Argument Mining: A Combined Approach

John Lawrence

Centre for Argument Technology
University of Dundee

ARG-tech
Centre for Argument Technology

Argument Mining

ARG-tech
Centre for Argument Technology

Argument Mining: A Combined Approach

John Lawrence

Premise vs Conclusion

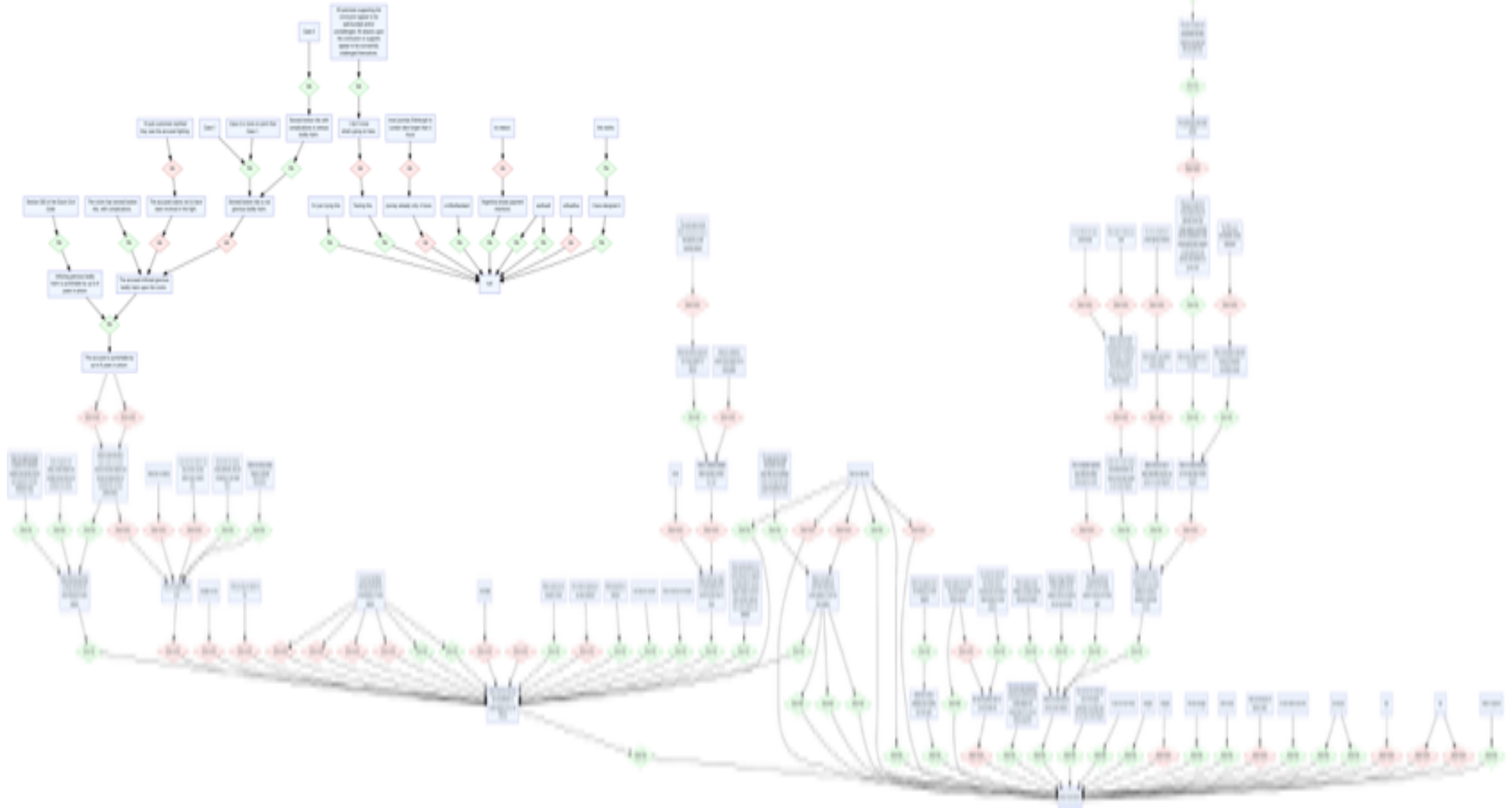
Argument vs Non-Argument

Support vs Conflict

…

Bigger Picture

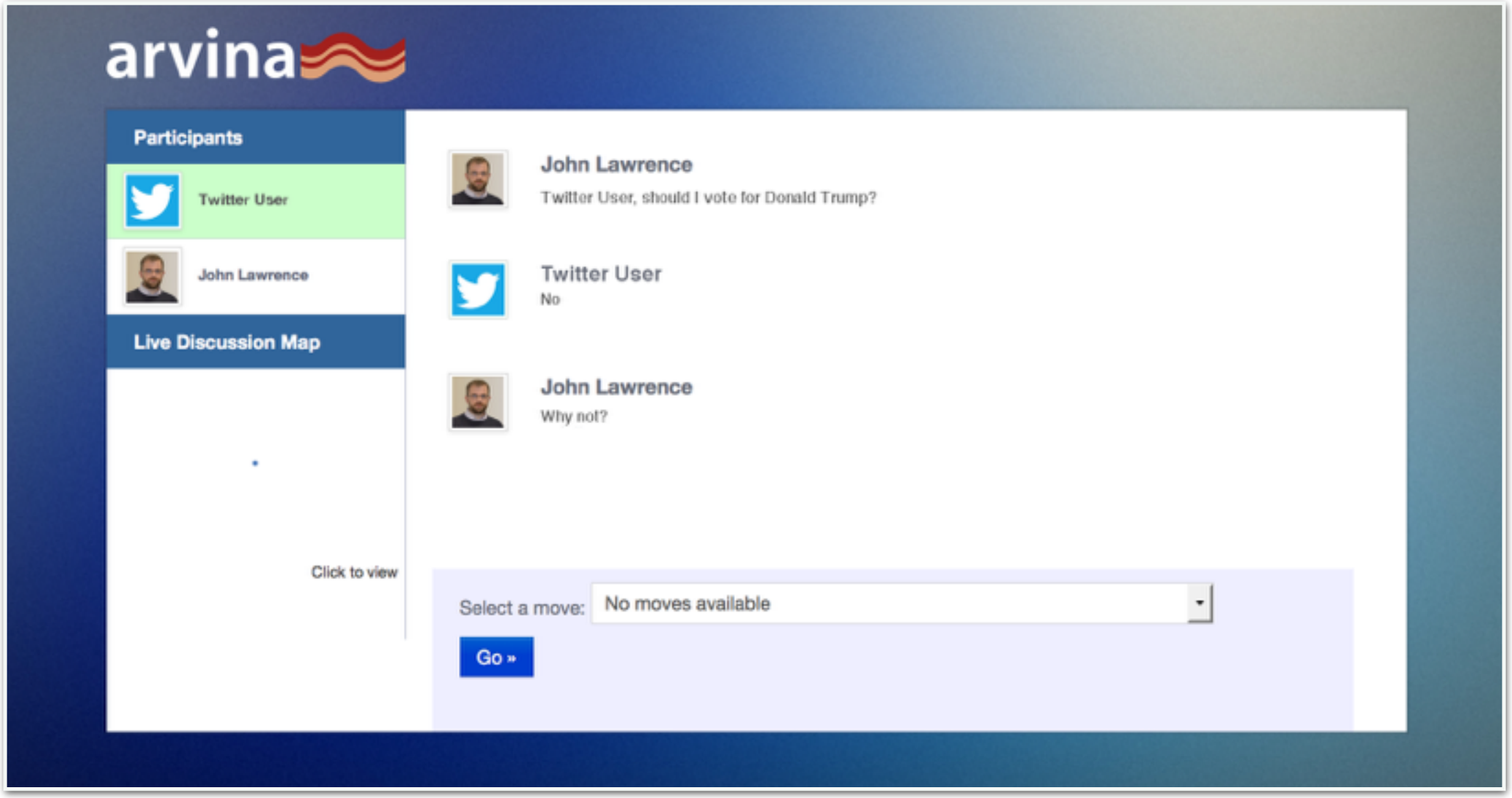
ELECTION 2016
Election 2016
HollywoodLife.com
Click For More Pics Of Presidental Debate '16
Donald J. Trump @realDonaldTrump · 4h
"@FrankLuntz: I'm getting a lot of @MegynKelly hatemail tonight. #GOPDebate" She is totally overrated and angry. She really bombed tonite
357
804
Donald J. Trump @realDonaldTrump · 4h
.@FrankLuntz is a low class slob who came
ce looking for consulting work and
interest. Now he picks anti-
anels!
948
@realDonaldTrump · 4h
negynkelly @FrankLuntz @realDonaldTrump Fox viewers
to bimbo @MegynKelly will consider other programs!"
476
@realDonaldTrump · 4h
reno: @realDonaldTrump The biggest loser in the
megynkelly. You can't out trump Donald Trump. You will
CNN
CNN
CNN
CNN
DEMOCRATIC PRESIDENTIAL DEBATE
TUESDAY 8:30 PM ET
THE DEMOCRATS
RUNNING
CHAFEE
CLINTON
LESSIG
O'MALLEY
SANDERS
WEBB
EXPLORING
BIDEN
NOT RUNNING
WARREN
THE REPUBLICANS
RUNNING
BUSH
CARSON
CHRISTIE
CRUZ
FIORINA
GILMORE
GRAHAM
HUCKABEE
JINDAL
KASICH
PATAKI
PAUL
RUBIO
SANTORUM
TRUMP
DROPPED OUT
PERRY
WALKER
NOT RUNNING
ROMNEY



