Graph Based Word Sense Disambiguation

Ma Yukun

Nanyang Technological University

mayu001@e.netu.edu.sg

September 12, 2014

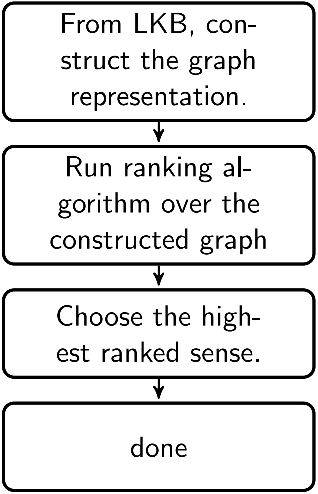
Overview

1. Graph-based word sense disambiguation
2. Traditional PageRank
3. Persoanlizing PageRank

Objective of WSD

- The task of Word Sense Disambiguation (WSD) is to automatically choose the intended sense of a given target word w in context.
- Graph-based WSD exploited the interrelations between senses underlying the graph representation of a particular Lexical Knowledge Base(LKB).

Graph-based WSD



Lexical Knowledge Based (LKB)

LKB is basicaly an undirected graph $G = (V, E)$

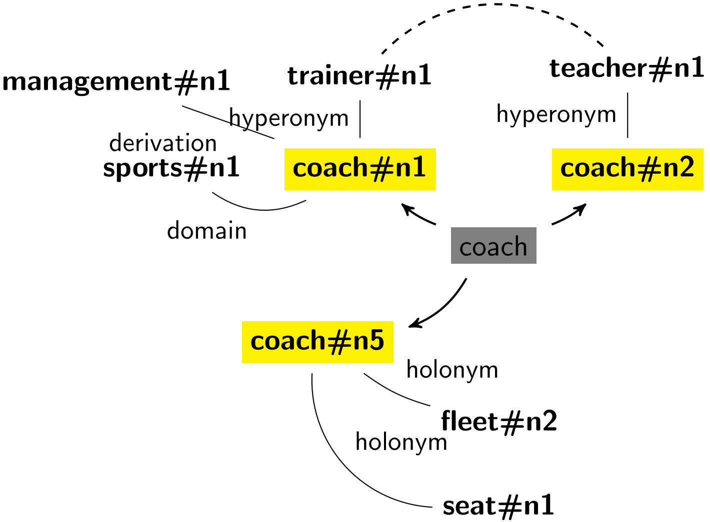
- each node in $V = v_i$ represents a concept.
- each undirected edge $e_{i,j}$ in E represents a relation between concept v_i and v_j.

LKB-example

- coach#n1 : coach, manager, handler ((sports) someone in charge of training an athlete or a team)
- coach#n2 : coach, private instructor, tutor (a person who gives private instruction (as in singing, acting, etc.))
- coach#n5 : bus, autobus, coach, charabanc, double-decker, jitney, motorbus, motorcoach, omnibus, passenger vehicle (a vehicle carrying many passengers; used for public transport)

LKB-example

"Our fleet comprises coaches from 35 to 58 seats."



Lexical Knowledge Based (LKB)

- MCR16+Xwn : WordNet 1.6 synsets and relations; WordNet 2.0 relations; eXtended WordNet relations.
- WNet17 + Xwn : WordNet 1.7 synsets and relations plus the exTended WordNet relations.
- WNet30+gloss : WordNet 3.0 synsets and relations.

Traditional PageRank (Brin and Page 1998)

- **Intuition**: A node v_i will vote for v_j when there is an link from v_i to v_j, which will increase the rank of v_j.
- The score vector will be updated iteratively until convergency or reaching max number of iterations.
- At each step, the score vector of a graph G can be computed as:

$$\mathbf{Pr} = cM\mathbf{Pr} + (1 - c)\mathbf{v}$$

  Let d_i be the outdegree of a vertice v_i. M is a $N \times N$ matrix where $M_{ij} = \frac{1}{d_i}$. v is a $N \times 1$ vector where each element is $\frac{1}{N}$, and c is a scaling constant. Accordingly, the neighboring nodes of these favored ones will also ranked relatively higher than others.

Traditional PageRank - problem

- Tranditional PageRank initialize v unifomly. If running it over the whole graph, the highest ranked senses are independent of context.
- Solution 1: construct a subgraph of senses that are most relevant to context words.
- Solution 2: Personalizing PageRank.

Traditional PageRank over Subgraph(Agirre and Soroa, 2008)

G_D:Knowledge base.
$Concept_i$: The set of concepts of context word w_i.

1. For each word v_i in the concept set $Concepts_i$, use breath-first-search to find minimum distance path between node v_i and concepts linked with the rest of the context words, i.e. the set of path including all the $v_j \in \cup_{j \neq i} Concepts_j$. Denote the path as mdp_{v_i}
2. Construct G_D with all the nodes incluede in mdp_{v_i};
3. Run traditional PageRank over G_D.

Persoanlizing PageRank

- *v* Since Traditional PageRank is unifomly initialized, it has to work with a subgraph. I
- Solutions: We can favor some certain type of nodes by assigning them more initial mass.

