Analogies Explained

Towards Understanding Word Embeddings

Carl Allen, Tim Hospedales

June 13 2019

School of Informatics, University of Edinburgh

The Problem: linking semantics to geometry

from:

"man is to king as woman is to queen"

The Problem: linking semantics to geometry

from:

"man is to king as woman is to queen"

explain:

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

The Problem: linking semantics to geometry

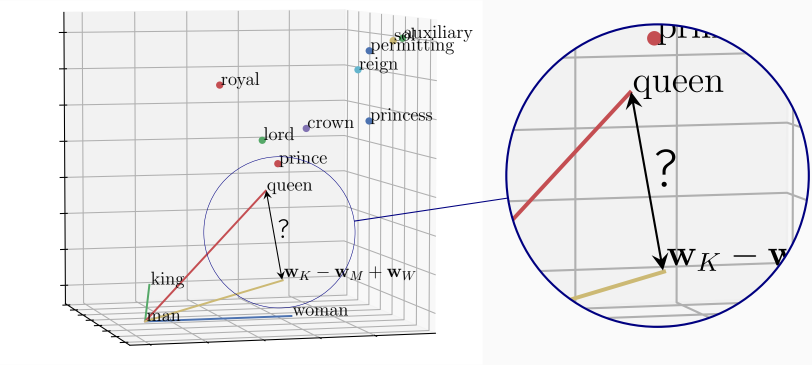
from:

“man is to king as woman is to queen”

explain:

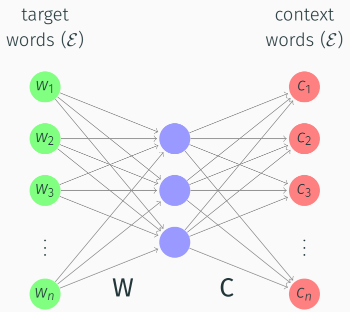
$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

or rather:



Word2Vec: SkipGram with Negative Sampling

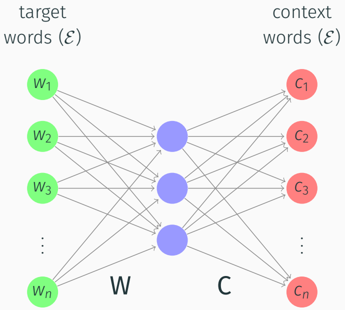
Mikolov et al. (2013a,b)



- $p(c_j|w_i)$ by **softmax** expensive

Word2Vec: SkipGram with Negative Sampling

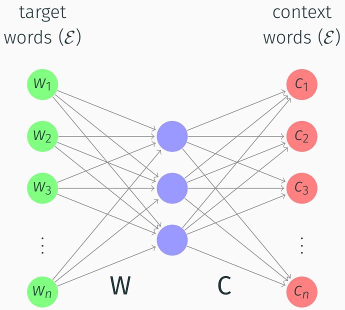
Mikolov et al. (2013a,b)



- $p(c_j|w_i)$ by **softmax** expensive
- use **sigmoid** with negative sampling (k)

Word2Vec: SkipGram with Negative Sampling

Mikolov et al. (2013a,b)

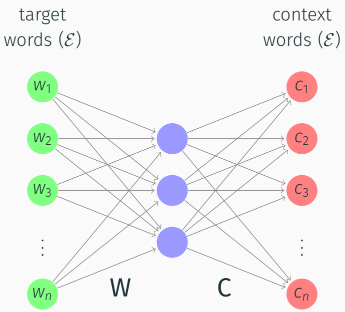


- $p(c_j|w_i)$ by **softmax** expensive
- use **sigmoid** with negative sampling (k)
- Levy and Goldberg (2014)

$$\mathbf{w}_i^\top \mathbf{c}_j \approx \log \frac{p(w_i, c_j)}{p(w_i)p(c_j)} - \log k$$
$$= \mathrm{PMI}(w_i, c_j) - \log k$$

