Conditional Entropy for Mining Syntagmatic Relations

- For each word W1
  - For every other word W2, compute conditional entropy $H(X_{W1}|X_{W2})$
  - Sort all the candidate words in ascending order of $H(X_{W1}|X_{W2})$
  - Take the top-ranked candidate words as words that have potential syntagmatic relations with W1
  - Need to use a threshold for each W1
- However, while $H(X_{W1}|X_{W2})$ and $H(X_{W1}|X_{W3})$ are comparable, $H(X_{W1}|X_{W2})$ and $H(X_{W3}|X_{W2})$ aren't!

How can we mine the strongest K syntagmatic relations from a collection?