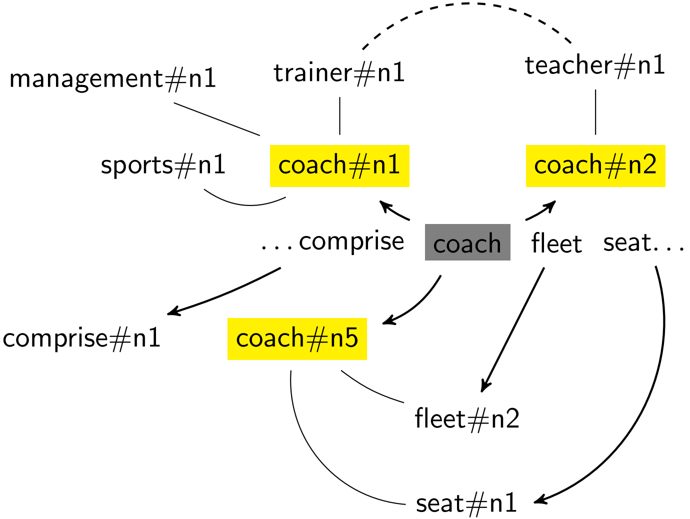
Personalizing PageRank

Steps:

1. Distribute the probability mass over the nodes of words including both the context words and target words.
2. Update the score vector iterateively using exactly the same formula as traditional PageRank.

Personalizing PageRank-example

"Our fleet comprises coaches from 35 to 58 seats."

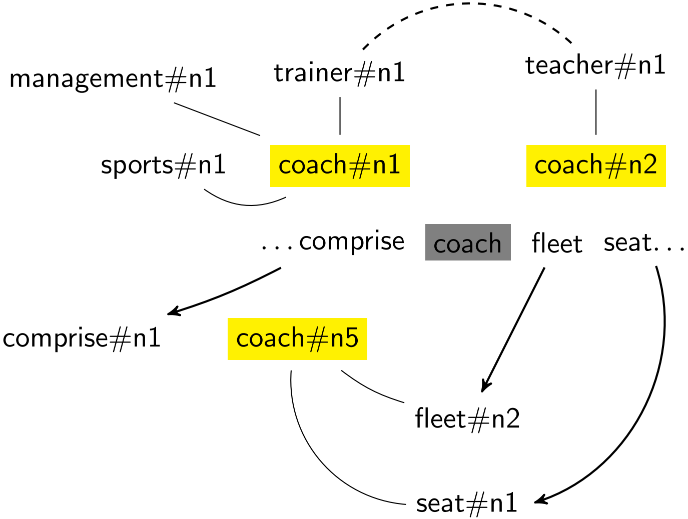


Personalizing PageRank-Problem

- **Problem**: if competing word senses are related with each other, they could be reinforced by each other.
- **Solution**: concentrate the initial probability mass in words surrounding target word.

Personalizing PageRank-w2w-example

"Our fleet comprises coaches from 35 to 58 seats."



Performance of Personalizing Pagerank

Senseval-2 dataset						
LKB	Method	All	N	V	Adj.	Adv.
MCR16 + Xwn	Ppr	51.1	64.9	38.1	57.4	47.5
MCR16 + Xwn	Ppr_w2w	53.3	64.5	38.6	**58.3**	38.1
MCR16 + Xwn	Spr	52.7	64.8	35.3	56.8	50.2
WNet17 + Xwn	Ppr	56.8	71.1	33.4	55.9	67.1
WNet17 + Xwn	Ppr_w2w	**58.6**	70.4	**38.9**	**58.3**	70.1
WNet17 + Xwn	Spr	56.7	66.8	37.7	57.6	**70.8**
WNet30 + gloss	Ppr	53.5	70.0	28.6	53.9	55.1
WNet30 + gloss	Ppr_w2w	55.8	**71.9**	34.4	53.8	57.5
WNet30 + gloss	Spr	54.8	68.9	35.1	55.2	56.5
MFS		60.1	71.2	39.0	61.1	75.4
SMUaw		68.6	78.0	52.9	69.9	81.7

Table: Results (recall) on Senseval-2 dataset

Performance of Personalizing Pagerank

Senseval-3 dataset						
LKB	Method	All	N	V	Adj.	Adv.
MCR16 + Xwn	Ppr	54.3	60.9	45.4	56.5	**92.9**
MCR16 + Xwn	Ppr_w2w	55.8	63.2	46.2	57.5	**92.9**
MCR16 + Xwn	Static	53.7	59.5	45.0	57.8	**92.9**
WNet17 + Xwn	Ppr	56.1	62.6	46.0	60.8	**92.9**
WNet17 + Xwn	Ppr_w2w	**57.4**	**64.1**	46.9	**62.6**	**92.9**
WNet17 + Xwn	Spr	56.20	61.6	**47.3**	61.8	**92.9**
WNet30 + gloss	Ppr	48.5	52.2	41.5	54.2	78.6
WNet30 + gloss	Ppr_w2w	51.6	59.0	40.2	57.2	78.6
WNet30 + gloss	Spr	45.4	54.1	31.4	52.5	78.6
MFS		62.3	69.3	53.6	63.7	92.9
GAMBL		65.2	70.8	59.3	65.3	100

Table: Results (recall) on Senseval-3 dataset

Performance of Personalizing Pagerank

Senseval-3 dataset	
Method	Time
Ppr	26m46
Spr	119m7
Ppr_w2w	164m4

Table: Elapsed time (in minutes) on Senseval-2 dataset

References

Eneko Agirre and Aitor Soroa (2009) Personalizing PageRank for Word Sense Disambiguation. Proceedings of the 12th conference of the European chapter of the Association for Computational Linguistics. Athens, Greece.

Eneko Agirre, Oier Lopez de Lacalle and Aitor Soroa (2013) Random Walks for Knowledge-Based Word Sense Disambiguation. Computational Linguistics.

S. Brin and L. Page. 1998. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30(1-7).

End

Thanks.