Word2Vec: SkipGram with Negative Sampling

Mikolov et al. (2013a,b)



- $p(c_j|w_i)$ by **softmax** expensive
- use **sigmoid** with negative sampling (k)
- Levy and Goldberg (2014)

$$\mathbf{w}_i^\top \mathbf{c}_j \approx \log \frac{p(w_i,c_j)}{p(w_i)p(c_j)} - \log k$$
$$= \mathrm{PMI}(w_i, c_j) - \log k$$

$$\boxed{\mathbf{W}^\top \mathbf{C} \;\approx\; \mathbf{PMI} - \log k}$$

Routemap

"man is to king as woman is to queen"

Routemap

"man is to king as woman is to queen"

$\Updownarrow$

man *transforms to* king as woman *transforms to* queen

Routemap

"man is to king as woman is to queen"

$\Updownarrow$

man *transforms to* king as woman *transforms to* queen

$\Updownarrow$

{woman, king} *paraphrases* {man, queen}

Routemap

"man is to king as woman is to queen"

$$\Updownarrow$$

man *transforms to* king as woman *transforms to* queen

$$\Updownarrow$$

{woman, king} *paraphrases* {man, queen}

$$\Downarrow$$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

Routemap

"man is to king as woman is to queen"

$$\Updownarrow$$

man *transforms to* king as woman *transforms to* queen

$$\Updownarrow$$

{woman, king} *paraphrases* {man, queen}

$$\Downarrow$$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$$\Downarrow$$

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

Routemap

"man is to king as woman is to queen"

$$\Updownarrow$$

man *transforms to* king as woman *transforms to* queen

$$\Updownarrow$$

{woman, king} *paraphrases* {man, queen}

$$\Downarrow$$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$$\Downarrow$$

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

semantic

geometric

Routemap

"man is to king as woman is to queen"

$\Updownarrow$

man *transforms to* king as woman *transforms to* queen

$\Updownarrow$

{woman, king} *paraphrases* {man, queen}

$\Downarrow$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$\Downarrow$

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

semantic

geometric

Routemap

"man is to king as woman is to queen"

$\Updownarrow$

man *transforms to* king as woman *transforms to* queen

$\Updownarrow$

{woman, king} *paraphrases* {man, queen}

$\Downarrow$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$\Downarrow \quad \mathrm{PMI}_i \approx \mathbf{w}_i^\top \mathbf{C}$

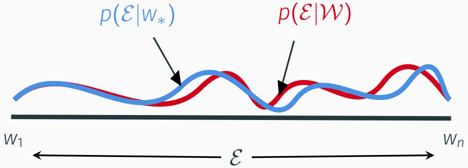
$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

semantic

geometric

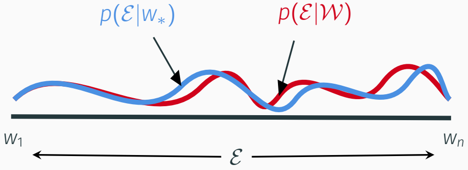
Paraphrase† of $\mathcal{W}$ by w_*

Intuition: word $w_* \in \mathcal{E}$ **paraphrases** word set $\mathcal{W} = \{w_1, \ldots, w_m\} \subseteq \mathcal{E}$, if w_* and $\mathcal{W}$ are *semantically interchangeable.*



Paraphrase† of $\mathcal{W}$ by w_*

Intuition: word $w_* \in \mathcal{E}$ **paraphrases** word set $\mathcal{W} = \{w_1, \ldots, w_m\} \subseteq \mathcal{E}$, if w_* and $\mathcal{W}$ are *semantically interchangeable.*



Definition (D1): $w_* \in \mathcal{E}$ paraphrases $\mathcal{W} \subseteq \mathcal{E}$, $|\mathcal{W}| < l$, if paraphrase error $\boldsymbol{\rho}^{\mathcal{W}, w_*} \in \mathbb{R}^n$ is (element-wise) small:

$$\boldsymbol{\rho}_j^{\mathcal{W}, w_*} = \log \frac{p(c_j|w_*)}{p(c_j|\mathcal{W})}, \quad c_j \in \mathcal{E}$$

†Inspired by Gittens et al. (2017)

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_*$?

