#### Argument Corpora, and Argument Mining and Analytics (PART 1)

Katarzyna Budzynska<sup>1,2</sup> & Chris Reed<sup>2,1</sup>

Centre for Argument Technology (www.arg.tech)

<sup>1</sup>Institute of Philosophy & Sociology, Polish Academy of Sciences

<sup>2</sup>University of Dundee, UK



#### **ARG-tech (Warsaw & Dundee)**



Kasia Budzynska (PL, UK)

Rory Duthie (UK)

Basia Konat (UK, PL)

Marcin Koszowy (PL)

Olena Yaskorska (PL)



Chris Reed (UK, PL)

Mathilde Janier (UK)

John Lawrence (UK)

Martín Pereira Fariña (UK)

Mark Snaith (UK)

Jacky Visser (UK)



## Argument Mining (Budzynska & Villata 2017)

**Before 2014:** fewer than 10 papers

#### After 2014:

- 1) From Real Data to Argument Mining, 12th ArgDiaP Conference, Polish Academy of Sciences, Warsaw, Poland, 23-24 May 2014. (22 papers)
- 2) 1st Workshop on Argumentation Mining, 52nd Annual Meeting of the Association for Computational Linguistics, Baltimore, US, June 26, 2014 (18 papers)
- 3) SICSA Workshop on Argument Mining: Perspectives from Information Extraction, Information Retrieval and Computational Linguistics, ARG-tech, Dundee, Scotland, 9-10 July 2014. (ca. 25 participants)
- *4)* Frontiers and Connections between Argumentation Theory and Natural Language *Processing*, Bertinoro (Forlì-Cesena), Italy, 21-25 July 2014.
- *5)* 2nd Workshop on Argumentation Mining, 53nd Annual Meeting of the Association for Computational Linguistics, Denver, US, June 04, 2015 (16 papers)
- 6) Arguments in Natural Language: The Long Way to Analyze the Reasons to Believe and the Reasons to Act, 1st European Conference on Argumentation, Lisbon, 9-12 July 2015
- 7) Natural Language Argumentation: Mining, Processing, and Reasoning over Textual Arguments, Dagstuhl Seminar, April 17 22, 2016
- 8) 3rd Workshop on Argumentation Mining, 54nd Annual Meeting of the Association for Computational Linguistics), Berlin, Aug 12, 2016 (20 papers)

#### (1/4) Foundations



# (1a/4) Arguments and Dialogue: Theoretical Foundations (Budzynska & Reed, 2011)



#### Inference Anchoring Theory, IAT

#### **DIALOGUE:**

- (a) Alexander Brown said, If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage.
- **(b)** Melanie Philips said, Why?
- **(c)** Alexander Brown said, *Evidence shows that those who are in poverty are more likely to be facing all of these other forms of disadvantage.*

#### ARGUMENTATION (INFERENCE):

(Arg1) If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage because Evidence shows that those who are in poverty are more likely to be facing all of these other forms of disadvantage.

If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage

rule applic. inst. #1

Evidence shows that those who

are in poverty are more likely to

be facing all of these other forms

of disadvantage

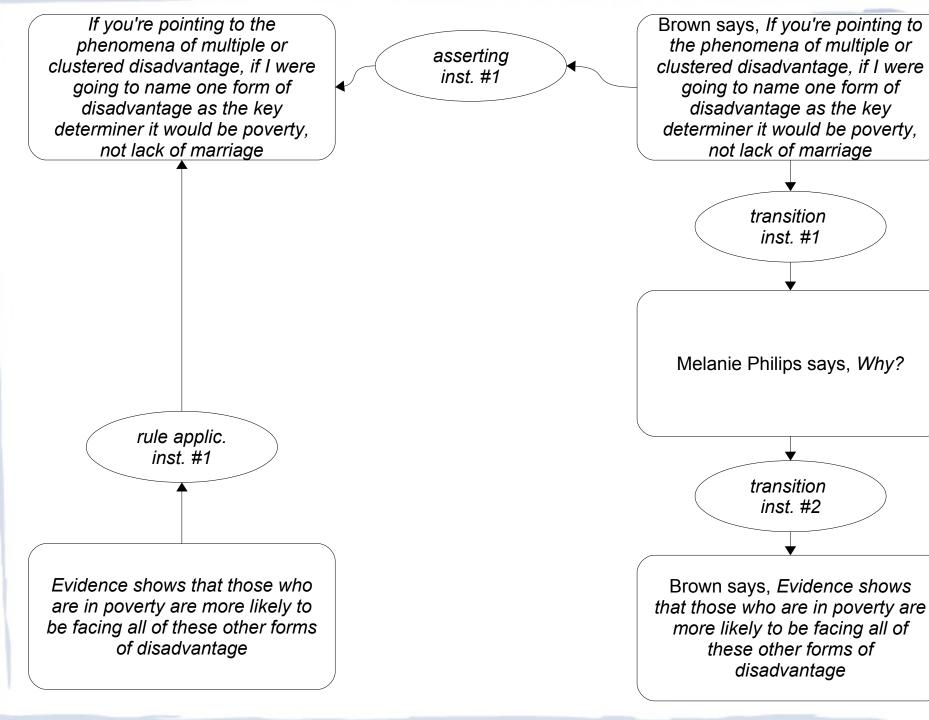
Brown says, If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage

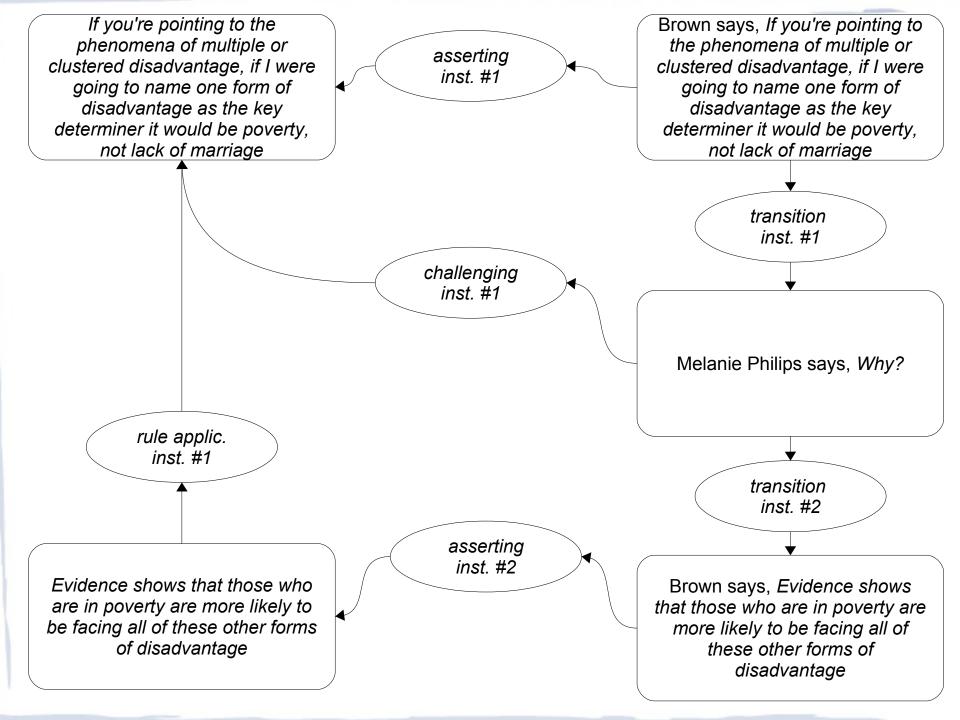
> transition inst. #1

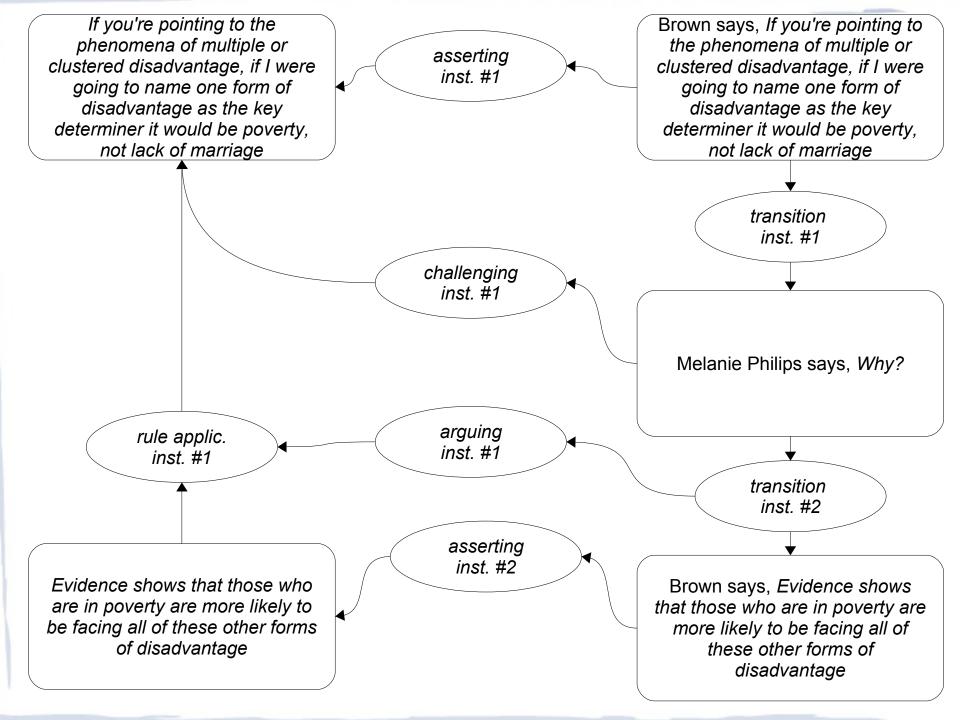
Melanie Philips says, Why?