ARG-tech
Centre for Argument Technology

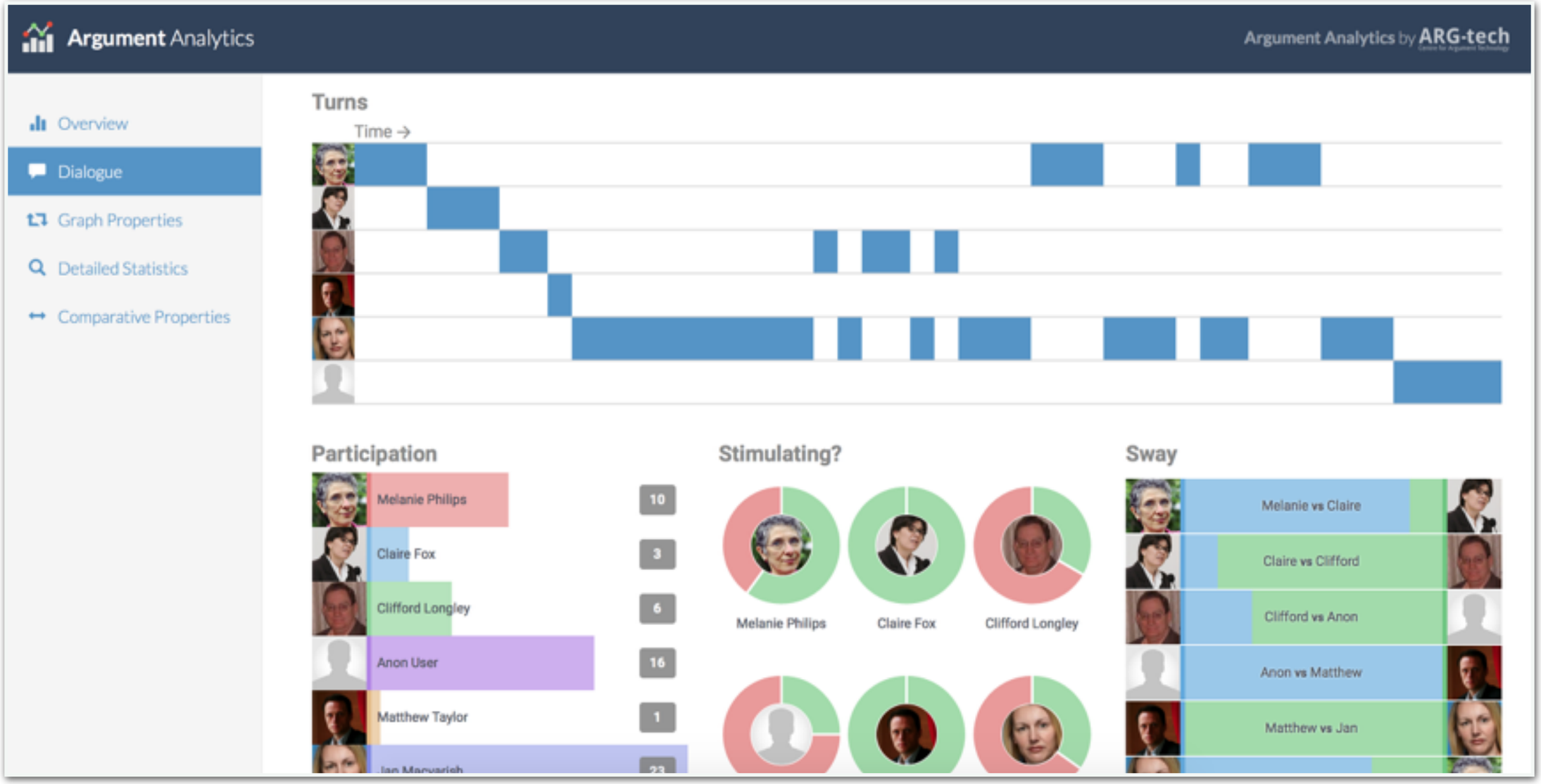
Argument Mining: A Combined Approach

John Lawrence



arvina

ARG-tech
Centre for Argument Technology



Argument Analytics

ARG-tech
Centre for Argument Technology

Argument Mining: A Combined Approach

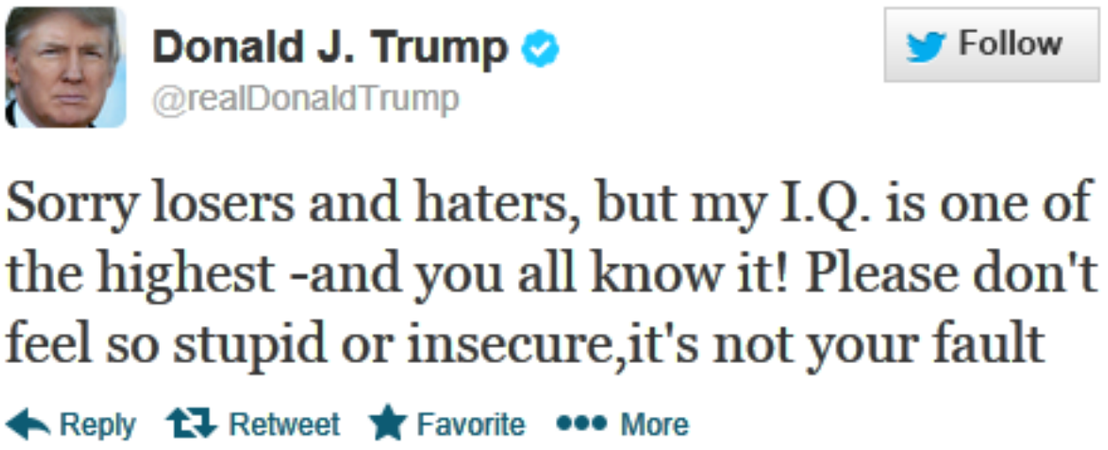
John Lawrence

Human Annotation

ARG-tech
Centre for Argument Technology

Example

Trump's speech was poor. The speech was "lacking in policy prescriptions," and its "strident rhetoric masked a lack of depth," said Robert McFarlane, a former national security adviser. However, his popularity in the polls continues to rise, perhaps because of his recently self-declared high IQ.



Example

Trump's speech was poor. The speech was "lacking in policy prescriptions," and its "strident rhetoric masked a lack of depth," said Robert McFarlane, a former national security adviser. **However**, his popularity in the polls continues to rise, perhaps **because** of his recently self-declared high IQ.

Discourse Indicators (“However”, “because”)

Example

Trump's speech was poor. The speech was "lacking in policy prescriptions," and its "strident rhetoric masked a lack of depth," said Robert McFarlane, a former national security adviser.

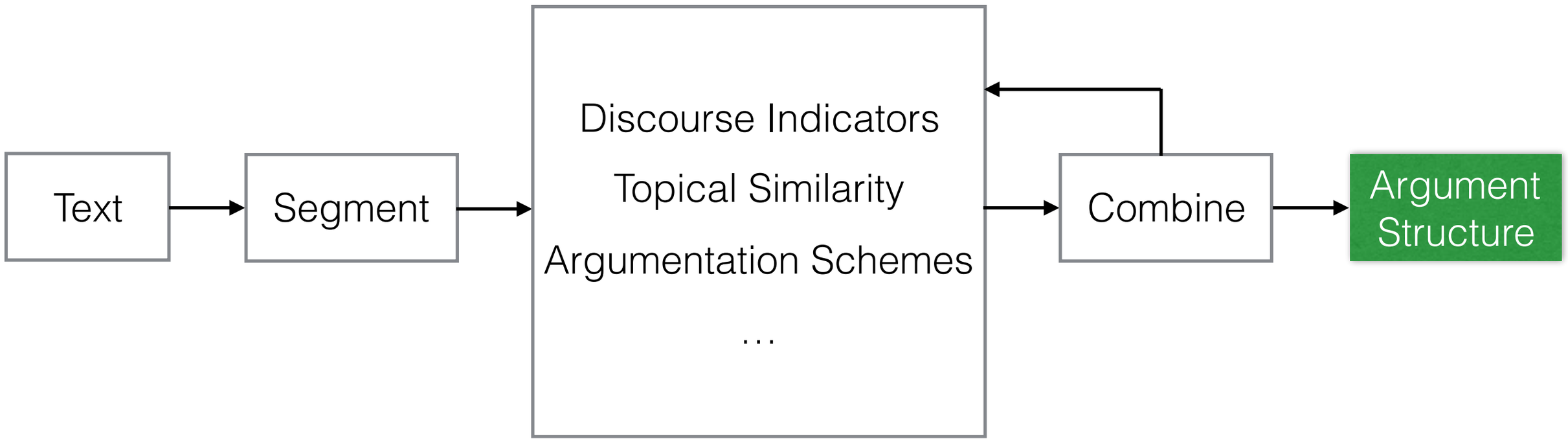
However, his popularity in the polls continues to rise, perhaps because of his recently self-declared high IQ.

Topical Similarity

Example

Trump's speech was poor. The speech was "lacking in policy prescriptions," and its "strident rhetoric masked a lack of depth," said Robert McFarlane, a former national security adviser. However, his popularity in the polls continues to rise, perhaps because of his recently self-declared high IQ.

Argumentation Schemes
(Argument from expert opinion)


Text
Segment
Discourse Indicators
Topical Similarity
Argumentation Schemes
…
Combine
Argument Structure

Discourse Indicators

ARG-tech
Centre for Argument Technology

John Lawrence

Discourse Indicators

Relation Type	Words
Support	because, therefore, after, for, since, when, assuming, so, accordingly, thus, hence, then, consequently,…
Conflict	however, but, though, except, not, never, no, whereas, nonetheless, yet, despite,…