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_*$?

$\mathrm{PMI}(w_*, c_j) - \big(\mathrm{PMI}(w_1,c_j) + \mathrm{PMI}(w_2,c_j)\big)$

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_*$?

$$\begin{aligned}
&\mathrm{PMI}(w_*, c_j) - \big(\mathrm{PMI}(w_1, c_j) + \mathrm{PMI}(w_2, c_j)\big) \\
&\qquad = \; \log \frac{p(w_*|c_j)}{p(w_*)} - \log \frac{p(w_1|c_j)p(w_2|c_j)}{p(w_1)p(w_2)}
\end{aligned}$$

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_* \;?$

$$\mathrm{PMI}(w_*, c_j) - \big(\mathrm{PMI}(w_1,c_j) + \mathrm{PMI}(w_2,c_j)\big)$$

$$= \;\log \frac{p(w_*|c_j)}{p(w_*)} - \log \frac{p(w_1|c_j)p(w_2|c_j)}{p(w_1)p(w_2)} + \log \frac{p(\mathcal{W}|c_j)}{p(\mathcal{W}|c_j)} + \log \frac{p(\mathcal{W})}{p(\mathcal{W})}$$

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_*\ ?$

$$\mathrm{PMI}(w_*, c_j) - \big(\mathrm{PMI}(w_1,c_j) + \mathrm{PMI}(w_2,c_j)\big)$$

$$= \log \frac{p(w_*|c_j)}{p(w_*)} - \log \frac{p(w_1|c_j)p(w_2|c_j)}{p(w_1)p(w_2)} + \log \frac{p(\mathcal{W}|c_j)}{p(\mathcal{W}|c_j)} + \log \frac{p(\mathcal{W})}{p(\mathcal{W})}$$

$$= \underbrace{\log \frac{p(c_j|w_*)}{p(c_j|\mathcal{W})}}_{\boldsymbol{\rho}_j^{\mathcal{W},w_*}} + \underbrace{\log \frac{p(\mathcal{W}|c_j)}{p(w_1|c_j)p(w_2|c_j)}}_{\boldsymbol{\sigma}_j^{\mathcal{W}}} - \underbrace{\log \frac{p(\mathcal{W})}{p(w_1)p(w_2)}}_{\tau^{\mathcal{W}}}$$

paraphrase error

conditional independence error

independence error

Summing PMI vectors of a paraphrase

$\mathrm{PMI}_1 + \mathrm{PMI}_2 \;\approx\; \mathrm{PMI}_*$?

$$\mathrm{PMI}(w_*, c_j) - \big(\mathrm{PMI}(w_1,c_j) + \mathrm{PMI}(w_2,c_j)\big)$$

$$= \log \frac{p(w_*|c_j)}{p(w_*)} - \log \frac{p(w_1|c_j)p(w_2|c_j)}{p(w_1)p(w_2)} + \log \frac{p(\mathcal{W}|c_j)}{p(\mathcal{W}|c_j)} + \log \frac{p(\mathcal{W})}{p(\mathcal{W})}$$

$$= \underbrace{\log \frac{p(c_j|w_*)}{p(c_j|\mathcal{W})}}_{\boldsymbol{\rho}_j^{\mathcal{W},w_*}} + \underbrace{\log \frac{p(\mathcal{W}|c_j)}{p(w_1|c_j)p(w_2|c_j)}}_{\boldsymbol{\sigma}_j^{\mathcal{W}}} - \underbrace{\log \frac{p(\mathcal{W})}{p(w_1)p(w_2)}}_{\tau^{\mathcal{W}}}$$

paraphrase error

conditional independence error

independence error

Lemma 1: For any word $w_* \in \mathcal{E}$ and word set $\mathcal{W} \subseteq \mathcal{E}$, $|\mathcal{W}| < l$:

$$\mathrm{PMI}_* \;=\; \sum_{w_i \in \mathcal{W}} \mathrm{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},w_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \tau^{\mathcal{W}}\mathbf{1}$$