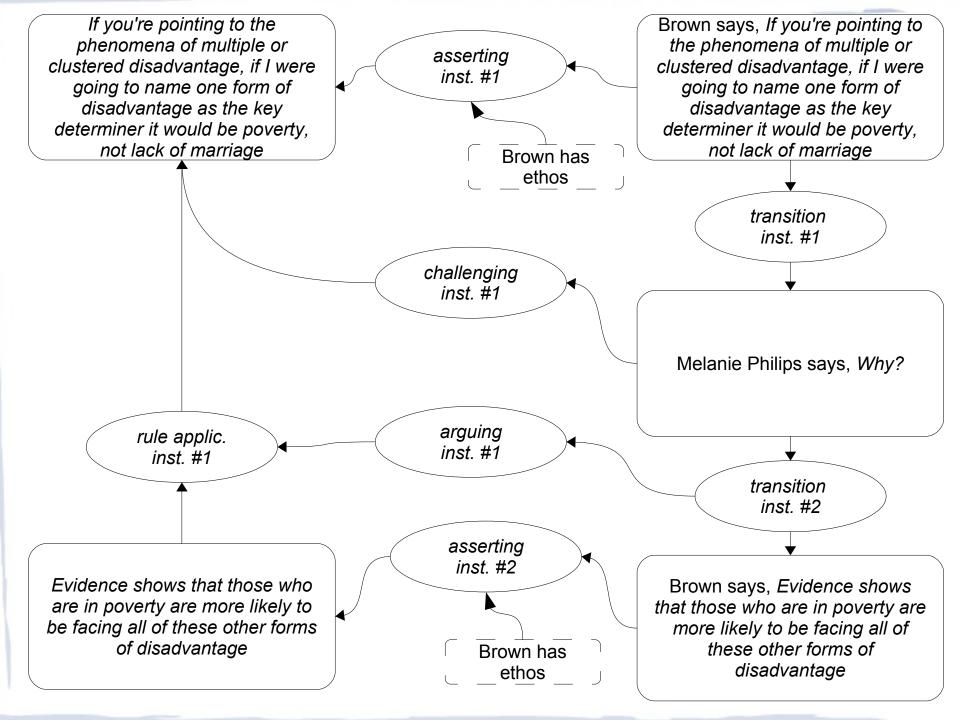
transition inst. #2

Brown says, Evidence shows that those who are in poverty are more likely to be facing all of these other forms of disadvantage









#### Inference Anchoring Theory, IAT

#### **DIALOGUE:**

- (a) Alexander Brown said, If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage.
- **(b)** Melanie Philips said, <u>Why?</u>
- **(c)** Alexander Brown said, *Evidence shows that those who are in poverty are more likely to be facing all of these other forms of disadvantage.*

#### ARGUMENTATION (INFERENCE):

(Arg1) If you're pointing to the phenomena of multiple or clustered disadvantage, if I were going to name one form of disadvantage as the key determiner it would be poverty, not lack of marriage because Evidence shows that those who are in poverty are more likely to be facing all of these other forms of disadvantage.

# (1b/4) Arguments and Dialogue: Technical Foundations (Lawrence et al. 2017b; Reed et al. 2017)



Treat it like any other text mining problem:

Annotate data, train classifier or build rules, tweak & repeat.

Moens, Boiy, Palau & Reed, 2007

#### Basic approach

Treat it like any other text mining problem:

Annotate <u>data</u>, train classifier or build rules, tweak & repeat.

