



**HAL**  
open science

## On Politics and Argumentation

Maria Boritchev

► **To cite this version:**

Maria Boritchev. On Politics and Argumentation. MALOTEC, Mar 2017, Nancy, France. hal-01666416

**HAL Id: hal-01666416**

**<https://hal.science/hal-01666416>**

Submitted on 18 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ON POLITICS AND ARGUMENTATION

---

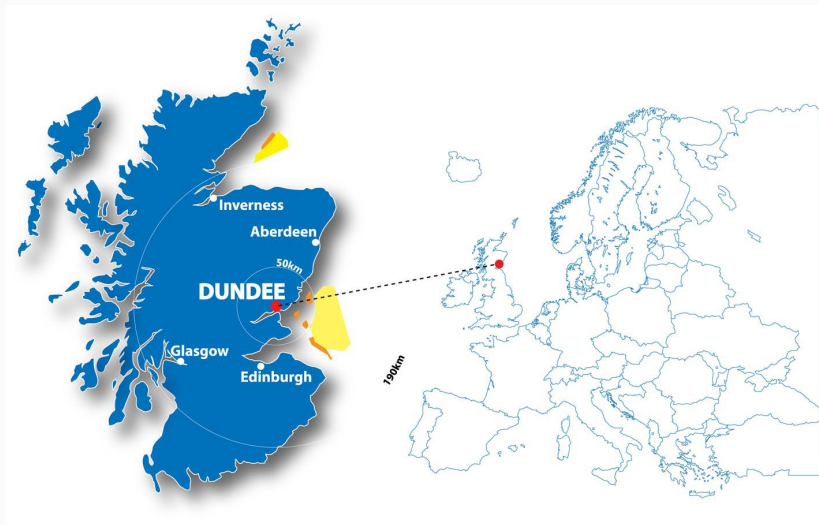
Maria Boritchev

March 15th, 2017

University of Lorraine

# Argument Mining

- Aim: identify argument structures that can be found within the discourse
  - Argumentative schemes
  - Relationships between pairs of arguments
- Introducing automation ~ 2000
- Applications:
  - Stock market
  - Computerized essay grading
  - Computer-supported peer review



## Projects

### Argument mining

- [Argument mining](#) provides algorithms for automatically extracting argument structures from monological text and from transcripts of dialogue
- The Centre for Argument Technology has won £700k from EPSRC for a project on [Argument Mining](#)
- The group ran and hosted the [Scottish Workshop on Argument Mining](#) 2014
- Members of the group have founded a series of workshops dedicated to Argument Mining, with editions at ACL in [2014](#), NAACL in [2015](#), and chaired by us in [2016](#) at ACL.
- Members of the group also co-delivered tutorials on Argument Mining at [IJCAI 2016](#) and [ACL 2016](#).

### Argument visualisation and analysis

- [AnalysisWall](#) – A very large touch screen supporting collaborative analysis in real time.
- [OVA: Online Visualisation of Argument](#) – A browser-based tool that supports analysis of online textual arguments and saves to the Argument Web. There are also two other tools in the suite: [OVAView](#), which renders AIF argument resources diagrammatically and [OVAgen](#) for manipulating abstract frameworks. [OVA+](#) supports analyses of dialogical arguments.
- [Araucaria](#) – A software tool for analysing arguments (also in a Polish version, [Araucaria-PL](#)) now largely replaced by OVA.

### Argument in dialogue

- The Dialogue Game Description Language, DGDL+, a language for describing the rules of dialogue games
- The Dialogue Game Execution Platform providing generalised execution services for arbitrary dialogue games

How to automatically mine intertextual correspondences between **U.S. candidates argumentation during primaries debates** and **Reddit users' comments** on those?

Some theory

Some politics

A spoonful of reality

Politics, people and argumentation

What now?

# SOME THEORY

---



## Example (A simple dialogue)

Alice: Q.

Bob: Why Q?

Alice: P.

## Modelisation (A simple dialogue)

$$\frac{P}{Q} (P \Rightarrow Q)$$

# Inference Anchoring Theory

- Aim: link dialogical and argumentative structures
- Combination of logic and argumentation theories
- Reference: Budzynska and Reed, 2011
- Contribution to the Argument Web
  - Mediation project
  - Ethos mining

### Example (Reddit comments)

Redditor1: Every American should be a capitalist.

