existing and proposed method for metaphor set expansion.

$x = \{Apple, Orange\}$	$h_x = \{fruit, company\}$
Hearst pairs	Metaphor pairs
$H(Apple, fruit)$ $H(Apple, company)$ $H(Orange, fruit)$	$m(Apple, fuel)$
$m(\text{"}orange, fuel\text{"}) = ?$	

high. We indicate this by introducing a predicate $HighH(a,b)$ as shown in Table 4, depicting Hearst pair $H(a,b)$ has high occurrence in evidence data set.

5.4 An Experiment on Larger Dataset

We next perform experiments on a relative larger data set. We use 344 Is-A pairs out of which 56 belongs to Γ_m and the rest in Γ_H. We experimented using single occurrence of Hearst pairs as well as multiple occurrences. The results are shown in Figure 6(a) and 6(b). We conclude that the average weight for the derived metaphors (h_x, y) are greater than that of the other extracted candidate metaphor pairs.

5.5 An Experiment on Weight Learning from Training Dataset

Weight learning in MLN is similar to inference program. MLN tries to learn the weights using the given evidence dataset as input and computes the weights of the formula by maximizing likelihood of given dataset (Doan et al., 2011).

We tried weight learning of metaphor expansion problem, with same MLN program and an instance of the larger dataset shown in figure 7. A specimen of the knowledge base used is given in Figure 6. We learn the weights by running the program for 50 iterations with 20 samples in each iteration. The average weights learnt are shown in Figure 8.

In Table 7, we compare the results using both unique as well as multiple occurrence of Hearst pairs in knowledge base Γ_H for fixed and learnt weights.

5.6 Comparison with the Existing Data-driven Approach

We compare our experimental results with previous data driven approach (Li et al., 2013) for metaphor set expansion. We performed our experiments using Type-I metaphor, where the source and target pairs are explicitly defined. The odds of a pair being a metaphor, probability values are considered. We can argue that MLN does a meaningful job in calculating the odds of pair being a metaphor, as MLN helps soften the constraints, we can get higher probability values, hence more number of new extracted metaphor pairs.

As shown in Table 6 with $(x,y) \equiv (Orange, fuel)$, if we compare the respective probability values for pair (x,y) by existing and MLN methods, it is observed from Table 8, that the proposed MLN based approach gives better probability values by softening the constraints.

6 CONCLUSION

In this work, we have shown a possibility of using MLN for metaphor set expansion. We have used a set of Is-A patterns divided into a metaphor set and a literal set, for identifying potential metaphors.

A single rule base was formulated and converted to the first order logic formulas. Further, based on the importance of rules the weights were assigned appropriately. From results of the experiments carried out it is evident that the proposed approach is able to draw a meaningful inference by assigning appropriate probability values, even in case of multiplicity of Hearst pairs.

It follows from our formulation that the estimated marginal probability values of resultant metaphors due to multiple occurrence of Hearst pairs in the knowledge base will differ from unique occurrence case. It also shows that information completion with MLN provides explainable and meaningful results in comparison with the existing completion methods.

We would like to work on a domain specific dataset and explore the possibility of automated metaphor generation. In addition, we also would like to extend the existing work to incorporate Type-II and Type-III metaphors.

Table 6: Weight learning of formulas with unique occurrence of Hearst pairs vs multiple occurrence of Hearst pairs.

	Unique occurrence	**Multiple occurrence**
Initial weights	$w_1 = 1, w_2 = 1$	$w_1 = 10, w_2 = 1$
Prob. using initial weights	$m(\text{“}fruit, fuel\text{”}) = 0.6600$ $m(\text{“}orange, fuel\text{”}) = 0.7600$ $m(\text{“}company, fuel\text{”}) = 0.7600$	$m(\text{“}fruit, fuel\text{”}) = 0.9990$ $m(\text{“}orange, fuel\text{”}) = 0.7460$ $m(\text{“}company, fuel\text{”}) = 0.8970$
Learned weights	$w_1 = 4.3, w_2 = 13.1$	$w_1 = 2.2, w_2 = 8.1$
Prob. using actual weights	$m(\text{“}fruit, fuel\text{”}) = 0.9400$ $m(\text{“}orange, fuel\text{”}) = 0.9400$ $m(\text{“}company, fuel\text{”}) = 0.9900$	$m(\text{“}fruit, fuel\text{”}) = 0.7900$ $m(\text{“}orange, fuel\text{”}) = 0.8300$ $m(\text{“}company, fuel\text{”}) = 0.8100$