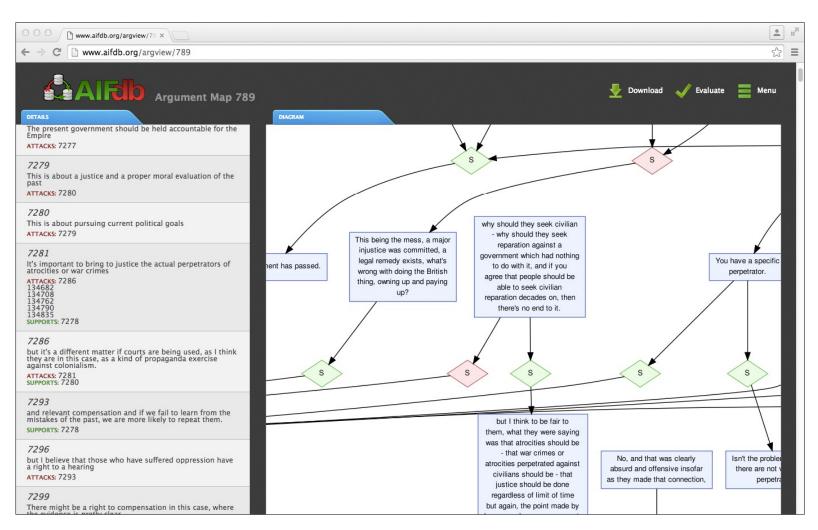
Moens, Boiy, Palau & Reed, 2007

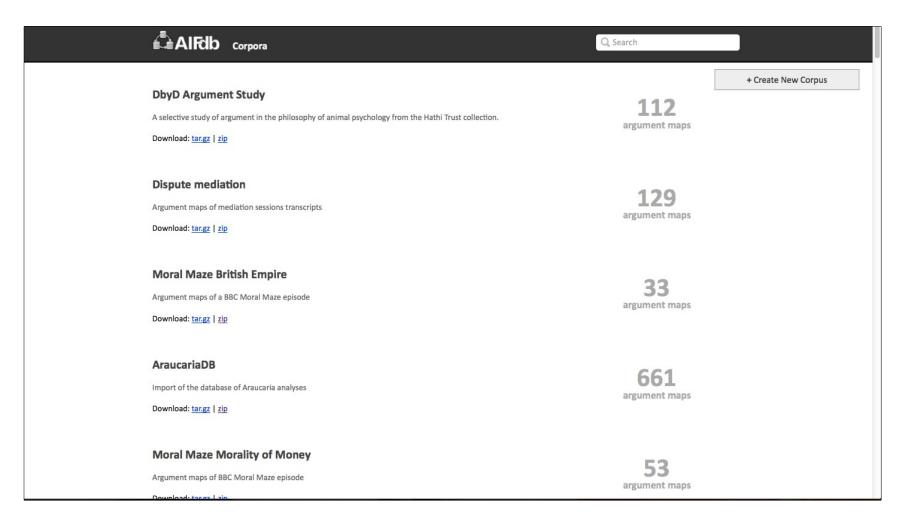
#### Basic approach

Treat it like any other text mining problem:

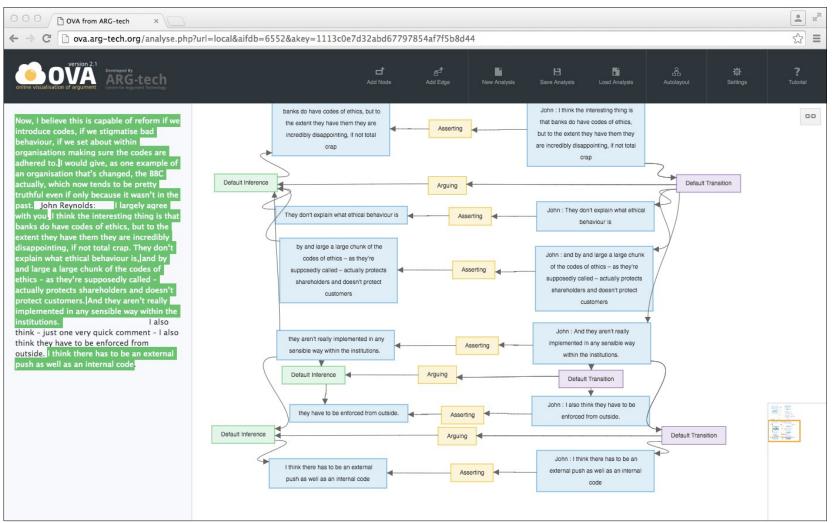
<u>Annotate</u> data, train classifier or build rules, tweak & repeat.

Moens, Boiy, Palau & Reed, 2007





corpora.aifdb.org



ova.arg.tech

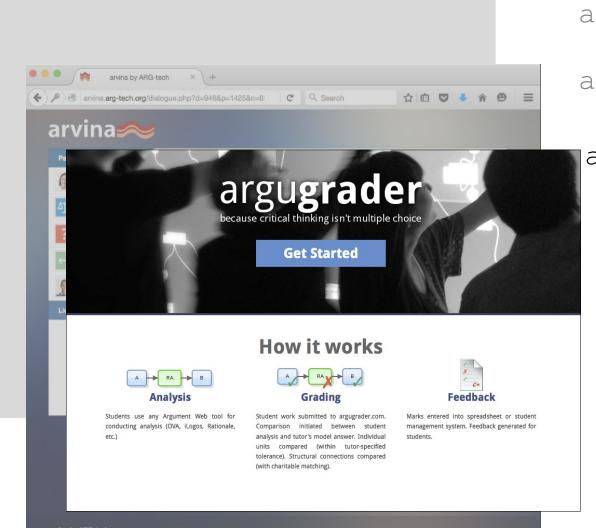
#### ArguBlo from ARG:dundee ArguBlogging allows you to agree and disagree with opinions expressed anywhere online. Your arguments are presented not just as text, but also as elements of the Argument Web: a richly structured interconnected web of debates and disputes, arguments and counterarguments. The arguments you contribute via argublogging can be used by other Argument Web tools to automatically express opinions, to teach students how to think better, to help the public navigate complex issues, and to conduct academic research into the way argument works. Try ArguBlogging as a guest Register for ArguBlogging Drag the guest bookmarklet to your bookmark toolbar: Give your email and a password so that we can store OAuth keys that allow argublogging to post to your blog ArguBlog This Email: chris@computing.dundee.ac Password: ....... Select some text on a web page, and click the ArguBlogging button to start argublogging and see the Register >>> result on the quest blog. Already registered? The application currently supports two popular blogging platforms: tumblr and Blogger. Publishing permissions are handled using OAuth, so there is never any need for you to provide us with your passwords - just grant the argublogging application access to your blog when prompted. When you encounter text online with which you want to agree or disagree, highlight it, and click the ArguBlogging button. After posting your opinion to your blog, you can also view it as a part of a conversation in the ever growing debate that is the Argument Web.

