

#Lecture 13: diffusion & influence

10/10/13

<Pop>

• How can looking @ language (computationally) help us understand issues of influence?

- ~~not~~ a useful distinction:
 - understanding who is influential
 - vs.
 - understanding what ~~is~~ ways of arguing types of lang. ~~are~~ influential

Up until very recently the computational work was generally in the first category, and going on @ places like WWW and ICWSM:

A sort of sign

A typical approach is to look @ Twitter, say,

Stuff people are doing on information diffusion is related to this question.

'Memetracker', Leskovec, Backstrom, Kleinberg '09.

Schneider, Hua, Gianfranceschi, Das, Heilman, Black, Crabbe, Smith '01:
Simmons, Adamic, Adar '01 (changes as propagation occurs)

Find a particular piece of language that's easy to track.

In these cases, direct quotes (as found in news articles).

Use those 'tracers' to see when information originates, who picks it up, how far out it gets.

- could imagine expanding on this work via ~~paraphrase~~ NLP techniques for recognizing paraphrases,

so not just tracking the particular expression of an idea, but the idea itself.

... or, plagiarism detection

~~report refs to us~~ post refs to website.

~~learning approaches vs~~

~~statistical approaches: parallel or comparable corpora~~

(Alternate "tracers" besides direct quotes: attachments (Liu et al, Physica A '04)

~~But we could also~~

~~But what if we were actually~~

But we could also be interested in the text or phrasing itself. words in titles vs. people moving in to areas
Paraphrases, after all, may 'convey the same info', but have different connotations. (Backstrom et al, KDD '06)

"We made mistakes" vs. "Mistakes were made" [Broder, NYT 2007]

Surveys: Anagnostopoulos et al JAIR '10
Madhani: Dore CL '10

(William Schneider: "past exonerative" < -

learning: distributional similarity (Liu/Pantel '01), // comparable corpora (Bergsma, McKeown '01), etc.

urban legends
e.g. Health, Bell, Steinberg
J. Personality 2001:
'disgustingness'

Yahoo has had to control for external influence

(people who moved ≠ people whose titles had "hot" words later on ...)

A big potential area is framing:

presentation of an issue that fronts certain perspectives or presents a particular way of thinking about it.

Examples summed up @ the lexical level:

"pro-choice" vs. "pro-life"

"war on drugs" vs. "drug health-crisis"

↓
a smokes; an army; winning.

↓
drugs; victims; alleviation

"frankenfoods" vs. "~~war on terror~~"

green revolution

Beyond the lexical:

syntactic packaging (Greene; Resnik '09)

hedging → scientific perspective → GMO debates (Choi, et al '12)

(present as going the other way)

Tan
Lee
CD/AM
Spindel

Where can influence ^{on actual actions} be observed?

Voting settings, particularly when external influence can be controlled for:

→ notabilia.net [Taraborelli; Campagna]

visualizes the longest discussions, ~~each time~~ re article deletion proposals.

- each tendril = article, twist "left" = "keep" vote, ~~and~~ twist "right" = "delete" vote.

- note about Debatopedia: does
Brooks Lindsay write the majority of
articles?