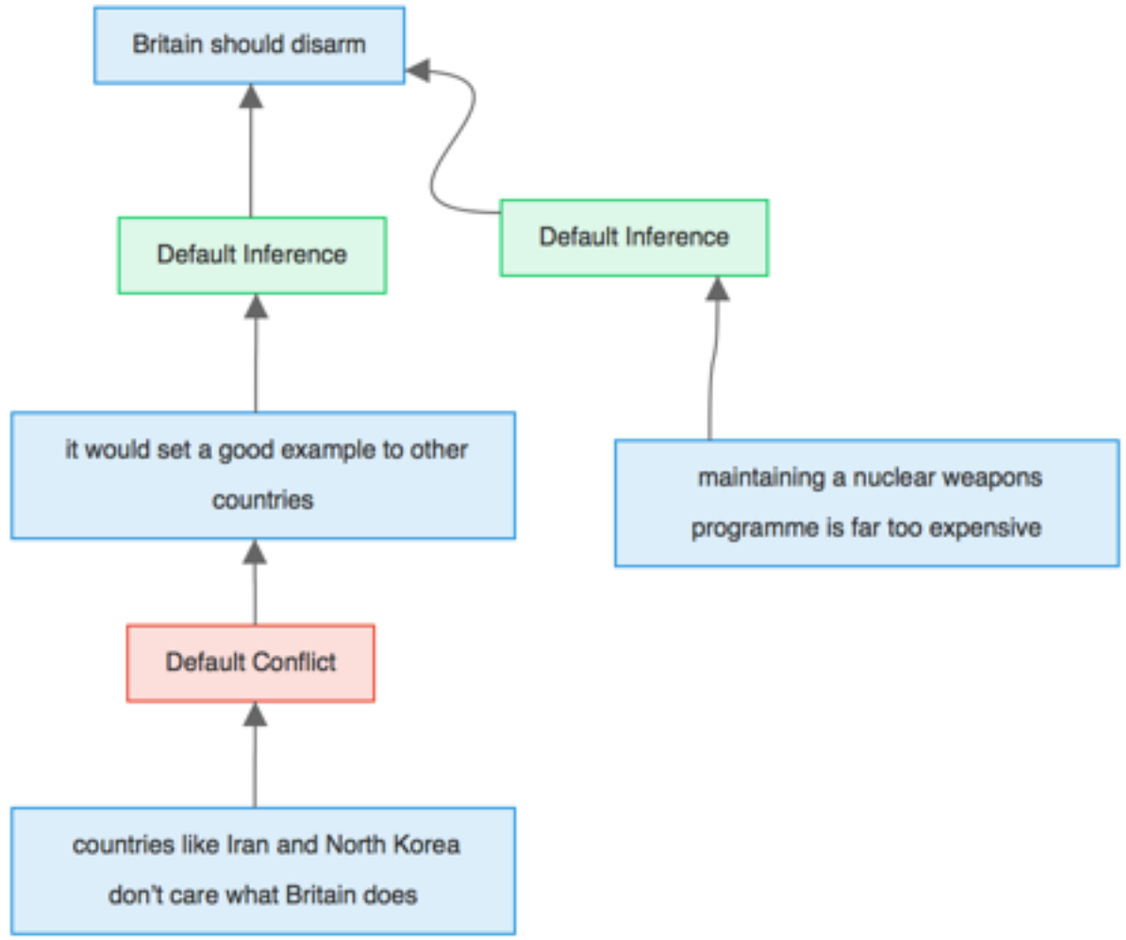
Discourse Indicators

Precision	Recall	f-score
0.89	0.04	0.07

Topical Similarity

Topical Similarity

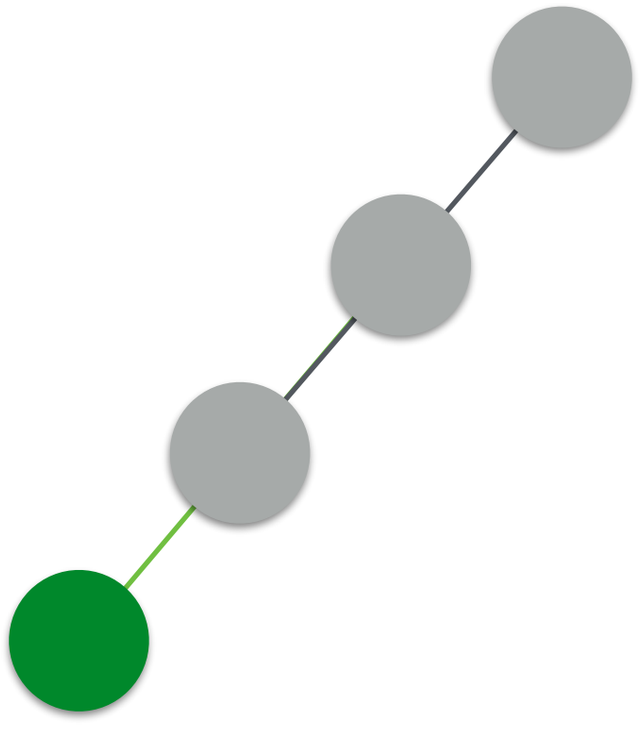
Britain should disarm because it would set a good example to other countries. However, countries like Iran and North Korea don't care what Britain does. And anyway, maintaining a nuclear weapons programme is far too expensive.



Topical Similarity

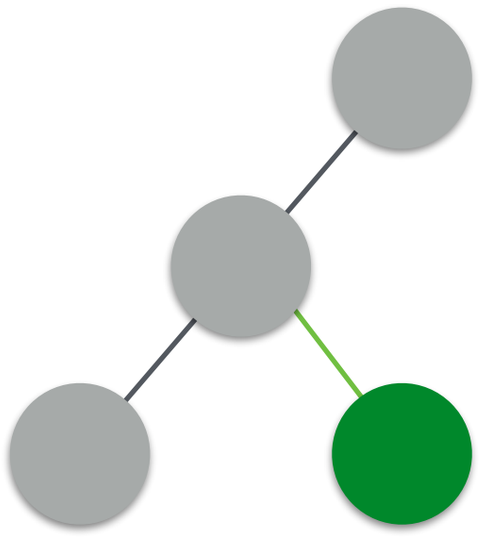
Look at propositions in order.

1. If proposition is similar to its predecessor, then they are connected.



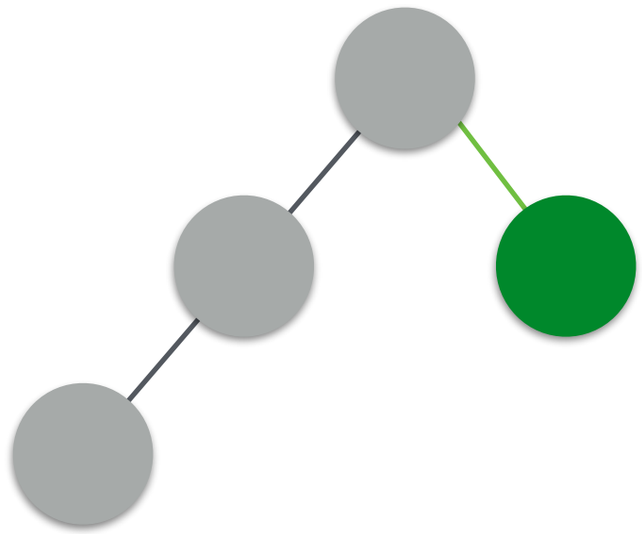
Topical Similarity

2. If not, then look at each of the previous propositions. If the new proposition is similar to a previous one, then they are connected.