Table 7: Inference results using the data driven approach(Li et al., 2013) vs the proposed approach for a MLN method using small example dataset.

Single occurrence	Existing Method (Li et al., 2013)	Proposed Method
Modern decision parameters	$\delta = 0.5714$	$w_1 = 10, w_2 = 1$
Marginal prob. of (x, y)	$m(\text{“}Orange, fuel\text{”}) = 0.1428$	**m(“Orange,fuel”) = 0.5800**
Prob. of derived metaphor (h_x, y)	$m(\text{“}fruit, fuel\text{”}) = 0.2857$ $m(\text{“}company, fuel\text{”}) = 0.1428$	**m(“fruit,fuel”)= 0.6600** **m(“company,fuel”)= 0.7600**
Multiple occurrence		
Modern decision parameters	$\delta = 0.5849$	$w_1 = 10, w_2 = 1$
Marginal prob. of (x, y)	$m(\text{“}Orange, fuel\text{”}) = 0.2830$	**m(“Orange,fuel”) = 0.7460**
Prob. of derived metaphor (h_x, y)	$m(\text{“}fruit, fuel\text{”}) = 0.5660$ $m(\text{“}company, fuel\text{”}) = 0.2325$	**m(“fruit,fuel”) = 0.9990** **m(“company,fuel”) = 0.8970**

REFERENCES

Cheng, J., Wang, Z., Wen, J.-R., Yan, J., and Chen, Z. (2015). Contextual text understanding in distributional semantic space. In *ACM International Conference on Information and Knowledge Management (CIKM)*. ACM - Association for Computing Machinery.

Doan, A., Niu, F., Ré, C., Shavlik, J., and Zhang, C. (2011). User manual of tuffy 0.3. Technical report.

Domingos, P., Kok, S., Lowd, D., Poon, H., Richardson, M., and Singla, P. (2008). *Markov Logic*, pages 92–117. Springer Berlin Heidelberg, Berlin, Heidelberg.

Fass, D. (1991). met*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1):49–90.

Lakoff, G. and Johnson, M. (1980). *Metaphors we Live by*. University of Chicago Press, Chicago.

Li, H., Zhu, K. Q., and Wang, H. (2013). Data-driven metaphor recognition and explanation. *Transactions of the Association for Computational Linguistics*, 1:379–390.

Niu, F., Ré, C., Doan, A., and Shavlik, J. W. (2011). Tuffy: Scaling up statistical inference in markov logic networks using an RDBMS. *CoRR*, abs/1104.3216.

Richardson, M. and Domingos, P. (2006). Markov logic networks. *Machine learning*, 62(1-2):107–136.

Schulder, M. and Hovy, E. (2014). Metaphor detection through term relevance. In *Proceedings of the Second Workshop on Metaphor in NLP*, pages 18–26, Baltimore, MD. Association for Computational Linguistics.

Shutova, E. (2010). Models of metaphor in NLP. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 688–697, Uppsala, Sweden. Association for Computational Linguistics.

Tsvetkov, Y., Boytsov, L., Gershman, A., Nyberg, E., and Dyer, C. (2014). Metaphor detection with cross-lingual model transfer. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 248–258, Baltimore, Maryland. Association for Computational Linguistics.

Wu, W., Li, H., Wang, H., and Zhu, K. Q. (2012). Probase: A probabilistic taxonomy for text understanding. In *ACM International Conference on Management of Data (SIGMOD)*.