



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Alexander Panchenko

---

**FROM UNSUPERVISED INDUCTION OF  
LINGUISTIC STRUCTURES FROM TEXT  
TOWARDS APPLICATIONS IN DEEP  
LEARNING**

# Motivation

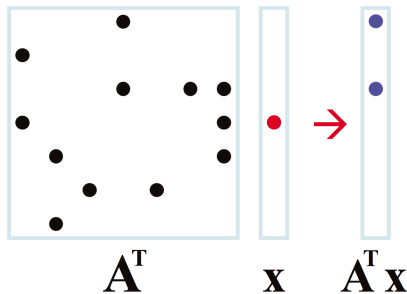
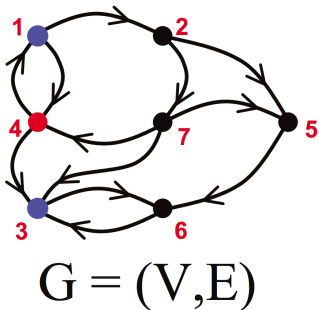
# Deep Learning: everything is a vector



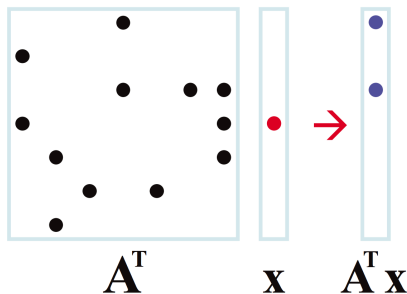
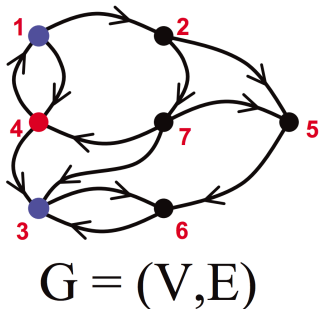
# Linguistic Structures and Graphs

- (Written) language is a **symbolic system**

# Graph Matrix Duality

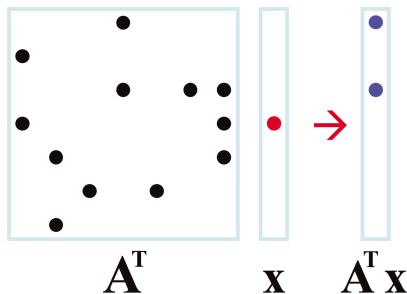
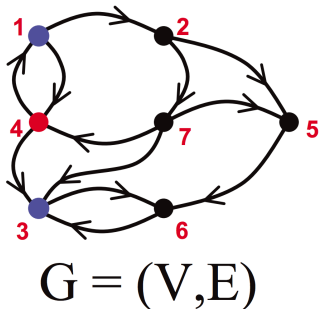


# Graph Matrix Duality



- Adjacency matrix  $A$  is dual with the corresponding graph  $G$ .

# Graph Matrix Duality



- Adjacency matrix  $A$  is dual with the corresponding graph  $G$ .
- Vector matrix multiply  $A^T x$  is dual with breadth-first search.

- 1 Learn the interpretable symbolic structures from text in an unsupervised way, which are **more complex than tokens and lemmas**.



- 1 Learn the interpretable symbolic structures from text in an unsupervised way, which are **more complex than tokens and lemmas**.
- 2 Represent the learned structures in the vector form.

- 1 Learn the interpretable symbolic structures from text in an unsupervised way, which are **more complex than tokens and lemmas**.
- 2 Represent the learned structures in the vector form.
- 3 Use the vector representations instead/in addition to word embedding the deep learning applications.

- 1 Learn the interpretable symbolic structures from text in an unsupervised way, which are **more complex than tokens and lemmas**.
- 2 Represent the learned structures in the vector form.
- 3 Use the vector representations instead/in addition to word embedding the deep learning applications.
- 4 More complex structures could improve performance, but also provide better interpretability of the deep learning models.