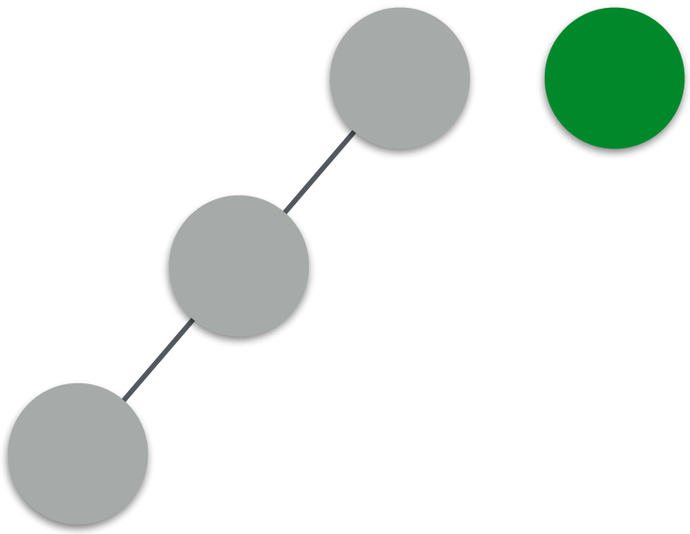
Topical Similarity

2. If not, then look at each of the previous propositions. If the new proposition is similar to a previous one, then they are connected.



Topical Similarity

3. If the new proposition is not similar to any of the previous ones it is left unconnected.



Topical Similarity

Precision	Recall	f-score
0.82	0.56	0.67

Success if two propositions are connected in the model analysis.

Topical similarity does not give direction of edges.

Argumentation Scheme Structures

ARG-tech
Centre for Argument Technology

John Lawrence

Analogy

Premise: Generally, case C1 is similar to case C2 [SimilarityOfCases]

Premise: A is true (false) in case C1 [Precedent]

Conclusion: A is true (false) in case C2

CauseToEffect

Premise: Generally, if A occurs, then B will (might) occur [Causal]

Premise: In this case, A occurs (might occur) [OccurrenceA]

Conclusion: Therefore, in this case, B will (might) occur

PracticalReasoning

Premise: I have a goal G [Goal]

Premise: Carrying out this action A is a means to realise G [GoalPlan]

Conclusion: Therefore, I ought (practically speaking) to carry out this action A

VerbalClassification

Premise: a has a property F [ContainsProperty]

Premise: For all x, if x has a property F, then x can be classified as having a property G [ClassificationProperty]

Conclusion: a has property G

Identify segments as scheme **components**

Classifiers for each component

Unigrams	Each word in the proposition
Bigrams	Each pair of successive words
Length	The number of words in the proposition
AvgWLength	The average length of words in the proposition
POS	The parts of speech contained in the proposition
Punctuation	The presence of certain punctuation characters, for example “ ” indicating a quote
Similarity	The maximum similarity of a word in the proposition to pre-defined words corresponding to each proposition type

Type	**Keywords**
AN Similar	similar, generally
AN Precedent	be (to be)
AN Conc	be (to be)
CE Causal	generally, occurs
CE Occurance	occurs
CE Conc	occurs
PR Goal	goal
PR GoalPlan	action
PR Conc	ought, perform
VC Property	be (to be)
VC Class	all, if
VC Conc	be (to be)