argublogging.com



argublogging.com

arvina.arg.tech



argublogging.com

arvina.arg.tech

argugrader.com



argublogging.com

arvina.arg.tech

argugrader.com

analytics.
arg-tech.org

#### Interoperably share & re-use:

- Language of Opposition Rutgers/Columbia
- Microtext corpus Potsdam
- Internet Argument Corpus (IAC) Santa Cruz
- Erulemaking corpus Cornell
- Araucaria Corpus Dundee
- Araucaria-PL Corpus Warsaw
- MM2012 Corpus Dundee
- plus over 30 other corpora

On the AIFdb servers at Dundee:

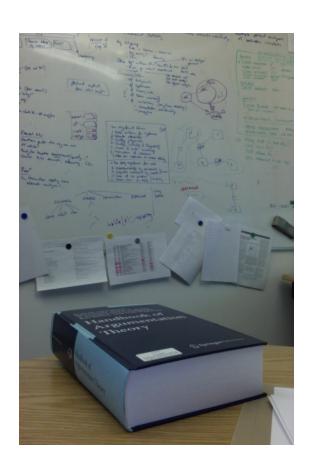
- 1.5m words
- 14 languages, 4 scripts
- 50k annotated arguments
- 300-500 unique users per month

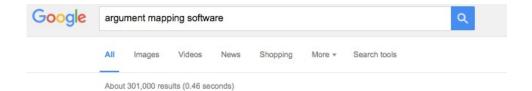
Frans H. van Eemeren Bart Garssen Erik C.W. Krabbe A. Francisca Snoeck Henkemans Bart Verheij Jean H.M. Wagemans