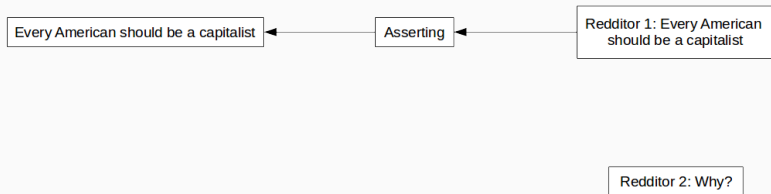
Redditor2: Why?

Redditor1: Our country was built on capitalism.

# IAT DIAGRAM (LOADING...)



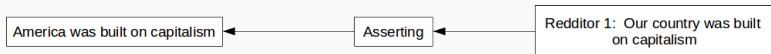
# IAT DIAGRAM (LOADING...)



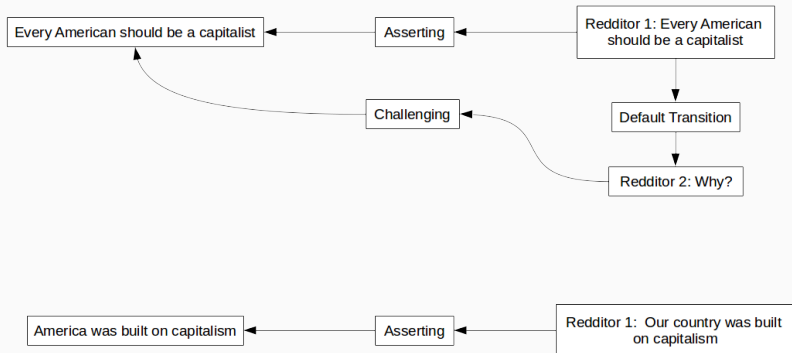
# IAT DIAGRAM (LOADING...)



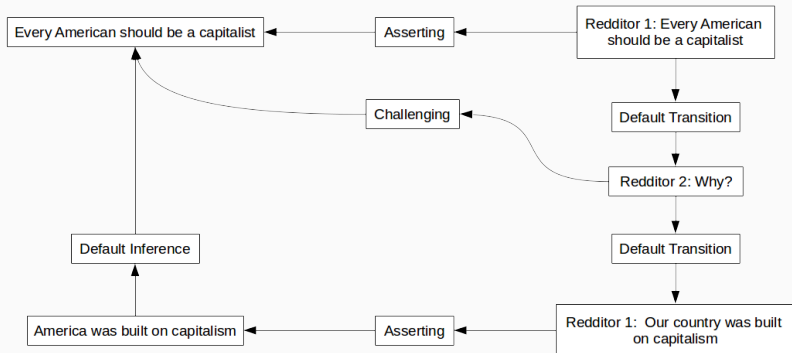
Redditor 2: Why?



# IAT DIAGRAM (LOADING...)

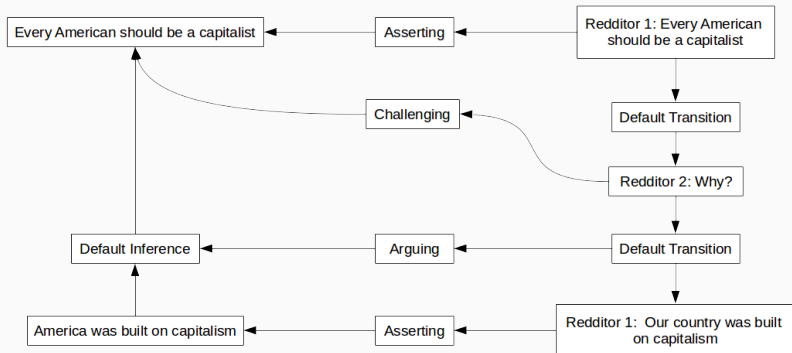


# IAT DIAGRAM (LOADING...)





# IAT DIAGRAM (LOADED!)



# SOME POLITICS

---

## U.S. presidential debates

- The American Presidency Project
- Web archive: American presidency related documents
- Transcripts of all presidential debates from 1960 to 2016
- 2016 elections:
  - 9 Democratic debates
  - 12 Republican debates
  - 3 General debates