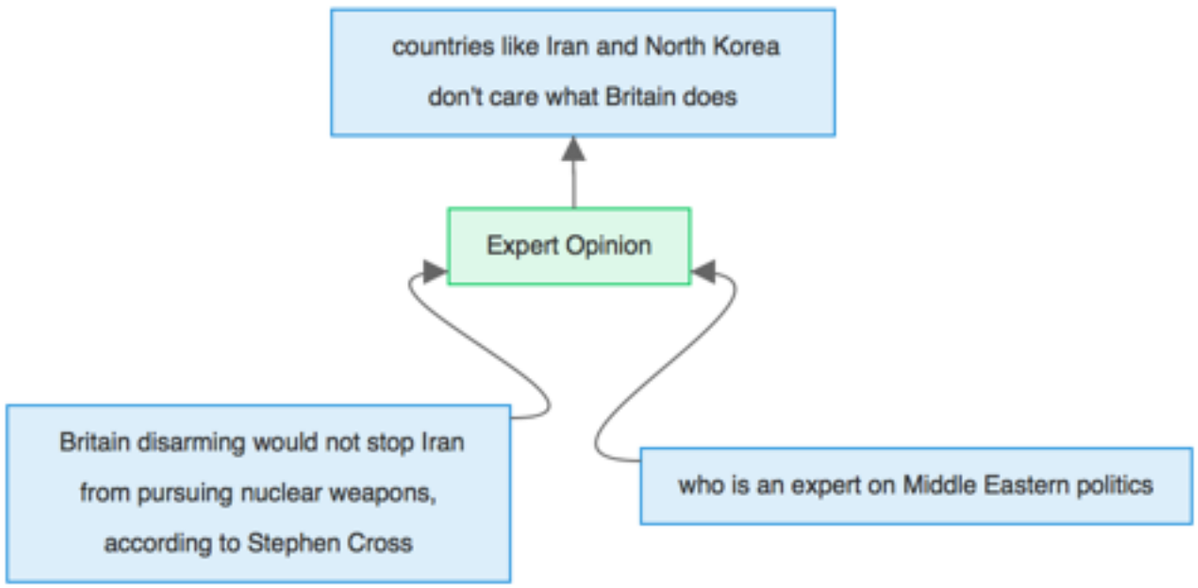
Argumentation Scheme Structure

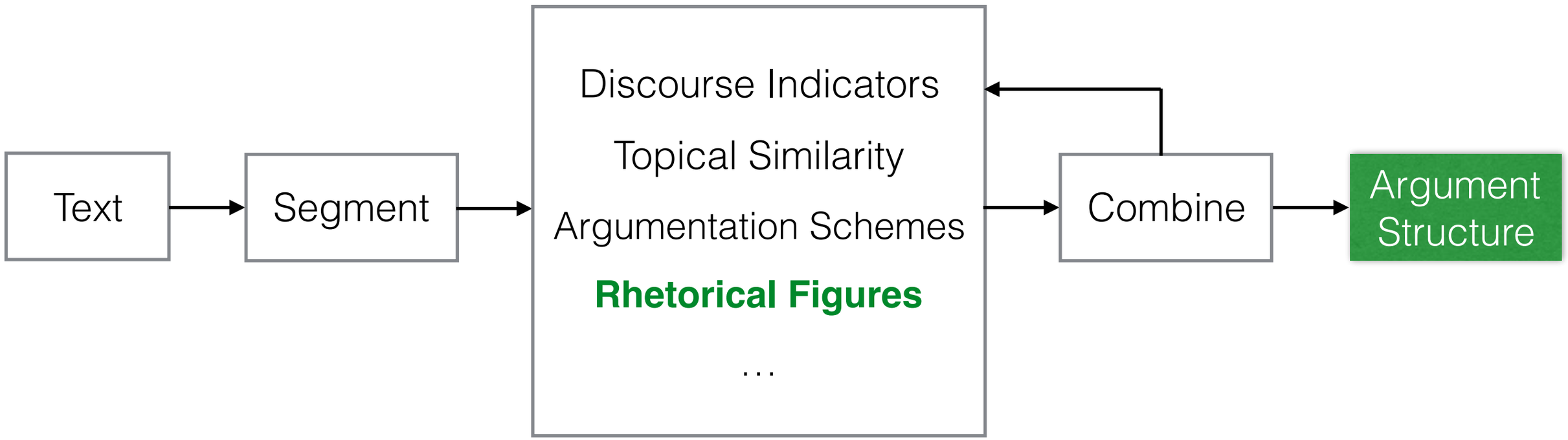
Expert Opinion

Premise: Source E is an expert in subject domain S containing proposition A [FieldExpertise]

Premise: E asserts that proposition A is true (false) [KnowledgeAssertion]

Conclusion: A is true (false) [KnowledgePosition]




Text
Segment
Discourse Indicators
Topical Similarity
Argumentation Schemes
Rhetorical Figures
…
Combine
Argument Structure

Rhetorical Figures

ARG-tech
Centre for Argument Technology

- Selected rhetorical figures from a range of different classes (all easily automatically identified)

- Searched the Moral Maze 2012 corpus for instances of each figure

- Compared identified instances of figures with the pre-annotated argumentative structure

Dirimens Copulatio

A figure by which one balances or outweighs one statement with a contrary, qualifying statement.

"not only...but also"

Trust **not only** requires moral integrity, **but also** competence. And I think that competence is lacking in the banking industry.

Epizeuxis

Repetition of a word or small group of words, with no other words in between

Well, as we just heard, because the system seems to be **very, very** heavily weighted against them, many savers have spent much of their life putting money by for old age

Polyptoton

repetition of words derived from the same root but with different endings

They ought to be leaning towards the common good and not towards personal and **individual** interest. Because that's a hard concept to get hold of in a very **individualised** society, and of course the free market system itself is based up on a very **individualistic** anthropology.

Antithesis

The juxtaposition of contrasting words or ideas

Do we really need more rules to behave **better**? Or, will that make us **worse**, reduce our moral agency to box ticking

Eutrepismus

Numbering and ordering of the parts under consideration

The growth of what you call ‘welfarism’ is to do with, **firstly** the disappearance of so many skilled middle income jobs; and **secondly** due to the failure to build social housing.

Anaphora

The use of the same word or group of words at the beginning of successive clauses, sentences, or lines

It's not necessarily a moral good. **It can be** a way of demonstrating independence. **It can be** risk averse. **It can be** conservative.

	I-nodes	Support	Incoming	Outgoing	Conflict	Incoming	Outgoing
MM2012c Corpus	2037	0.795	-	-	0.159	-	-
Dirimens Copulatio	8	0.929	5	3	0.571	2	2
Epizeuxis	15	0.854	7	7	0.176	0	3