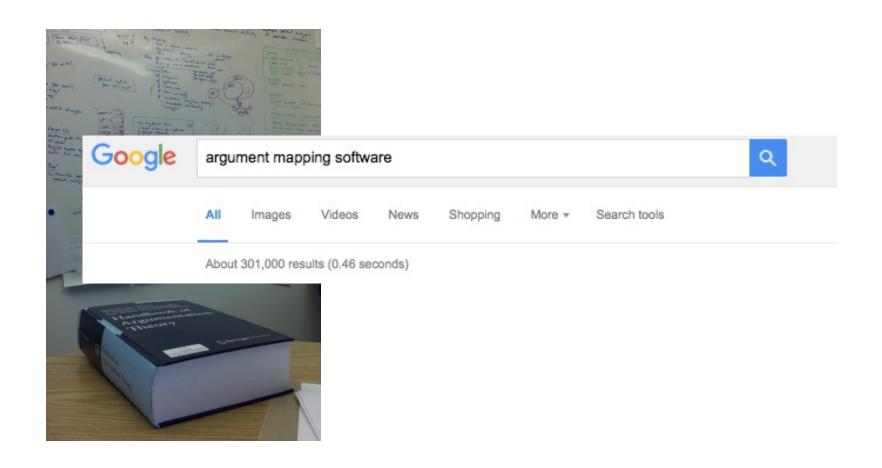
Handbook of Argumentation Theory

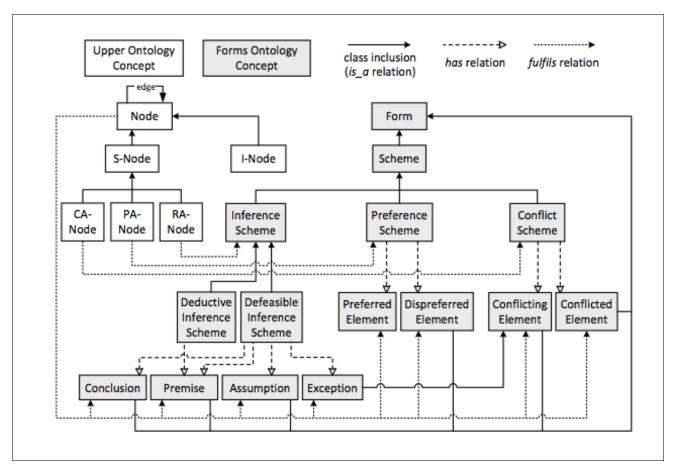








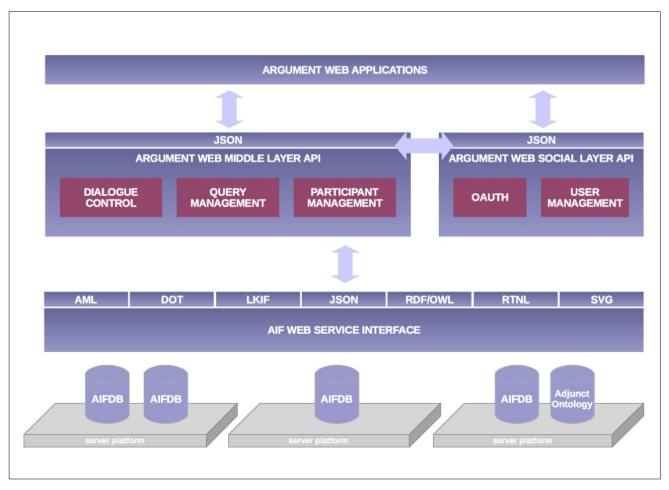




AIF: Argument Interchange Format Artificial Intelligence 2007

<b>Node Type</b>	Description
I	propositional information contained in an argument, such as a conclusion, premise, data etc.
L	subset of I-nodes referring to propositional reports specifically about discourse events
RA	application of a scheme of reasoning or inference
CA	application of a scheme of conflict
MA	application of a scheme of rephrasing
YA	application of a scheme of illocution describing communicative intentions which speakers use to introduce propositional contents
TA	application of a scheme of interaction or protocol describing relations between locutions

AIF and IAT



Argument Web **CACM** 2013

#### (2/4) Argument Structures for Corpus Analysis



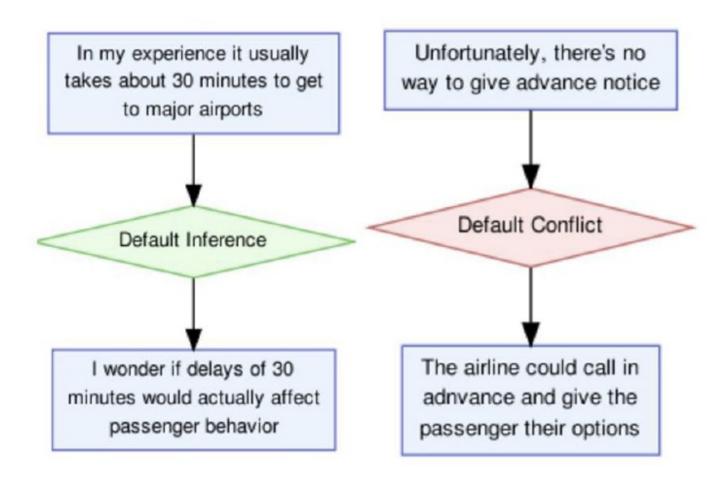
#### (2a/4) Basic Argument Structures



## Two types

- **Support** (argumentation, Default Inference, RA) holds between two propositions when one proposition provides a reason to accept another proposition. In other words, for a given claim p, a supporting claim q can be used to reply to the question "Why p?" ("Because q").
- **Conflict** (Attack, Default Conflict, CA) holds between two propositions which cannot be both true at the same time. Speakers use conflicting propositions to attack another speaker's claims by means of providing counter-claims.

## Two types of arguments



#### Example 1

Discussion Board in BB: Is technology destroying imagination?

### Argument map in OVA+ (ova.argtech.org)



# How are arguments signalled linguistically?

Is technology destroying imagination?

#### (1) Discourse indicators?

Is technology destroying imagination?

### (2) Topic similarity?

Is **technology** destroying imagination?

### (2) Topic similarity?

Is **technology** destroying imagination?

# (1+2) Discourse indicators + topic similarity?

Is **technology** destroying imagination?

### (3) Sentiment?

Is technology destroying imagination?

### (3) Sentiment?

Is technology destroying imagination?

Where should we go for holidays?

Warsaw has many great bars which are cheap.

Where should we go for holidays?

Warsaw has many great bars which are cheap.

- "We should go to Warsaw." "Warsaw has many great bars which are cheap." Support
- "We should go to Copenhagen." "Warsaw has many great bars which are cheap."

- "We should go to Warsaw." "Warsaw has many great bars which are cheap." Support
- "We should go to Copenhagen." "Warsaw has many great bars which are cheap." Attack

Is technology destroying imagination?

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

Is technology destroying imagination?

- 1. The internet can be used to inspire people.
- 2. Things can be seen, read about, watched, listened to at the user's convenience.
- 3. The internet can also be used to learn an entire spectrum of new information.
- 4. There is a valid point though that some websites do not offer many positive outcomes.