Generalised Paraphrase (of $\mathcal{W}$ by $\mathcal{W}_*$)

Lemma 1: For any word $w_* \in \mathcal{E}$ and word set $\mathcal{W} \subseteq \mathcal{E}$, $|\mathcal{W}| < l$:

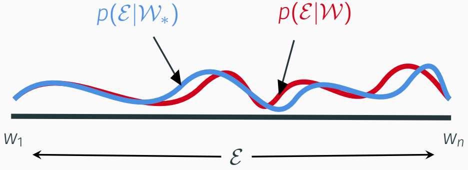
$$\text{PMI}_* = \sum_{w_i \in \mathcal{W}} \text{PMI}_i + \boldsymbol{\rho}^{\mathcal{W}, w_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \tau^{\mathcal{W}} \mathbf{1}$$

Generalised Paraphrase (of $\mathcal{W}$ by $\mathcal{W}_*$)

Lemma 1: For any word $w_* \in \mathcal{E}$ and word set $\mathcal{W} \subseteq \mathcal{E}$, $|\mathcal{W}| < l$:

$$\mathbf{PMI}_* = \sum_{w_i \in \mathcal{W}} \mathbf{PMI}_i + \boldsymbol{\rho}^{\mathcal{W}, w_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \tau^{\mathcal{W}} \mathbf{1}$$

Replace word w_* with word set $\mathcal{W}_* \subseteq \mathcal{E}$:

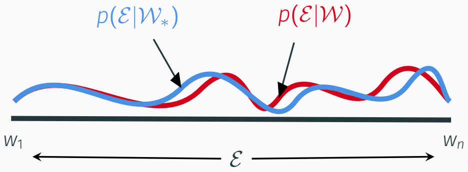


Generalised Paraphrase (of $\mathcal{W}$ by $\mathcal{W}_*$)

Lemma 1: For any word $w_* \in \mathcal{E}$ and word set $\mathcal{W} \subseteq \mathcal{E}$, $|\mathcal{W}| < l$:

$$\mathsf{PMI}_* = \sum_{w_i \in \mathcal{W}} \mathsf{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},w_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \tau^{\mathcal{W}}\mathbf{1}$$

Replace word w_* with word set $\mathcal{W}_* \subseteq \mathcal{E}$:



Lemma 2: For any word sets $\mathcal{W}, \mathcal{W}_* \subseteq \mathcal{E}$, $|\mathcal{W}|, |\mathcal{W}_*| < l$:

$$\sum_{w_i \in \mathcal{W}_*} \mathsf{PMI}_i = \sum_{w_i \in \mathcal{W}} \mathsf{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},\mathcal{W}_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \boldsymbol{\sigma}^{\mathcal{W}_*} - (\tau^{\mathcal{W}} - \tau^{\mathcal{W}_*})\mathbf{1}$$

Paraphrase: the link from semantics to geometry

Lemma 2: For any word sets $\mathcal{W}$, $\mathcal{W}_* \subseteq \mathcal{E}$, $|\mathcal{W}|$, $|\mathcal{W}_*| < l$:

$$\sum_{w_i \in \mathcal{W}_*} \mathsf{PMI}_i \;=\; \sum_{w_i \in \mathcal{W}} \mathsf{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},\mathcal{W}_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \boldsymbol{\sigma}^{\mathcal{W}_*} - (\tau^{\mathcal{W}} - \tau^{\mathcal{W}_*})\mathbf{1}$$