Polyptoton	24	1.119	4	7	0.202	1	3
Antithesis	98	0.853	23	28	0.147	3	8
Eutrepismus	5	1.334	2	4	0.167	0	1
Anadiplosis	7	0.334	0	1	1.667	1	4
Anaphora	21	0.894	4	10	0	0	0
Epistrophe	20	0.916	5	7	0.134	1	2

Dirimens Copulatio

- Introduce a new concept often in contrast to an existing point
- Or, amplify a topic, making it seem larger and more striking
- Both the support rate (0.929) and especially the conflict rate (0.571) are significantly above that of the corpus as a whole
- The only figure, out of those that we are considering, where the number of incoming supporting connections is greater than that of outgoing supporting connections

Epizeuxis

- By far the most common usage of this figure in our corpus is “very, very” with eleven of seventeen examples using this wording

- Often in conflict with a previous point, but are themselves conflicted with less

- All of the conflict relations are outgoing

Polyptoton

- The frequency of support relations connected to polyptotonic nodes is strikingly high (1.119)

- Often multiple I-nodes which are comprising the structure of this figure, and furthermore that the majority of the support is occurring between these nodes

- 24 I-nodes in just 7 turns

- In all cases these nodes are being used to build an argument, which although it may, as a whole, be in conflict with external issues, is strongly internally supported

Antithesis

- Although we might expect that the usage of antithesis would commonly indicate the presence of conflict, the frequency of conflict connections is actually slightly below that for the corpus as a whole

Investment banks that have highly public **ethical** problems don't lose their clients. So people are still going to them, probably because they are **unethical** rather than despite

ethical problems = unethical

There are **good** nation states and **bad** nation states

no real connection being made between these two concepts, they are neither supporting or conflicting with each other but are completely independent.

Eutrepismus

- Readily correspond to a list of reasons being provided for a particular viewpoint
- Eutrepismus is a strong indicator of support, with an overall rate of 1.334

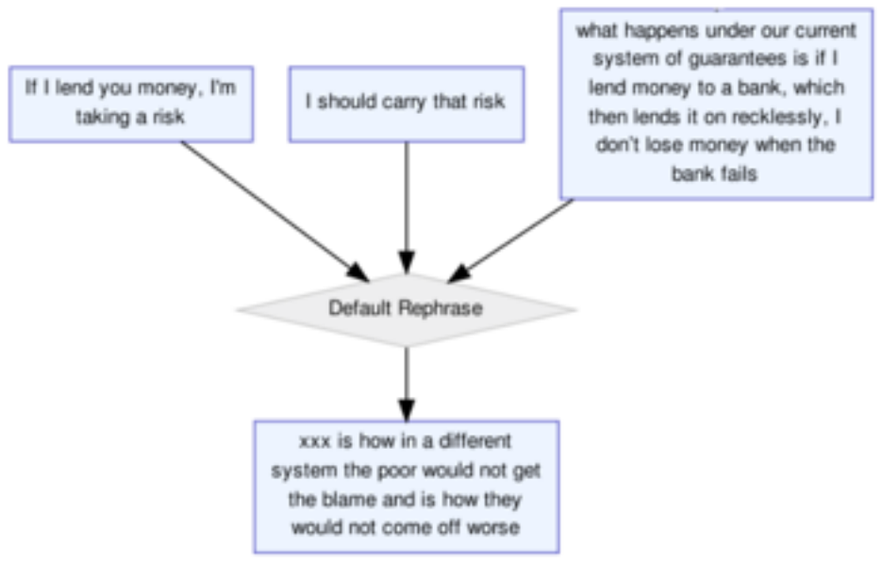
Anaphora

- Good indicator of several premises working together to support a conclusion

Epistrophe

- Indicator of nodes that are siblings in the argumentative structure

- 6 out of the 20 nodes identified (30%) are siblings. . Significantly above the rate at which siblings occur in the corpus as a whole (15.1%)



