Equivalence of Holographic and Complex Embeddings

Katsuhiko Hayashi[a] **Masashi Shimbo**[b]

[a] NTT Communication Science Laboratories [b] Nara Institute of Science and Technology

Background

Two state-of-the-art models for embedding knowledge graphs:

- Holographic embeddings (HolE) [Nickel et al. 2015]
- Complex embeddings (ComplEx) [Trouillon et al. 2016]

Contributions

- Spectral version of HolE reduces time complexity of score computation from $O(n \log n)$ to $O(n)$
- Equivalence of HolE and ComplEx

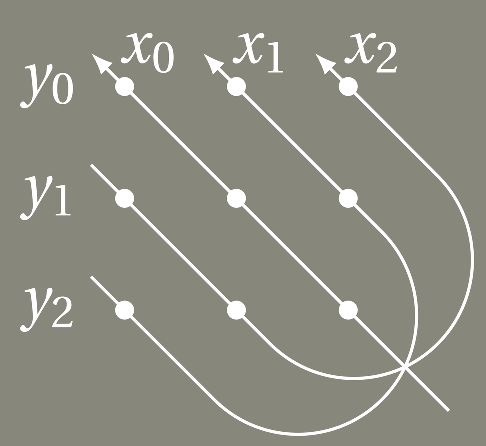
Holographic embeddings (HolE) [Nickel et al. 2015]

Use *circular correlation* between vectors to define a score function

Circular correlation $\mathbf{x} \star \mathbf{y}$

$$[\mathbf{x} \star \mathbf{y}]_j = \sum_{k=0}^{n-1} x_{[(k-j) \bmod n]} \, y_k$$

(for $j = 0, \ldots, n-1$)



HolE score function

$$f_{\text{HolE}}(\mathbf{r}, \mathbf{e}_1, \mathbf{e}_2) = \mathbf{r} \cdot (\mathbf{e}_1 \star \mathbf{e}_2)$$

Note: Convolution is commutative/correlation is not
Correlation thus allows for modeling asymmetric relations

Use discrete Fourier transform for efficient computation

$$\mathbf{e}_1 \star \mathbf{e}_2 = \mathrm{DFT}^{-1}(\overline{\mathrm{DFT}(\mathbf{e}_1)} \odot \mathrm{DFT}(\mathbf{e}_2))$$

where $\odot$: elementwise (Hadamard) product
$\overline{\mathbf{x}}$: complex conjugate of $\mathbf{x}$

➡ Efficient $\because$ DFT, DFT^{-1} computable in $O(n \log n)$ time

Proposal: Spectral computation of HolE

Exploit duality in Fourier transform for further speed-up

Observations:

- Every operation required for computing/training HolE has an equivalent counterpart in Fourier (frequency) space
- Dot product is preserved over two spaces (up to scaling)

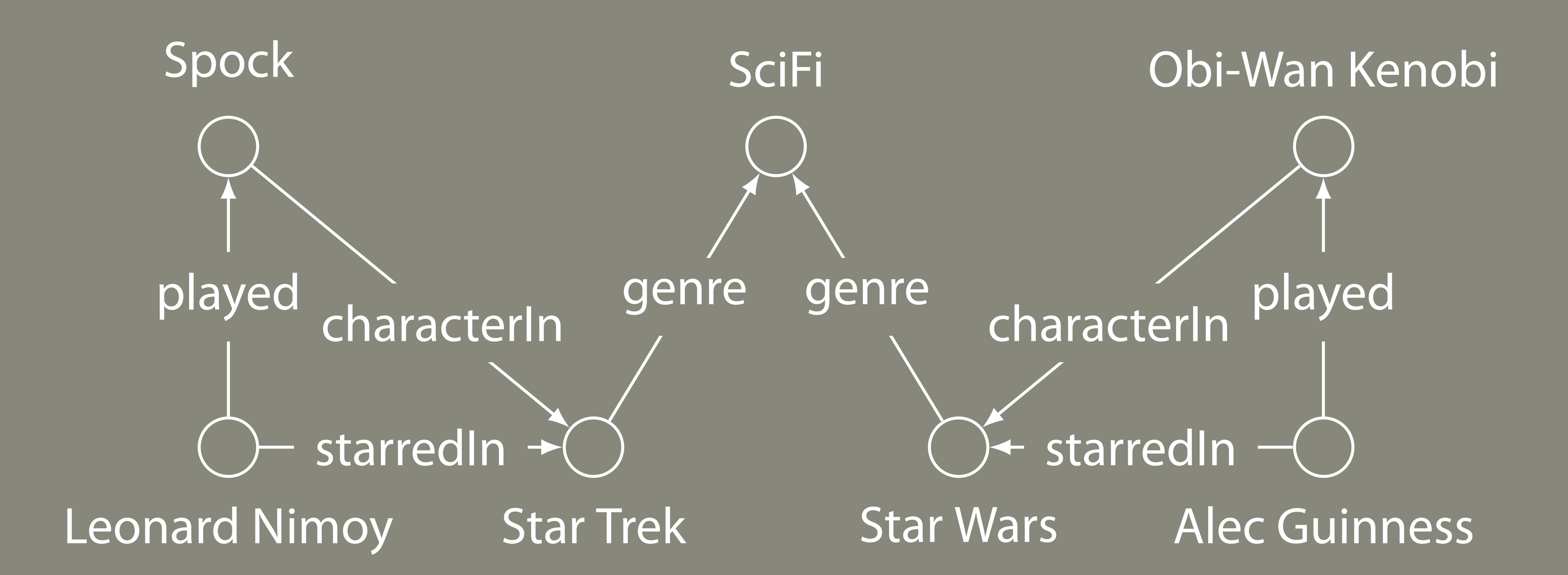
operation in original space			frequency space
scalar multiplication	$\alpha\mathbf{x}$	$\longleftrightarrow$	$\alpha\,\mathrm{DFT}(\mathbf{x})$
summation	$\mathbf{x}+\mathbf{y}$	$\longleftrightarrow$	$\mathrm{DFT}(\mathbf{x}) + \mathrm{DFT}(\mathbf{y})$
correlation	$\mathbf{x} \star \mathbf{y}$	$\longleftrightarrow$	$\overline{\mathrm{DFT}(\mathbf{x})} \odot \mathrm{DFT}(\mathbf{y})$
dot product	$\mathbf{x} \cdot \mathbf{y}$	$=$	$(1/n)\,\mathrm{DFT}(\mathbf{x}) \cdot \mathrm{DFT}(\mathbf{y})$

