GloVe: Global Vectors for Word Representation[7]

How can we capture words related to **ice**, but not to **steam**?

Prob. or Ratio	w_k = **solid**	w_k = **gas**	w_k = **water**	w_k = **fashion**
$P(w_k \mid \text{ice})$	1.9×10^{-4}	6.6×10^{-5}	3.0×10^{-3}	1.7×10^{-5}
$P(w_k \mid \text{steam})$	2.2×10^{-5}	7.8×10^{-4}	2.2×10^{-3}	1.8×10^{-5}
$\frac{P(w_k \mid \text{ice})}{P(w_k \mid \text{steam})}$	8.9	8.5×10^{-2}	1.36	0.96

Probability ratios are most informative:

- **solid** is related to **ice** but not **steam**
- **gas** is related to **steam** but not **ice**
- **water** and **fashion** do not discriminate between **ice** or **steam** (ratios close to 1)

[7] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. "GloVe: Global Vectors for Word Representation". In: *EMNLP*. 2014, pp. 1532–1543.