Paraphrase: the link from semantics to geometry

Lemma 2: For any word sets $\mathcal{W}$, $\mathcal{W}_* \subseteq \mathcal{E}$, $|\mathcal{W}|$, $|\mathcal{W}_*| < l$:

$$\sum_{w_i \in \mathcal{W}_*} \mathsf{PMI}_i \;=\; \sum_{w_i \in \mathcal{W}} \mathsf{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},\mathcal{W}_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \boldsymbol{\sigma}^{\mathcal{W}_*} - (\tau^{\mathcal{W}} - \tau^{\mathcal{W}_*})\mathbf{1}$$

So, if

$$\mathcal{W} = \{\textit{woman}, \textit{king}\} \quad \text{paraphrases} \quad \mathcal{W}_* = \{\textit{man}, \textit{queen}\},$$

Paraphrase: the link from semantics to geometry

Lemma 2: For any word sets $\mathcal{W}, \mathcal{W}_* \subseteq \mathcal{E}$, $|\mathcal{W}|, |\mathcal{W}_*| < l$:

$$\sum_{w_i \in \mathcal{W}_*} \mathrm{PMI}_i \;=\; \sum_{w_i \in \mathcal{W}} \mathrm{PMI}_i + \boldsymbol{\rho}^{\mathcal{W},\mathcal{W}_*} + \boldsymbol{\sigma}^{\mathcal{W}} - \boldsymbol{\sigma}^{\mathcal{W}_*} - (\tau^{\mathcal{W}} - \tau^{\mathcal{W}_*})\mathbf{1}$$

So, if

$$\mathcal{W} = \{\textit{woman}, \textit{king}\} \quad \text{paraphrases} \quad \mathcal{W}_* = \{\textit{man}, \textit{queen}\},$$

then:

$$\mathrm{PMI}_{\textit{queen}} \;\approx\; \mathrm{PMI}_{\textit{king}} - \mathrm{PMI}_{\textit{man}} + \mathrm{PMI}_{\textit{woman}} + \underbrace{\boldsymbol{\sigma}^{\mathcal{W}} - \boldsymbol{\sigma}^{\mathcal{W}_*} - (\tau^{\mathcal{W}} - \tau^{\mathcal{W}_*})\mathbf{1}}_{\text{net dependence error}}$$

Routemap

"man is to king as woman is to queen"

$\Updownarrow$

man *transforms to* king as woman *transforms to* queen

$\Updownarrow$

{woman, king} *paraphrases* {man, queen}

$\Downarrow$ dependence error

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$\Downarrow$ $\mathrm{PMI}_i \approx \mathbf{w}_i^\top \mathbf{C}$

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

semantic

geometric

Word Transformation: a change of perspective

A paraphrase w_* of $\mathcal{W}$ can be thought of as a **word transformation** from some $w \in \mathcal{W}$ to w_* by *adding* $\mathcal{W}^+ = \{w_i \in \mathcal{W},\ w_i \neq w\}$, e.g.

$$\{man, royal\} \approx_p king \quad \Longrightarrow \quad man \xrightarrow{+royal} king$$

Word Transformation: a change of perspective

A paraphrase w_* of $\mathcal{W}$ can be thought of as a **word transformation** from some $w{\in}\mathcal{W}$ to w_* by *adding* $\mathcal{W}^+{=}\{w_i \in \mathcal{W},\ w_i{\neq}w\}$, e.g.

$$\{man, royal\} \approx_p king \quad \Longrightarrow \quad man \xrightarrow{+royal} king$$

Added words ***contextualise*** w, such that the induced distribution better aligns with that of w_*.

Word Transformation: a change of perspective

A paraphrase w_* of $\mathcal{W}$ can be thought of as a **word transformation** from some $w \in \mathcal{W}$ to w_* by *adding* $\mathcal{W}^+ = \{w_i \in \mathcal{W},\ w_i \neq w\}$, e.g.

$$\{man, royal\} \approx_p king \quad \Longrightarrow \quad man \xrightarrow{+royal} king$$

Added words ***contextualise*** w, such that the induced distribution better aligns with that of w_*.