## Reddit

- Social media and news aggregation, web content rating and discussion English-speaking website
- **subreddits**: news, science, politics, gaming, movies, ...
- In /r/politics, October 13 DNC Primary Debate - During-debate Discussion Megathread

[-] **Excusezy** 528 points 10 months ago  
"Pretty soft Granite..."  
Damn Cooper, you're not playing at all.  
permalink embed

[-] **LarkerLeo** 224 points 10 months ago  
Pretty good recovery, "the party left me!"  
permalink embed parent

[-] **mactownflameater** 85 points 10 months ago  
He's absolutely right  
permalink embed parent

[-] **bugcatcher\_billy** 48 points 10 months ago  
The man doesnt have much stage presence. But he has made excellent points and remained calm throughout.  
permalink embed parent

[-] **solmakou** 7 points 10 months ago  
left  
permalink parent

[-] **mactownflameater** 1 point 10 months ago  
Hahahaha  
permalink parent

1. For each debate, select the corresponding thread.
2. Sort the comments by time-stamp (oldest on top).
3. Remove all comments having no children.
4. Remove all comments trees beginning with irony or wordplay (rethoric structures are not handled by IAT).
5. Keep comments trees classified by excerpts (time-stamp identification), discard all others.

# A SPOONFUL OF REALITY

---

# MODEL


Moderator 

Participant<sub>1</sub> 

Moderator 

Participant<sub>2</sub> 

Participant<sub>1</sub> 

Participant<sub>2</sub> 

# IAT IN PRACTICE

The screenshot displays the OVA (Open Video Analysis) software interface. On the left, a transcript of a debate is shown with various segments highlighted in green. On the right, a network diagram maps the relationships between these segments, using nodes and arrows to represent logical connections like 'Asserting', 'Restating', and 'Default Transition'.

**OVA Version 1.1** | Developed by **ARG-tech**  
where visualization of arguments

Navigation icons: Add Nodes, Add Edge, New Analysis, Save Analysis, Load Analysis, Autolayout, Settings, Tutorial

**Transcript (Left):**

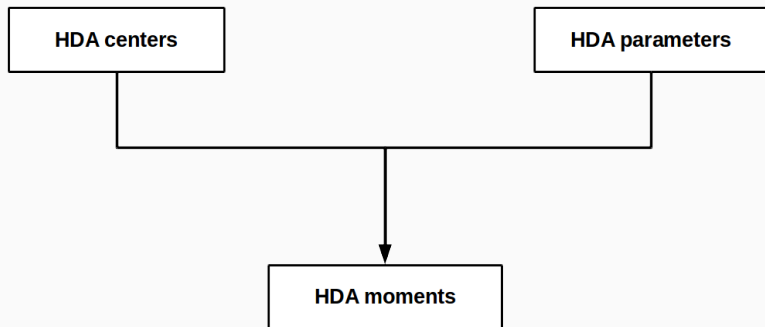
- everybody in on the question of electing Governor Chafee, you've been everything but socialist. When you were senator from Rhode Island, you were a Republican. When you were elected governor, you were an independent. You've only been a Democrat for little more than two years. Why should Democratic voters trust you won't change again?
- CHAFEE: Anderson, you're looking at a block of grants when it comes to the issues. Whether it's... [crosstalk]
- COOPER: It seems like pretty soft grants. I mean, you've been a Republican, you've been an independent.
- CHAFEE: Did you hear what I said? On the issues, I have not changed on the issues. I was a liberal Republican. That I was an independent and now I'm a proud Democrat. But I have not changed on the issues. And I open my record to scrutiny. Whether it's on the environment, a woman's right to choose, gay marriage, fiscal responsibility, aversion to foreign entanglements, using the tools of government to help the less fortunate. Time and time again, I have never changed. You're looking at a block of grants when it comes to the issues. So, I have not changed.
- COOPER: Then why change labels?
- CHAFEE: The party left me. There's no doubt about that. There was no room for a liberal moderate Republican in that party. I even had a primary for my reelection in 2006. I won it. But the money poured in to defeat me in Rhode Island as a Republican. That's what we were up against.