- "We should go to Warsaw." "Warsaw has many great bars which are cheap." Support
- "We should go to Copenhagen." "Warsaw has many great bars which are cheap." Attack

- "We should go to Warsaw." "Warsaw has many great bars which are cheap." Support
- "We should go to Copenhagen." "Warsaw has many great bars which are cheap." Attack

- "We should go to Warsaw." "Warsaw has many great bars which are cheap." Support
- "We should go to Copenhagen." "Warsaw has many great bars which are cheap." Attack
- "Warsaw has many great bars which are cheap." ???

# (4) Txt span or relation? (argument = txt span?)

"I love bananas"

# (4) Txt span or relation? (argument = txt span?)

- "What fruits do you like?" "I love bananas."
- "We should visit Philippines. I love bananas and in Philippines they grow amazing baby bananas, best in the world."
- "Chris hates bananas." "He loves them!"

# (4) Txt span or relation? (argument = txt span?)

- "What fruits do you like?" "I love bananas." Not Argument
- "We should visit Philippines. I love bananas and in Philippines they grow amazing baby bananas, best in the world." Support
- "Chris hates bananas." "He loves them!" Conflict

# (4) Txt span or relation? (If argument = relation, then...)

- Not only convergent, but also linked and serial arguments, or even unrelated text
- Long-distance relations

## (4) Txt span or relation? (If argument = relation, then...)

- Not only convergent, but also linked and serial arguments, or even unrelated text
- Long-distance relations

- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

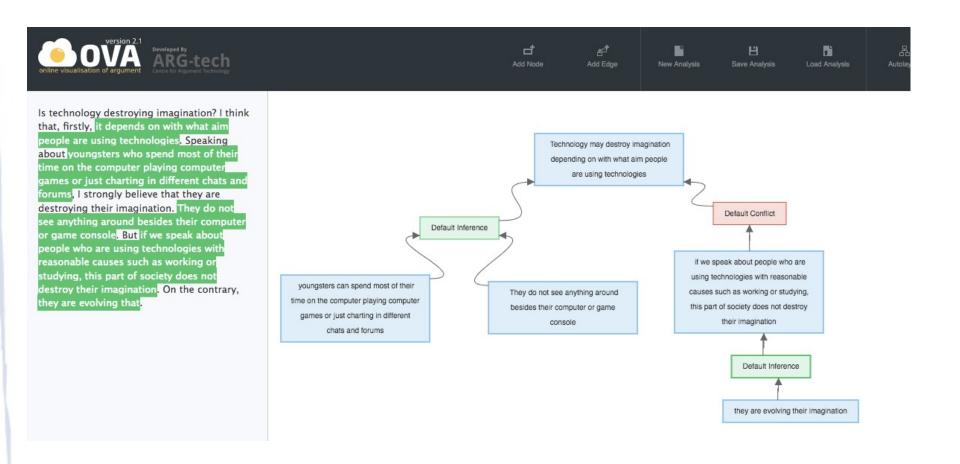
- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

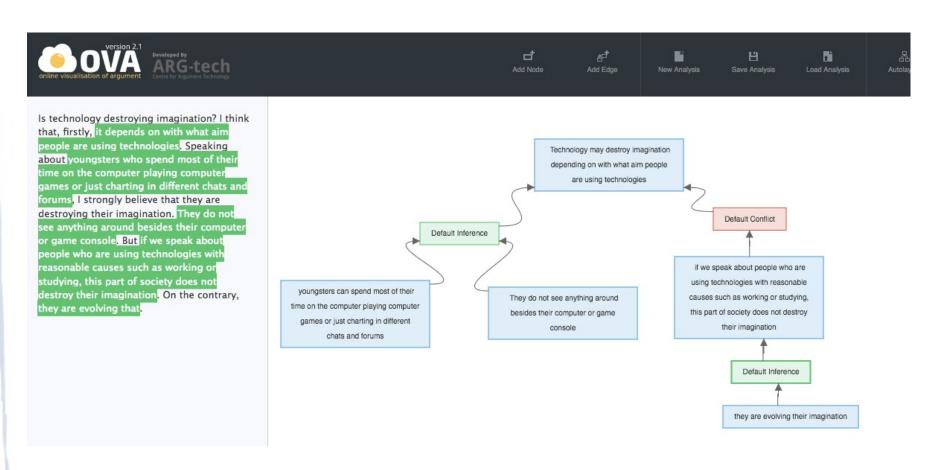
### (4) Analysis/SenseMaking/Summary: What the speaker is advocating?

- Speaking about youngsters who spend most of their time on the computer playing computer games or just charting in different chats and forums, I strongly believe that they are destroying their imagination
- 2. They do not see anything around besides their computer or game console
- 3. if we speak about people who are using technologies with reasonable causes such as working or studying, this part of society does not destroy their imagination
- 4. they are evolving that

### (4) Graph-based AIF argument map



### (4) Analysis/SenseMaking/Summary: What the speaker is advocating?



## (4) Txt span or relation? (If argument = relation, then...)

- Not only convergent, but also linked and serial arguments, or even unrelated text
- Long-distance relations

### (4) Txt span or relation? (If argument = relation, then...)

"Meta-talk in mediation" Corpus (Janier & Reed, 2016a; 2016b; 2017; Janier et al. 2016)

• http://corpora.aifdb.org/metatalk

• 2,862 ADUs, 446 supports, 73 conflicts

Eric: It's my team.

