

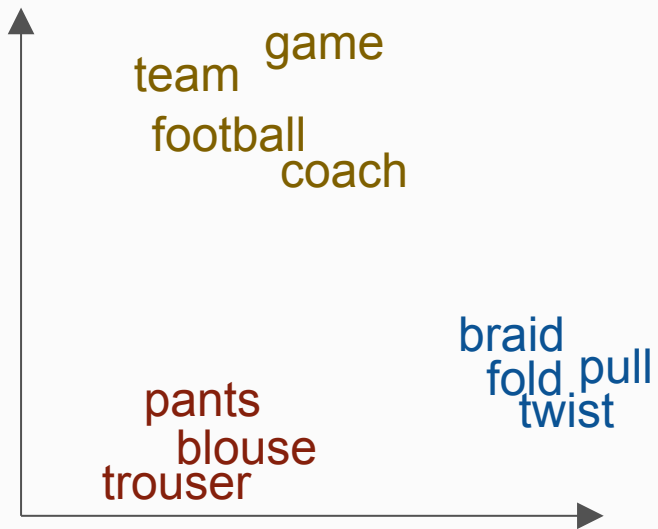
Linear Algebraic Structure of Word Senses, with Applications to Polysemy

Sanjeev Arora, Yuanzhi Li, Yingyu Liang, Tengyu Ma, Andrej Risteski

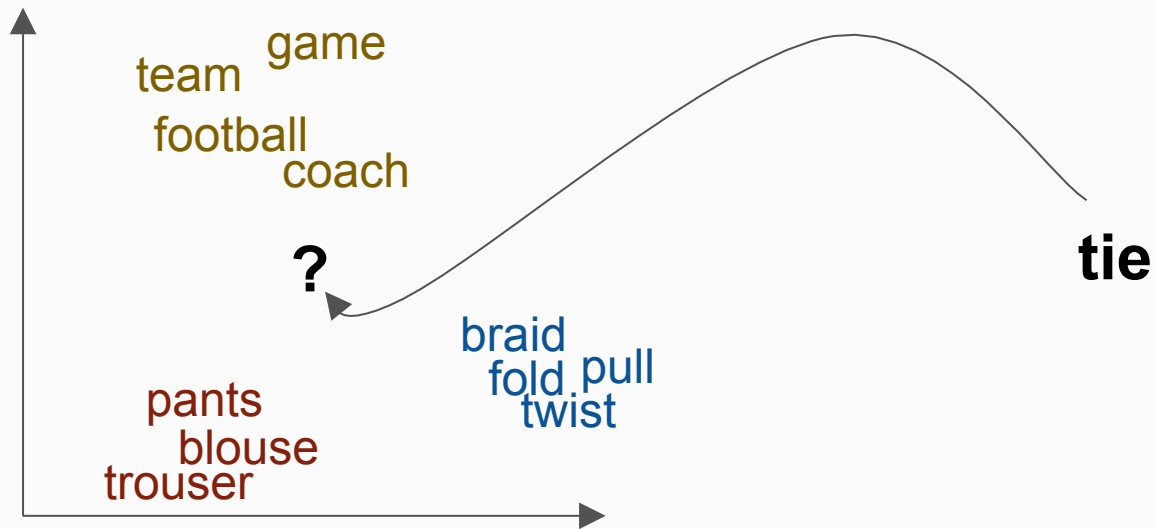
Presented by: **Arun Chaganty**



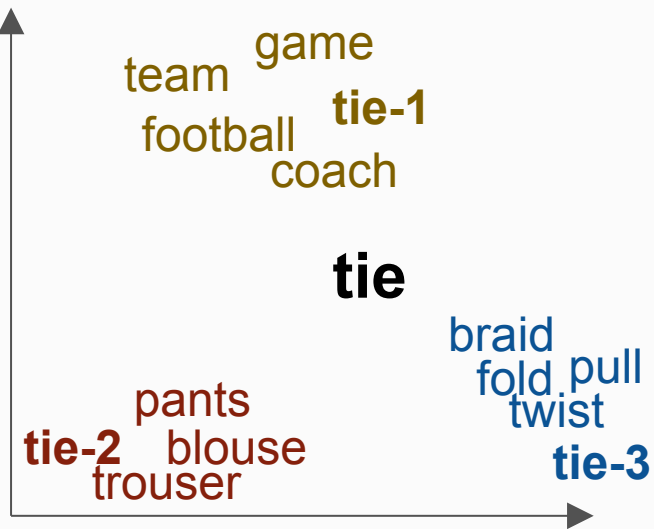
Word vectors encode similarity.



What about polysemy?



1. Polysemous vectors are superpositioned.



2. Senses can be recovered by sparse coding

$$v = \sum_{i=0}^D \alpha_i A_i + \eta$$

Word vector **Context vectors (~2000)**
Selectors (< 5) **noise**

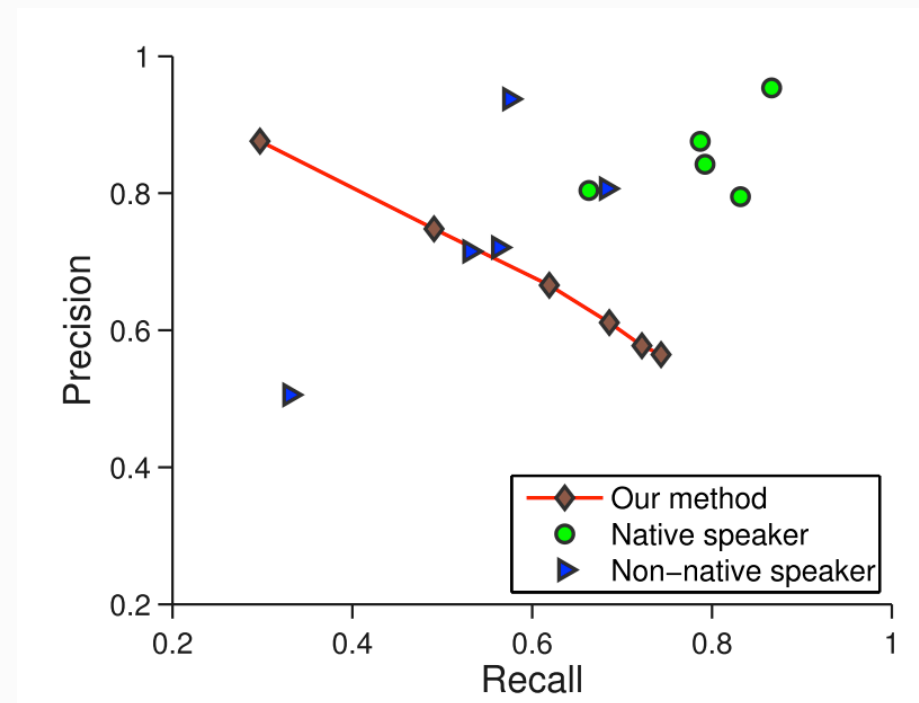
2. Senses can be recovered by sparse coding

tie				
trousers	season	scoreline	wires	operatic
blouse	teams	goalless	cables	soprano
waistcoat	winning	equaliser	wiring	mezzo
skirt	league	clinching	electrical	contralto
sleeved	finished	scoreless	wire	baritone
pants	championship	replay	cable	coloratura

3. Senses recovered are non-native English level

tie

1. Trousers, blouse, pants
2. Season, teams, winning
3. Computer, mouse, keyboard
4. Bulb, light, flash



Summary

Word vectors can capture polysemy!

Word vectors are linear superposition of each sense vector.

Sense/context vectors can be recovered by sparse coding.

The senses recovered are about as good as a non-native English speakers!

Thanks!