**Network Diagram (Right):**

The diagram illustrates the logical structure of the transcript. Nodes represent segments, and arrows represent relationships:

- CHAFEE: I have not changed on the issues** (Top Right) connects to **Asserting** and **Default Transition**.
- CHAFEE: I was a liberal Republican** (Middle Right) connects to **Restating** and **Default Transition**.
- CHAFEE: I was an independent** (Middle Right) connects to **Restating** and **Default Transition**.
- CHAFEE: now I'm a proud Democrat** (Middle Right) connects to **Disagreeing** and **Default Transition**.
- CHAFEE: I have not changed on the issues** (Bottom Right) connects to **Default Transition**.
- CHAFEE: I open my record to scrutiny** (Bottom Right) connects to **Default Transition**.
- CHAFEE has not changed on the issues** (Top Left) connects to **Asserting**.
- CHAFEE was a liberal Republican** (Middle Left) connects to **Default Rephrase**.
- CHAFEE was an independent** (Middle Left) connects to **Default Rephrase**.
- now CHAFEE is a proud Democrat** (Middle Left) connects to **Default Conflict**.
- CHAFEE opens his record to scrutiny** (Bottom Left) connects to **Default Inference**.
- Asserting** nodes connect to **Default Inference**.
- Restating** nodes connect to **Default Rephrase**.
- Disagreeing** nodes connect to **Default Conflict**.
- Default Transition** nodes connect to **Default Inference**.



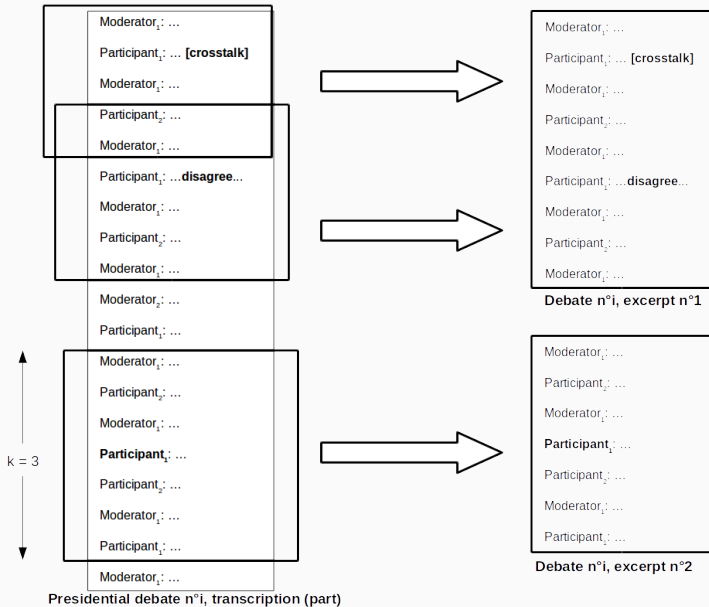


## Example

CHAFEE: Anderson, you're looking at a block of granite when it comes to the issues. Whether it's...*[crosstalk]*

COOPER: It seems like pretty soft granite. I mean, you've been a Republican, you've been an independent.

# HDA MOMENTS EXTRACTION



# POLITICS, PEOPLE AND ARGUMENTATION

---

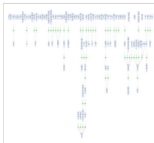
## US2016

Download full corpus: [tar.gz](#) | [zip](#) || View [analytics](#)

### Argument Map 10436

THANK YOU Chafee was the only Republican in the Senate to vote against the Iraq war it is good enough it 's the worst decision in American history That 's very significant if you 're going to make those poor judgment calls a critical time in our history we just finished with the Vietnam era getting...

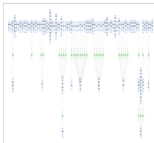
Download: [SVG](#) | [PNG](#) | [DOT](#) | [JSON](#) | [LKIF](#) | [RTNL](#) | [RDF](#) | [PL](#) Edit: [OVA](#) | [OVA+](#)



### Argument Map 10437

Clinton should have seen that attack coming let's remember what was going on we had a murderous dictator, Gadhafi, who had American blood on his hands Gadhafi was threatening to massacre large numbers of the Libyan people we had our closest allies in Europe burning up the phone lines begging us to...