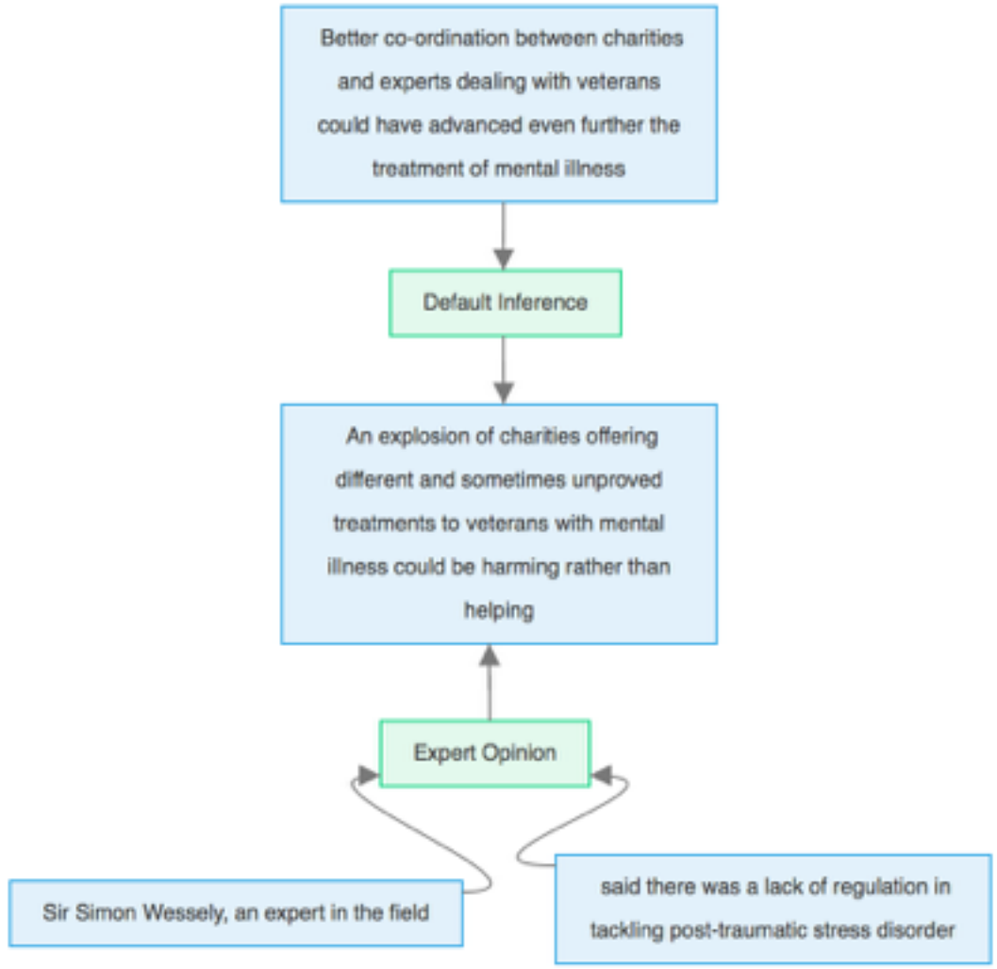
Extra Material

Combining Techniques

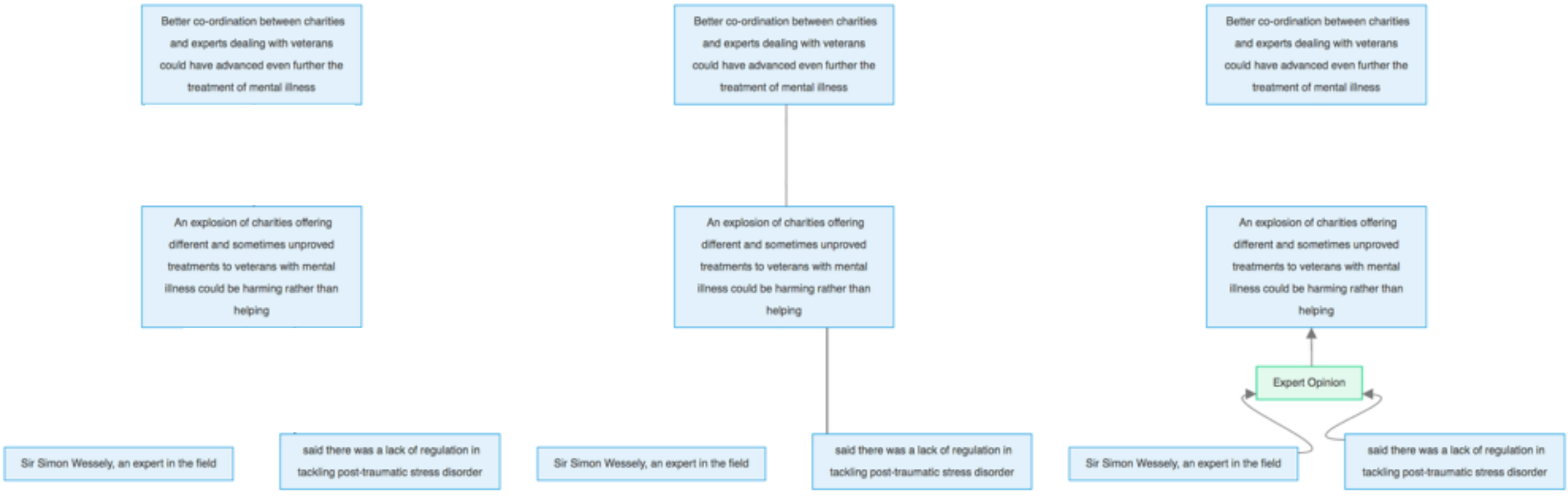
ARG-tech
Centre for Argument Technology

An explosion of charities offering different and sometimes unproved treatments to veterans with mental illness could be harming rather than helping, it was claimed last night.