Paraphrase $\mathcal{W}$ by $\mathcal{W}_*$ can be thought of as a word transformation from some $w \in \mathcal{W}$ to some $w_* \in \mathcal{W}_*$ by adding to both ...

$$w \xrightarrow{+\mathcal{W}^+} \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \xleftarrow{+\mathcal{W}^-} w_*$$

Word Transformation: a change of perspective

A paraphrase w_* of $\mathcal{W}$ can be thought of as a **word transformation** from some $w \in \mathcal{W}$ to w_* by *adding* $\mathcal{W}^+ = \{w_i \in \mathcal{W},\ w_i \neq w\}$, e.g.

$$\{man, royal\} \approx_p king \quad \Longrightarrow \quad man \xrightarrow{+royal} king$$

Added words ***contextualise*** w, such that the induced distribution better aligns with that of w_*.

Paraphrase $\mathcal{W}$ by $\mathcal{W}_*$ can be thought of as a word transformation from some $w \in \mathcal{W}$ to some $w_* \in \mathcal{W}_*$ by adding to both …

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* & \\
w \arrow[ur, "+\mathcal{W}^+"] & & w_* \arrow[ul, "+\mathcal{W}^-"']
\end{tikzcd}$$

… or adding to one side and *subtracting* from the other:

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
w \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & w_*
\end{tikzcd}$$

Word Transformation: a change of perspective (cont.)

$$
\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
w \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & w_*
\end{tikzcd}
$$

A generalised paraphrase ***is*** a word transformation from $w \in \mathcal{W}$ to $w_* \in \mathcal{W}_*$, where:

- added words *narrow* context
- subtracted words *broaden* context

Word Transformation: a change of perspective (cont.)

$$\begin{tikzcd}
& \mathcal{W} \approx_{\mathsf{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
w \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & w_*
\end{tikzcd}$$

A generalised paraphrase ***is*** a word transformation from $w \in \mathcal{W}$ to $w_* \in \mathcal{W}_*$, where:

- added words *narrow* context
- subtracted words *broaden* context

Providing a “richer dictionary” to explain the difference between w and w_*, or rather, how **“w is to $\mathbf{w_*}$”**.

Word Transformation: a change of perspective (cont.)

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
w \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & w_*
\end{tikzcd}$$

A generalised paraphrase ***is*** a word transformation from $w \in \mathcal{W}$ to $w_* \in \mathcal{W}_*$, where:

- added words *narrow* context
- subtracted words *broaden* context

Providing a "richer dictionary" to explain the difference between w and w_*, or rather, how **"w is to $\mathbf{w_*}$"**.

Definition (D4): We say *"w_a is to w_{a*} as w_b is to w_{b*}" iff* there exist $\mathcal{W}^+, \mathcal{W}^- \subseteq \mathcal{E}$ that simultaneously transform w_a to w_{a*} and w_b to w_{b*}.

Word Transformation: a change of perspective (cont.)

That is, we say:

"man is to king as woman is to queen"

iff there exist $\mathcal{W}^+, \mathcal{W}^- \subseteq \mathcal{E}$ that simultaneously transform man to king and woman to queen.

Word Transformation: a change of perspective (cont.)

That is, we say:

"man is to king as woman is to queen"

iff there exist $\mathcal{W}^+, \mathcal{W}^- \subseteq \mathcal{E}$ that simultaneously transform man to king and woman to queen.

That is:

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
\text{man} \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{king}
\end{tikzcd}$$

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
\text{woman} \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{queen}
\end{tikzcd}$$

Word Transformation: a change of perspective (cont.)

That is, we say:

"man is to king as woman is to queen"

iff there exist $\mathcal{W}^+, \mathcal{W}^- \subseteq \mathcal{E}$ that simultaneously transform man to king and woman to queen.