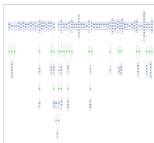
Download: [SVG](#) | [PNG](#) | [DOT](#) | [JSON](#) | [LKIF](#) | [RTNL](#) | [RDF](#) | [PL](#) Edit: [OVA](#) | [OVA+](#)



### Argument Map 10438

WEBB put the African American soldier on the Mall WEBB made that recommendation WEBB fought for it if you want someone who can stand up in front of you right now and say I have done the hard job, I have taken the risks, WEBB is your person wages and incomes are flat the gap between rich and poor is...

Download: [SVG](#) | [PNG](#) | [DOT](#) | [JSON](#) | [LKIF](#) | [RTNL](#) | [RDF](#) | [PL](#) Edit: [OVA](#) | [OVA+](#)

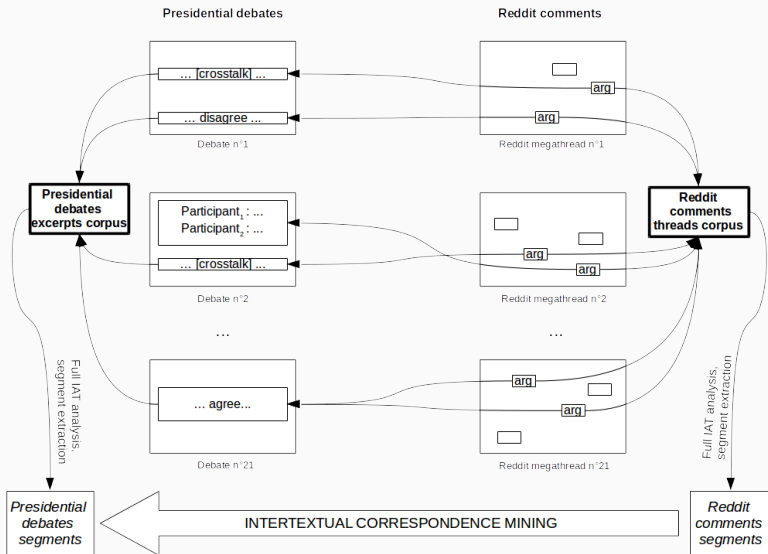


### Argument Map 10439

I Secretary Clinton , you have to be able to respond Sanders brought you up CLINTON respects the constitution integrity CLINTON represented Wall Street , as a senator from New York CLINTON went to



# CORRESPONDENCE MINING



## Input

$R = \{r_1, \dots, r_i, \dots, r_n\}$  set of Reddit comments segments.

$D = \{d_1, \dots, d_j, \dots, d_m\}$  set of presidential debates segments.

## Output

$C = (c_{i,j})$ , matrix of the correspondence coefficients between  $r_i$  and  $d_j$ .

Natural Language Processing tools:

- Speakers similarity
- Frequent wordsets similarity
- Semantic similarity

$$c_{i,j} = w_1 \cdot \text{Sp}(r_i, d_j) + \frac{w_2}{\text{length}(r_i)} \cdot \sum_{u \in r_i} \text{tf-idf}(u, d_j, D) \\ + \frac{w_3}{\text{length}(r_i) \cdot \text{length}(d_j)} \cdot \sum_{u \in r_i, v \in d_j} \text{Semsim}(u, v)$$



Natural Language Processing tools:

- **Speakers similarity**
- Frequent wordsets similarity
- Semantic similarity

$$c_{i,j} = w_1 \cdot \mathbf{Sp}(r_i, d_j) + \frac{w_2}{\text{length}(r_i)} \cdot \sum_{u \in r_i} \text{tf-idf}(u, d_j, D) \\ + \frac{w_3}{\text{length}(r_i) \cdot \text{length}(d_j)} \cdot \sum_{u \in r_i, v \in d_j} \text{Semsim}(u, v)$$

## Example

$r_i$  = “Hillary Clinton knows Bernie Sanders’ gun control record isn’t his strong suit”

$d_j$  = “CLINTON thinks what Senator SANDERS is saying certainly makes sense in the terms of the inequality that we have”

$S(t) = \{\text{CLINTON, SANDERS}\}$

Natural Language Processing tools:

