

Concept hierarchies and Credibility

Jean Mark Gawron
Alex Dodge
Kathleen Burke

August 7, 2013

1 Introduction

- Group studies
- Ontologies and concept hierarchies

2 Groups and conceptual frameworks

3 Credibility and importance

4 Tools



Studying groups organized for collective action



Project task

Task 3: Acquire the vocabulary and concept hierarchies for groups organized around ideas or themes or for collective action.

- 1 What are the defining features of the group? Do they constitute a real world community? An online community?
- 2 What are their goals?
- 3 What are their ideas?
- 4 What are they likely to do?
- 5 What are the characteristic features of their language and their discourse? What do they talk about? How do they talk about it?



Three groups (kinds of groups)



Website data collected from Google queries, mixture of blogs, forums, and organization websites

- 1 White militants
- 2 Vaccine debates [anti vaccine movement, pro vaccine websites, mostly professional or government]
- 3 Climate change [pro and anti]



Two applications



- 1 Word clustering
- 2 Relation extraction with Random Fields



Ontologies



Wikipedia definition

... represents knowledge as a set of concepts within a domain, and the relationships between pairs of concepts. It can be used to model a domain and support reasoning about concepts. ... a “formal specification of a shared conceptualisation” ... provides a shared vocabulary, which can be used to model ... the type of objects and/or concepts that exist, and their properties and relations.

consists-of(water, [oxygen, hydrogen])

subtype(car-make, [Volvo, Volkswagen, BMW])



Relation extraction

CONSIST-OF

- 1 Water consists of hydrogen and oxygen.
- 2 Water is made up of hydrogen and oxygen.
- 3 Hydrogen and oxygen combine to make water.
- 4 Hydrogen bonds with oxygen to make water.
- 5 Hydrogen combined with oxygen makes water.

Group discourse analogues: Ideology components

What are the salient **relations** in the data? What are the texts talking about?

The vocabulary-based method



- 1 Collect data on group documents (websites, blogs, Tweets)
- 2 Characterize each word usage patterns in a set of documents, using syntactic analysis
- 3 Clustering words with with similar distributions together
- 4 A hierarchy of words suggesting concepts and conceptual connections

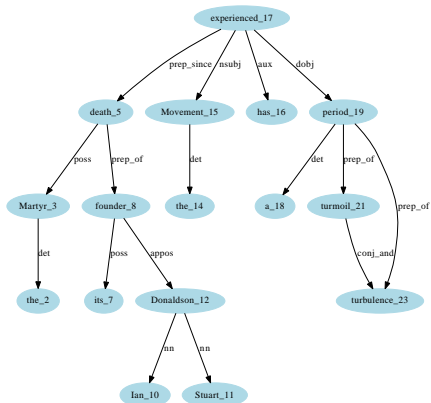
White Militant Data I



An example

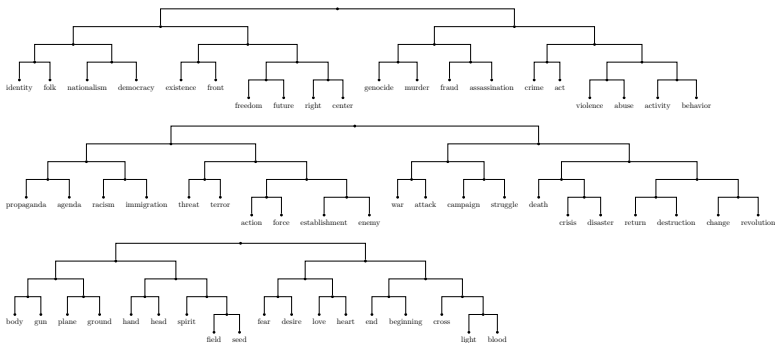
Blood & Honour is a world-wide pan Aryan organization dedicated to the struggle for survival and prosperity of the White Race. Since the Martyr's death of its founder, Ian Stuart Donaldson, the Movement has experienced a period of turmoil and turbulence. Profiteers and pop stars have -unsuccessfully- tried to abuse B&H for their own selfish puposes. ZOG has geared up the campaign of oppression and the media has constantly smeared the national revolutionary forces. Members and supporters all over the White world have often been misinformed through infiltrators or left on their own due to the problems of the main organization...

White Militant Data II



Since the Martyr 's death of its founder , Ian Stuart Donaldson , the Movement has experienced a period of turmoil and turbulence .

White militant data III



Climate change judgments



- Target/Importance** Global climate change, conversely, may not have such a substantially negative effect on the world's population, standards of living, health, and survival.
- Target/Sentiment** When you start to explore the issue, climate change can be terrifying.
- Information/Source** Million Climate Jobs Campaign has released a new booklet describing how the transformation of the economy to one that protects people and the environment, can confront two of South Africa's biggest threats: climate change and unemployment
- Target/credibility** What is true is that climate change is real and its already happening.