George: It's your team, exactly.

Mildred: Maybe.

George: There are obviously discussions that need to happen around the

team. Would that be a fair statement?

Viv: Yes.

Eric: Oh, I want Viv to be a productive member, but it's not for me to

actually

George: The team...

Mildred: Contribution to the team? Would that be a [Crosstalk 0:16:12].

Viv: [I would have assumed] they're my team. I mean, I thought we were equals.

### (4) Txt span or relation? (If argument = relation, then...)

Eric: <u>It's my team</u>.

George: It's your team, exactly.

Mildred: Maybe.

George: There are obviously discussions that need to happen around the

team. Would that be a fair statement?

Viv: Yes.

Eric: Oh, I want Viv to be a productive member, but it's not for me to

actually

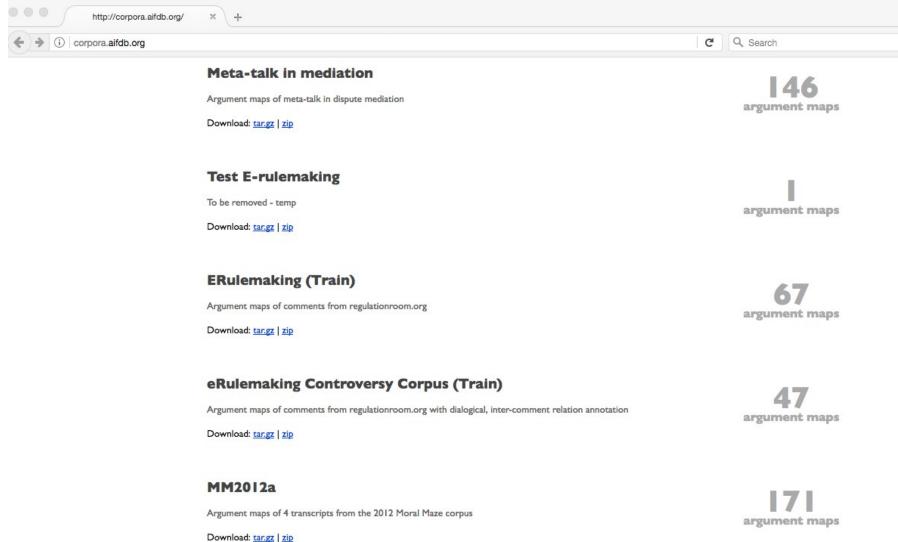
George: The team...

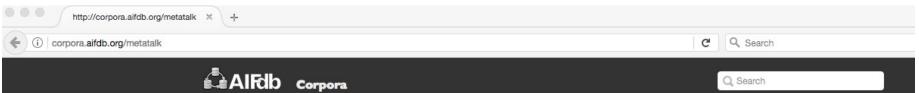
Mildred: Contribution to the team? Would that be a [Crosstalk 0:16:12].

Viv: [I would have assumed] they're my team. I mean, I thought we were

equals.

# (4) Txt span or relation? (If argument = relation, then...)





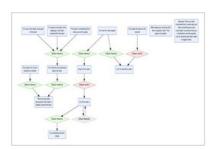
#### Meta-talk in mediation

Download full corpus: tar.gz | zip || View analytics

#### **Argument Map 4815**

Viv and Eric are concerned about the team Viv very much wants to be part of the team Viv wants to be seen to be adopting a role that's valued within the team The team is something Eric's built up over the years It's something that Eric values It's Eric's team There are obviously discussions...

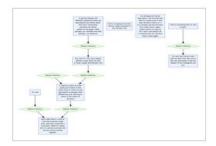
Download: SVG | PNG | DOT | ISON | LKIF | RTNL | RDF | PL Edit: OVA | OVA+



#### **Argument Map 4820**

Eric and Viv may have slightly different views about the role of Team Leader and Number Two In gaining George and Mildred's experience when job descriptions are written down, they don't necessarily translate into what's written on the paper. What perhaps you intended and what perhaps, you...

Download: SVG | PNG | DOT | JSON | LKIF | RTNL | RDF | PL Edit: OVA | OVA+



#### **Argument Map 4823**

xxx is how Viv and Eric communicate that and the aspects of the job Eric tries to speak to Viv Viv's assertive to Eric there's an example there, Viv just doesn't want to listen Viv, Eric, Mildred and George should put even something as general as communication and how that is achieved between...

Download: SVG | PNG | DOT | ISON | LKIF | RTNL | RDF | PL Edit: OVA | OVA+