Idea:

- Compute everything in frequency space
  - ➡ Training begins with random "frequency" vectors that are *conjugate symmetric* (see right panel)
- Fourier/inverse Fourier transforms not necessary
  - ➡ $O(n)$ time for score computation

Knowledge graph embeddings

Knowledge graphs (DBpedia, Freebase, Wordnet, ...)
vertices = entities / edges = relational facts



Example from Nickel et al., *Proc. IEEE* 101(1), 2016

Knowledge graph completion (KGC)

- Graph usually incomplete = many edges missing
- Want to augment missing edges automatically

Graph embedding approach to KGC

From existing vertices and edges, learn

- vector representation $\mathbf{e}$ of each entity e
- representation $\mathbf{r}$ of each relation r
- score function $f(\mathbf{r}, \mathbf{e}_1, \mathbf{e}_2)$ = likelihood of edge $e_1 \xrightarrow{r} e_2$

Many different models of function f exist, e.g.,

TransE [Bordes et al. 2013] $\quad f_{\text{TransE}}(\mathbf{r}, \mathbf{e}_1, \mathbf{e}_2) = -\|\mathbf{e}_1 + \mathbf{r} - \mathbf{e}_2\|$

Complex embeddings (ComplEx) [Trouillon et al. 2016]

Similar to DistMult [Yang et al. 2015] but embedding space = *complex* space (not real space)

$$f_{\text{ComplEx}}(\mathbf{r}, \mathbf{e}_1, \mathbf{e}_2) = \mathrm{Re}\left(\sum_{j=0}^{n-1} [\mathbf{r}]_j [\mathbf{e}_1]_j \overline{[\mathbf{e}_2]_j}\right)$$

where $\mathbf{r}, \mathbf{e}_1, \mathbf{e}_2$: complex vectors ($\in \mathbb{C}^n$, *not* $\mathbb{R}^n$)
$\mathrm{Re}(x)$: real part of complex value x

Finding: HolE and ComplEx are equivalent

Spectral HolE $\subset$ ComplEx

- Every spectral HolE vector is a ComplEx vector satisfying *conjugate symmetry*:

$$\mathbf{x} = \begin{cases} [x_0 \; x_1 \; \cdots \; x_{(n-1)/2} \; \overline{x_{(n-1)/2}} \; \cdots \; \overline{x_1}] & \text{if } n \text{ odd} \\ [x_0 \; x_1 \; \cdots \; x_{n/2-1} \; x_{n/2} \; \overline{x_{n/2-1}} \; \cdots \; \overline{x_1}] & \text{if } n \text{ even} \end{cases}$$

with $x_0, x_{n/2} \in \mathbb{R}$; all other $x_j \in \mathbb{C}$

- Score functions for spectral HolE and ComplEx are equivalent

ComplEx $\subset$ Spectral HolE

An n-dim. ComplEx vector can be transformed into a $(2n+1)$-dim. spectral HolE vector

$$\underbrace{[x_0 \cdots x_{n-1}]}_{\text{ComplEx}} \mapsto \underbrace{[0 \; x_0 \cdots x_{n-1} \; \overline{x_{n-1}} \cdots \overline{x_0}]}_{\text{Spectral HolE (conjugate symmetric)}}$$

➡ **HolE $\equiv$ Spectral HolE $\equiv$ ComplEx**