Credibility judgments



- 1 X is credible
 - 1 **informational targets:** pronouncements, claims, stories
 - 2 **sources**
- 2 Kinds of judgments
 - Debunking** Calling into question the credibility or reliability of a story
 - Legimitizing** Affirming the credibility of a source

Tasks



Wiebe (2003) [sources, predicates, MPQA corpus]

- 1 Identify $\left\{ \begin{array}{l} \text{Sources} \\ \text{Targets} \\ \text{Predicates} \end{array} \right.$
- 2 Infer polarities (positive/negative credibility)

Key sentences: Climate change



Anyone who says they can confidently predict global climate change or effects is either a fool or a fraud.

Tags	Definition
Source	Source of information (thought, speech, writing)
Predicate	Word or phrase expressing credibility, importance, or sentiment opinion (positive, negative)
Target	The message or phenonemon about which the opinion is held, by the source, if there is a source, or by the author; simultaneous judgments possible: “lies emanating from the White House”

Example 1



[_{source} Anyone who says
 [_{source} they] can
 [_{pred} OPINION(credibility,positive) confidently]
 [_{pred} OPINION(credibility,positive) predict]
 [_{target} global climate changes or effects]
]
 is either a [_{pred} OPINION(credibility,negative) fool]
 or a [_{pred} OPINION(credibility, negative) fraud] .

Example II



[*target* That is how [*source* they]
 [*pred* (credibility,positive,negative) make it seem]
 [*target* that all the top scientists are agreed,]
]
 [*source* he] [*pred*(credibility,positive) says] .



Some annotated data



fully_annotated_test_preprocessed_1.xml

6 | I think we're seeing almost weekly, or even daily, scientists that are coming forward and questioning the original idea

7 | He said some want billions or trillions of taxpayer dollars spent to address the issue, but he added: "I do n't think from

8 | His comments came at a packed breakfast meeting with local business leaders in a region known for its strong environmental pr

9 | And he made his global warming comment in response to a question by an audience member who cited evidence from the National

10 | But Perry's opinion runs counter to the view held by an overwhelming majority of scientists that pollution released from the

11 | Perry's home state of Texas releases more heat-trapping pollution carbon dioxide -- ; the chief greenhouse gas -- ; than any

12 | Global warming has become an issue for contenders for the Republican nomination to run away from, since many conservatives

13 | Jon Huntsman, Newt Gingrich and Mitt Romney -- ; who all at one point supported steps to curb global warming pollution -- ;

14 | But unlike Perry, both Romney and Huntsman acknowledge the scientific evidence.

15 | On Wednesday, Perry promised to return regularly to a state that was not kind to a past Texas governor; Arizona Sen. John

16 | For many New Hampshire voters, Wednesday offered their first close look at the longtime Texas politician, who formally laur

17 | At the breakfast, Perry also questioned the loyalty of the Federal Reserve, just days after saying that if the Federal Rese

18 | He noted the criticism he took for the comment, but did not back away from them.

19 | And he called on the institution to open its books.

20 | "It would go a long way toward either finding out whether or not there is some activities that are improper of that they

21 | But until they do that, I think there will continue to be questions about their activity and what their true goal is for

Save Find Quit Help

Line no 10 annotator b edits? 1



Predication by predication



Line Editor

File Edit Help

Line no: 18

Line	He			noted			the criticism he took for the comment			, but did not back away from them	
CF	source			predicate			target				
opinion_type				credibility							
polarity				positive							
Line	He			noted the criticism he took for the comment , but did not			back away			from them	
CF	source						predicate			target	
opinion_type							credibility				
polarity							negative				

Quit Commit

edits? 0

Factoring the problem



“He said she said global warming is a myth.”

Nested brackets:

[*source* He] [*pred* said] [*target* [*source* she] [*pred* said] [*target* [*target* global warming] is a [*predicate* myth].]]

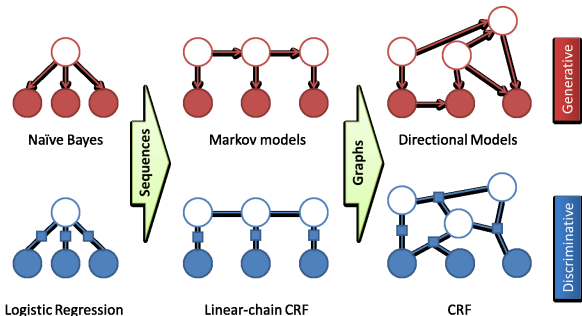
Shallow brackets:

[*source* He] [*pred* said][*source* she] [*pred* said] [*target* global warming] is a [*predicate* myth].

Tags:

He/B-Source said/B-Pred she/B-Source said/B-Pred
global/B-Target warming/I-Target is/O a/O myth/B-Pred ./O

Conditional Random Fields



Adapted from C. Sutton, A. McCallum, "An Introduction to Conditional Random Fields", ArXiv, November 2010