- Speakers similarity
- Frequent wordsets similarity
- Semantic similarity

$$c_{i,j} = w_1 \cdot \text{Sp}(r_i, d_j) + \frac{w_2}{\text{length}(r_i)} \cdot \sum_{u \in r_i} \text{tf-idf}(u, d_j, D) \\ + \frac{w_3}{\text{length}(r_i) \cdot \text{length}(d_j)} \cdot \sum_{u \in r_i, v \in d_j} \text{Semsim}(u, v)$$

### Definition (Term frequency)

Let  $w$  be a word of a (non-empty) segment  $t$  in a corpus  $C$ . Term frequency of  $w$  in  $t$  is defined as

$$\text{Tf}(w, t) = \frac{|\{v \in t, v = w\}|}{|\{v \in t\}|}.$$

### Definition (Inverse document frequency)

Inverse document frequency of  $w$  in  $C$  is defined as

$$\text{Idf}(w, C) = \begin{cases} 0 & \text{if } w \notin t, \\ \log \frac{|s, s \in C|}{|s \in C, w \in s|} & \text{otherwise.} \end{cases}$$

## Definition (Term frequency–inverse document frequency)

Let  $w$  be a word of a segment  $t$  in a corpus  $C$ .

tf-idf of  $w$  in  $t$  in  $C$  is defined as

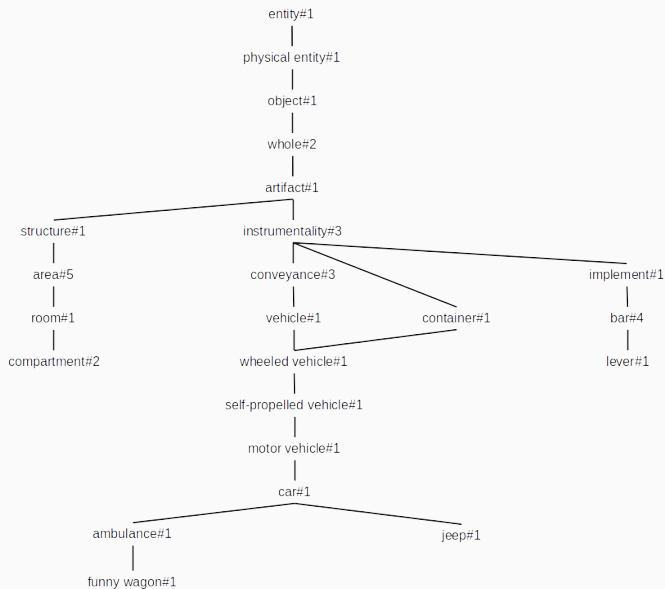
$$\text{tf-idf}(w, t, C) = \text{Tf}(w, t) \cdot \text{Idf}(w, C).$$

Natural Language Processing tools:

- Speakers similarity
- Frequent wordsets similarity
- Semantic similarity

$$c_{i,j} = w_1 \cdot \text{Sp}(r_i, d_j) + \frac{w_2}{\text{length}(r_i)} \cdot \sum_{u \in r_i} \text{tf-idf}(u, d_j, D) \\ + \frac{w_3}{\text{length}(r_i) \cdot \text{length}(d_j)} \cdot \sum_{u \in r_i, v \in d_j} \text{Semsim}(u, v)$$

# WORDNET HIERARCHY (PART.)



## Definition

Semantic similarity of  $w_1$  and  $w_2$  is defined as

$$\text{Semsim}(w_1, w_2) = 1 - \frac{\min_{i,j} \{\text{length}(\text{path}(w_1\#i, w_2\#j))\}}{\max_{v,k} \{\text{length}(\text{path}(\text{entity}, v\#k))\}}.$$



Natural Language Processing tools:

- Speakers similarity
- Frequent wordsets similarity
- Semantic similarity

$$c_{i,j} = w_1 \cdot \text{Sp}(r_i, d_j) + \frac{w_2}{\text{length}(r_i)} \cdot \sum_{u \in r_i} \text{tf-idf}(u, d_j, D) \\ + \frac{w_3}{\text{length}(r_i) \cdot \text{length}(d_j)} \cdot \sum_{u \in r_i, v \in d_j} \text{Semsim}(u, v)$$

WHAT NOW?

---

- Stating the **intertextual correspondence task**
- Developing the **corpora**
- Sketching the **intertextual correspondence algorithm**
  
- Future work
  - Other corpora?
  - Other methods?

**THANK YOU FOR YOUR ATTENTION**