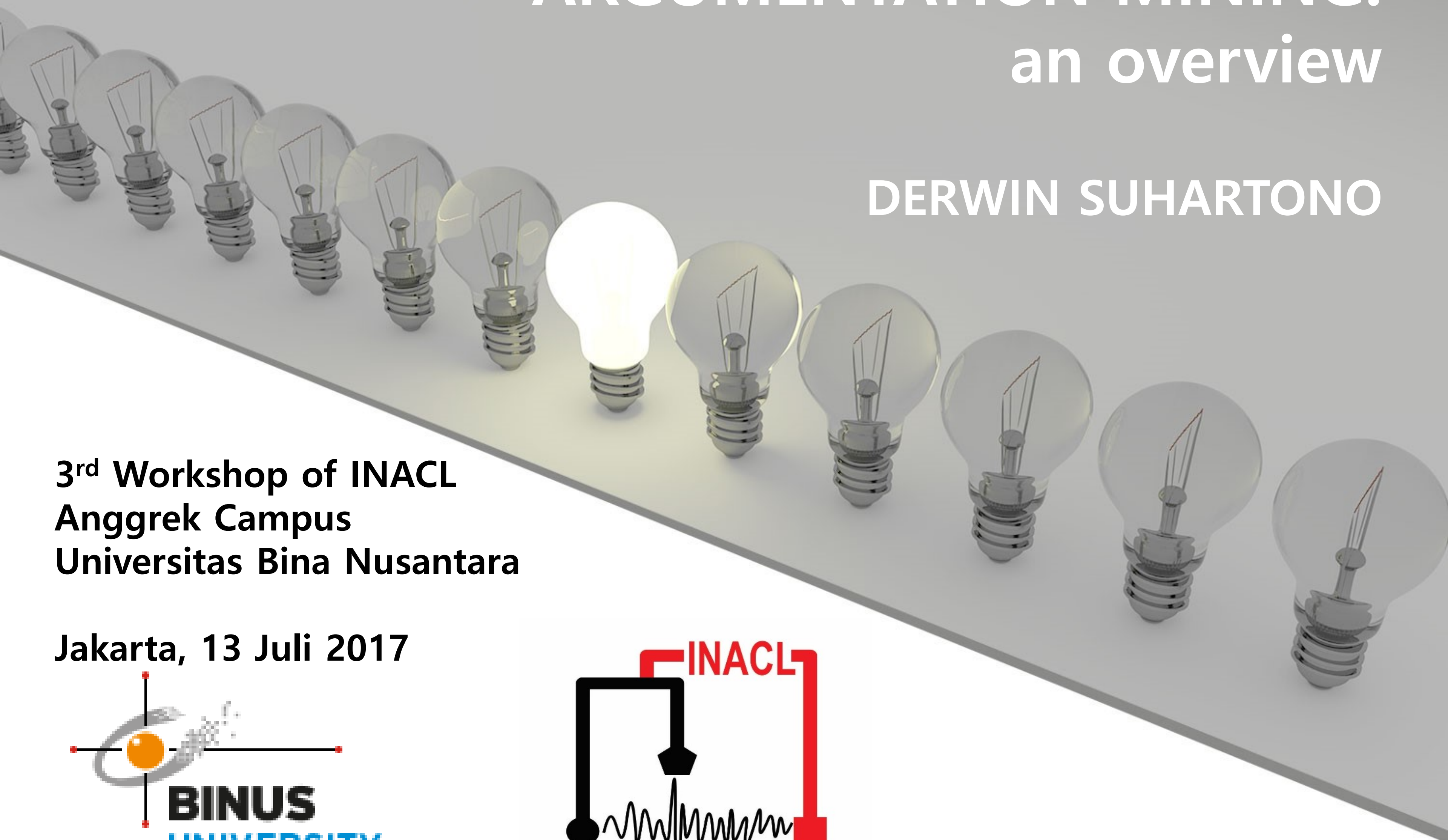
ARGUMENTATION MINING: an overview

DERWIN SUHARTONO



3rd Workshop of INACL
Anggrek Campus
Universitas Bina Nusantara

Jakarta, 13 Juli 2017

BINUS UNIVERSITY
School of Computer Science

INACL
Indonesian Association for Computational Linguistics

Today's Meal

- Argumentation
- Argumentation Annotation
- Argumentation Analysis
- some Experiments
- Tools
- Conference/Workshop

Argumentation

Argumentasi

alasan untuk memperkuat atau menolak suatu pendapat, pendirian, atau gagasan

Argumentation Mining

sebuah studi yang berfokus pada **ekstraksi** dan **analisa** argumen yang ada dalam natural language text.

Contoh data:

tulisan esai, komentar user pada sebuah blog/tulisan, naskah debat, artikel ilmiah dan lain sebagainya

Argumentation

Argumentation mining memiliki 2 (dua) level task yaitu:

- Argumentation annotation
- Argumentation analysis

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Hal yang paling fundamental untuk mengelola kalimat argumentasi adalah bagaimana **menemukan lokasi kalimat argumentasi** dalam kumpulan dokumen.

Beberapa pendekatan supervised machine learning sudah dilakukan dalam rangka memilah secara binary menjadi: **komponen argumentasi** dan **komponen non-argumentasi**.