Sir Simon Wessely, an expert in the field there was a lack of regulation in tackling post-traumatic stress disorder. Better co-ordination between charities and experts dealing with veterans could have advanced even further the treatment of mental illness



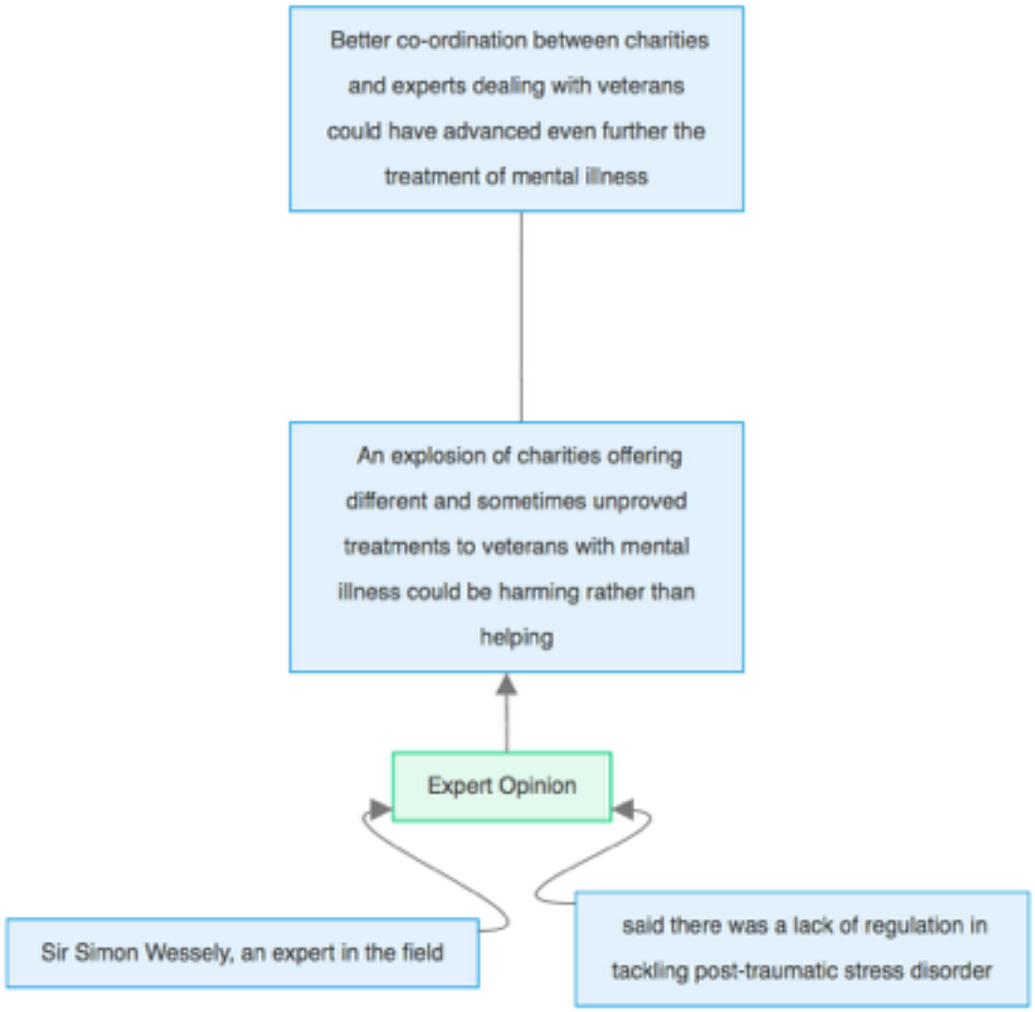
Manual Analysis



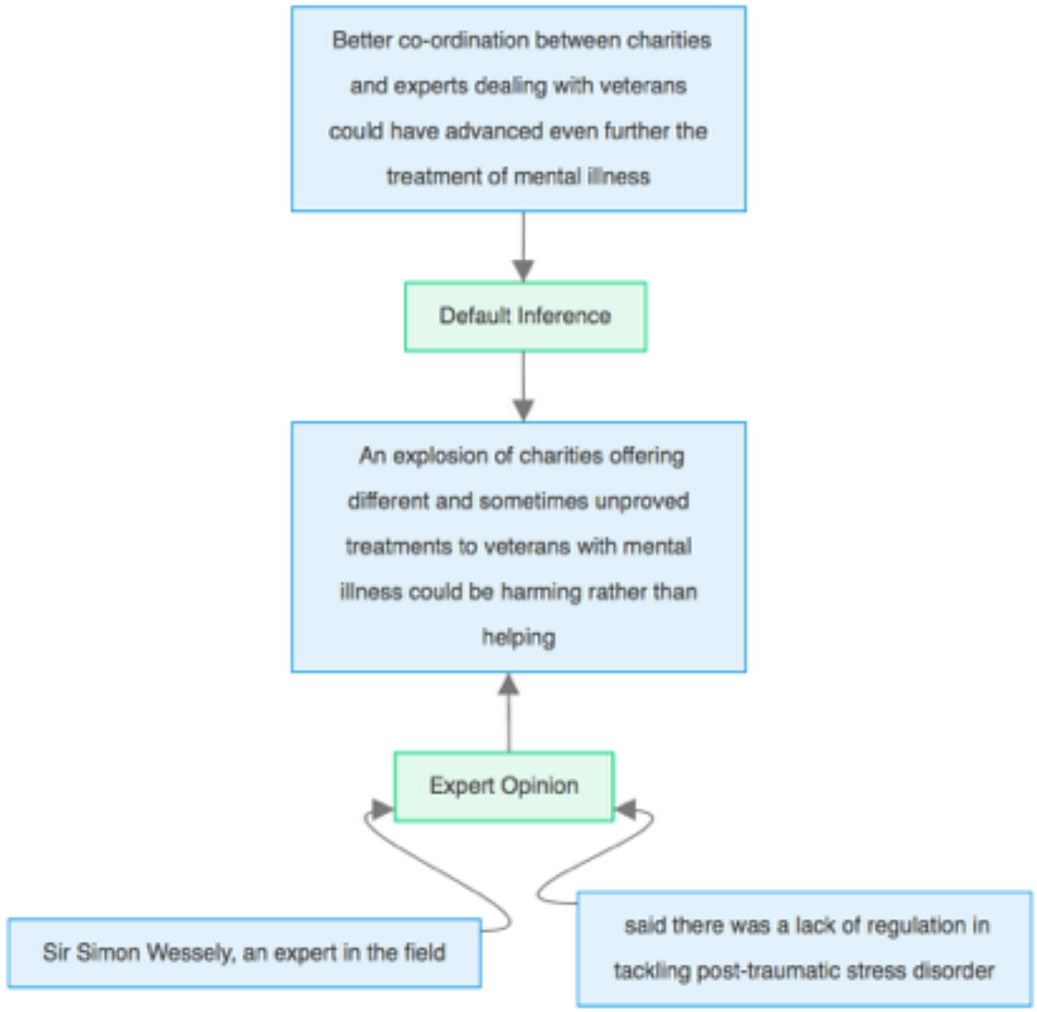
Discourse Indicators

Topical Similarity

Scheme Structure



Combined Techniques



Manual Analysis