

---

# Non-Euclidean Cross-Lingual Word Embeddings

---

## Abstract

## 1 Introduction

This is an alternative approach to the implicit learning of a shared cross-lingual embedding space as yielded in MT. The motivations are two fold a)trade-off model complexity for efficiency as demonstrated by Mikolov et al. [2013] and b)obtain more *composable* word/sentence representations.

## 2 Motivation

The biggest criticism of deep learning is its overt dependence on voluminous data. For human level concept learning and generalization, deep learning models need to become more data efficient. I would like to break down the task of compositionality into two parallel sub-tasks as follows:

### Storing Information meaningfully and efficiently

1. A more efficient data representation.
2. Deducing the manifold of the data.
3. A new RNN cell with internal bias towards composition.
4. The new RNN cell designed with efficient memory allocation considerations.
5. Understander Executor with a running stack.

### Retrieving Information efficiently

1. Attentive Guidance
2. Information Compression. E.g. an entire sentence work of representation recoded into a pronoun 3 sentences later.
3. Dynamic encoder states.
4. Curriculum Learning.

## References

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed Representations of Words and Phrases and their Compositionality. pages 1–9, 2013. ISSN 10495258. doi: 10.1162/jmlr.2003.3.4-5.951. URL <http://arxiv.org/abs/1310.4546>.