That is:

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
\text{man} \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{king}
\end{tikzcd}$$

$$\begin{tikzcd}
 & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\mathcal{W}^-"] & \\
\text{woman} \arrow[ur, "+\mathcal{W}^+"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{queen}
\end{tikzcd}$$

Let $\mathcal{W}^+ = \{\mathit{king}\}, \mathcal{W}^- = \{\mathit{man}\}$.

Word Transformation: a change of perspective (cont.)

That is, we say:

"man is to king as woman is to queen"

iff there exist $\mathcal{W}^+, \mathcal{W}^- \subseteq \mathcal{E}$ that simultaneously transform man to king and woman to queen.

That is:

$$\begin{tikzcd} & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\textit{man}"] & \\ \text{man} \arrow[ur, "+\textit{king}"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{king} \end{tikzcd}$$

$$\begin{tikzcd} & \mathcal{W} \approx_{\mathrm{P}} \mathcal{W}_* \arrow[dr, "-\textit{man}"] & \\ \text{woman} \arrow[ur, "+\textit{king}"] \arrow[rr, dashed, "\textit{word transformation}"'] & & \text{queen} \end{tikzcd}$$

Let $\mathcal{W}^+ = \{\textit{king}\}, \mathcal{W}^- = \{\textit{man}\}$.

Routemap

"man is to king as woman is to queen"

$$\Updownarrow$$

man *transforms to* king as woman *transforms to* queen

$$\Updownarrow$$

{woman, king} *paraphrases* {man, queen}

$$\Downarrow \quad \text{dependence error}$$

$$\mathrm{PMI}_{king} - \mathrm{PMI}_{man} + \mathrm{PMI}_{woman} \approx \mathrm{PMI}_{queen}$$

$$\Downarrow \quad \mathrm{PMI}_i \approx \mathbf{w}_i^\top \mathbf{C}$$

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \approx \mathbf{w}_{queen}$$

semantic

geometric

The Solution: linking semantics to geometry

"man is to king as woman is to queen"

The Solution: linking semantics to geometry

"man is to king as woman is to queen"

implies:

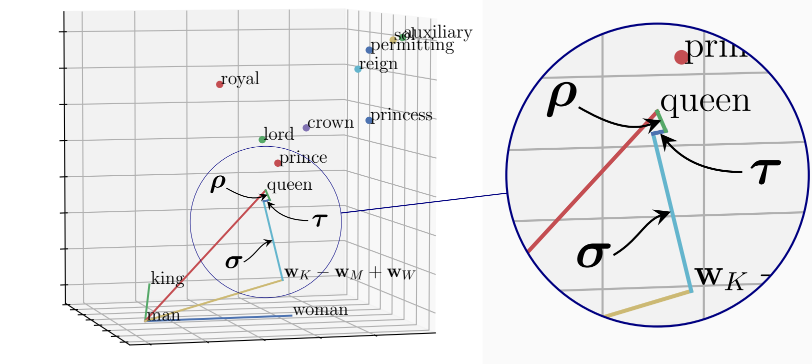
$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \overset{\boldsymbol{\rho},\boldsymbol{\sigma},\boldsymbol{\tau}}{\approx} \mathbf{w}_{queen}$$

The Solution: linking semantics to geometry

"man is to king as woman is to queen"

implies:

$$\mathbf{w}_{king} - \mathbf{w}_{man} + \mathbf{w}_{woman} \overset{\rho,\sigma,\tau}{\approx} \mathbf{w}_{queen}$$



References

Alex Gittens, Dimitris Achlioptas, and Michael W Mahoney. Skip-gram-zipf+ uniform= vector additivity. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 69–76, 2017.

Omer Levy and Yoav Goldberg. Neural word embedding as implicit matrix factorization. In *Advances in neural information processing systems*, 2014.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013a.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 2013b.