

Lexical Entailment

Rahul MS (IMT2017033)

Prateksha U (IMT2017517)

Ram S (IMT2017521)

Contents

- Introduction
- Literature Review
 - Initial Reading
 - LEAR
 - CLEAR
- Evaluation Criteria
- Solution Ideas

Introduction

Lexical Entailment: Semantic Relationships between words

Types: Hyponym-Hypernym, Correlation, Consequence

Classification: Binary, Graded

Literature Review - Initial Reading

Directional Similarity

Relation Classification with Context Vectors

Relation Classification with Similarity Difference

Multi-Step Linear Transformation (VecMap)

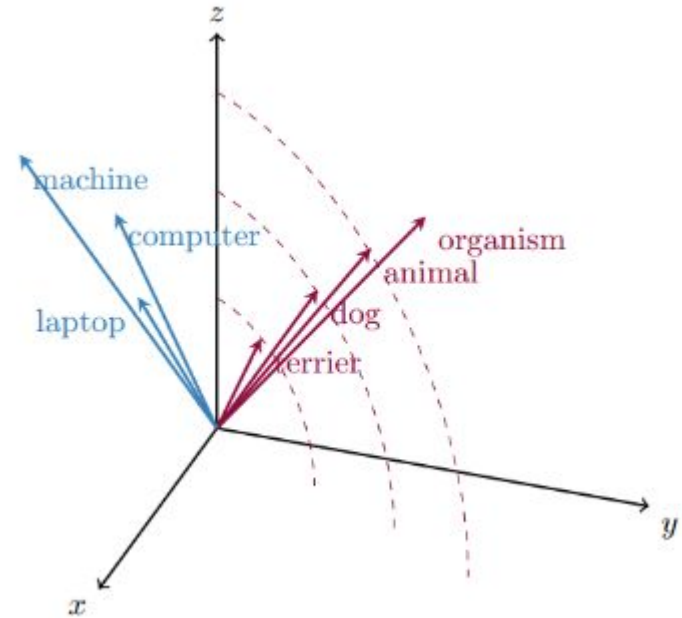
Unsupervised Machine Translation

LEAR

LEAR

Post-processing method that transforms an input word vector space to emphasise the asymmetric relation of lexical entailment (LE).

- Emphasising symmetric similarity of LE pairs through cosine distance
- Imposing the LE ordering using vector norms



Terms

V = Vocabulary

A = Set of Attract pairs

R = Set of repel pairs

T_A = Vectors closest to x_a, x_b form a pair

T_B = Vectors furthest away from x_a, x_b form a pair

L = Lexical constraints

ATTRACT-REPEL framework

$$\begin{aligned} 1 \quad Att(\mathcal{B}_A, T_A) = & \sum_{i=1}^{k_1} \left[\tau \left(\delta_{att} + \cos(\mathbf{x}_l^i, \mathbf{t}_l^i) - \cos(\mathbf{x}_l^i, \mathbf{x}_r^i) \right) \right. \\ & \left. + \tau \left(\delta_{att} + \cos(\mathbf{x}_r^i, \mathbf{t}_r^i) - \cos(\mathbf{x}_l^i, \mathbf{x}_r^i) \right) \right] \quad (1) \end{aligned}$$

$$\begin{aligned} 2 \quad Rep(\mathcal{B}_R, T_R) = & \sum_{i=1}^{k_2} \left[\tau \left(\delta_{rep} + \cos(\mathbf{x}_l^i, \mathbf{x}_r^i) - \cos(\mathbf{x}_l^i, \mathbf{t}_l^i) \right) \right. \\ & \left. + \tau \left(\delta_{rep} + \cos(\mathbf{x}_l^i, \mathbf{x}_r^i) - \cos(\mathbf{x}_r^i, \mathbf{t}_r^i) \right) \right] \quad (2) \end{aligned}$$

$$3 \quad Reg(\mathcal{B}_A, \mathcal{B}_R) = \sum_{\mathbf{x}_i \in V(\mathcal{B}_A \cup \mathcal{B}_R)} \lambda_{reg} \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2$$

$$4 \quad LE_j(\mathcal{B}_L) = \sum_{i=1}^{k_3} D_j(\mathbf{x}_i, \mathbf{y}_i) \quad \text{where, } D_2(\mathbf{x}, \mathbf{y}) = \frac{|\mathbf{x}| - |\mathbf{y}|}{|\mathbf{x}| + |\mathbf{y}|}$$

Cost Function

$$\begin{aligned} C(\mathcal{B}_A, T_A, \mathcal{B}_R, T_R, \mathcal{B}_L, T_L) = & Att(\mathcal{B}_S, T_S) \cdot \\ & + Rep(\mathcal{B}_A, T_A) + Reg(\mathcal{B}_A, \mathcal{B}_R, \mathcal{B}_L) + \\ & + Att(\mathcal{B}_L, T_L) + LE_j(B_L) \end{aligned}$$

LE decoding metric

$$I_{LE}(\mathbf{x}, \mathbf{y}) = dcos(\mathbf{x}, \mathbf{y}) + D_j(\mathbf{x}, \mathbf{y})$$

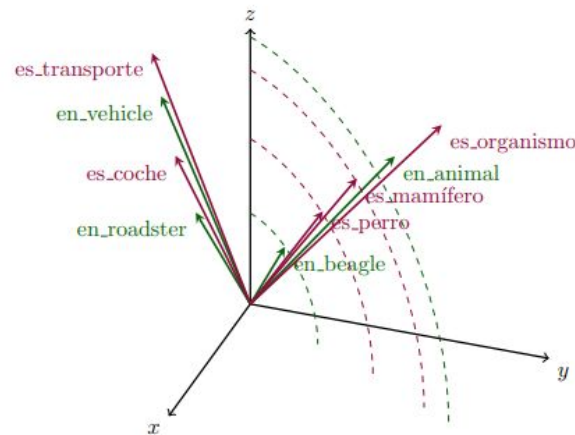
Solving Cross-Lingual Lexical Entailment

- Direct Translation
- Bilingual Word Mappings
 - VecMap
 - CLEAR
- LE Evaluation Criteria

CLEAR - Cross Lingual LEAR

New Cost Function:

$$\begin{aligned} J = & Att(B_S, T_S) + Rep(B_A, T_A) + \\ & Att(B_{Le}, T_{Le}) + LE(B_{Le}) + \\ & Att_D(B_D) + Reg(B_s, B_A, B_{Le}, B_D) \end{aligned}$$



LE Evaluation Criteria

A Distance function that reflects both the cosine distance between the vectors as well as the asymmetric distance between the vectors' norms

$$I_{LE}(x, y) = dcos(x, y) + \frac{||x|| - ||y||}{||x|| + ||y||}$$

Future Work

1. Enhancing/Modifying CLEAR
2. Enhancing the Bilingual word space
3. Experimenting with a hyperbolic word space
4. Multi Cross-lingual Lexical Entailment

Thank You