Argumentation Annotation

Table 1: Examples of argumentative sentences per text type.

Text type	Argument	Non-argument
Discussion fora	"On this occasion, however, I shall not vote for any individual or party but will spoil my paper."	"I have been voting since 1964 and at one time worked for my chosen party."
Legal judgments	"He is aware of the risks involved, and he should bear the risks."	"Let there be any misunderstanding one point should be clarified at the outset. "
Newspapers	"Labor no longer needs the Liberals in the Upper House."	"The independents were a valuable sounding board for Labor's reform plans."
Parliamentary records	"I have accordingly disallowed the notice of question of privilege."	"Copies of the comments of the Ministers have already been made available to Dr. Raghuvansh Prasad Singh."
Weekly magazines	"But for anyone who visits Rajasthan's Baran district, the apathy of the district administration and the failure of the Public Distribution System (pds) is clear to see"	"This time in Rajasthan."

Pada awalnya,
anotasi yang dilakukan adalah mengklasifikasikan kalimat dalam dokumen menjadi 2 kategori:

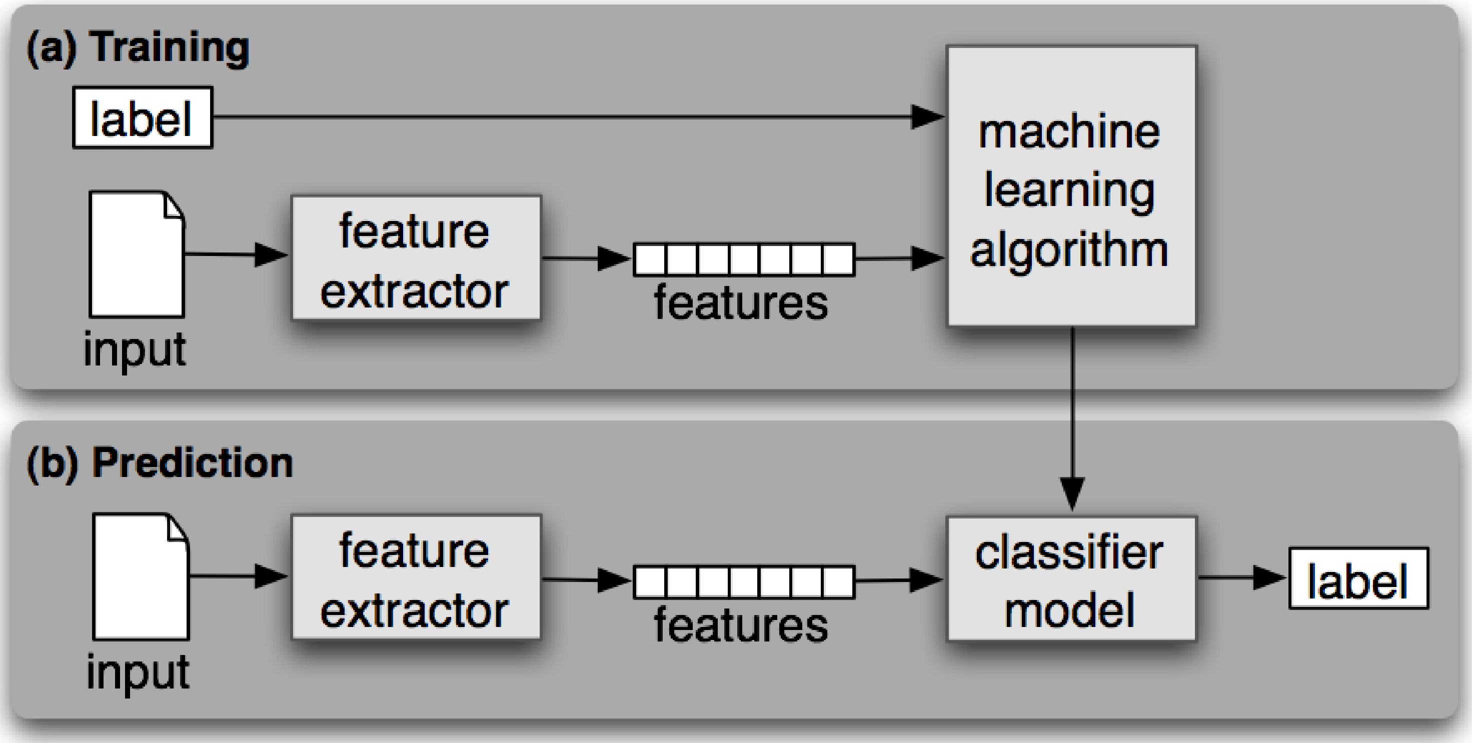
- kalimat argumentasi
- bukan kalimat argumentasi

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Feature Extraction



Argumentation Annotation

Feature Extraction

- Unigram
- Bigram
- Trigram
- Adverbs
- Verbs
- Modals Aux.
- Word Couples
- Text Statistics
- Punctuation
- Key Words
- Parse Features
- Tense and Mood

Table 1: Categories of morpho-syntactic features extracted from text segments.

Label	Description	Features
DM	Absolute number of occurrences of discourse markers from a given category	5 numerical features
Rel	Relative frequency of each of the 6 tenses and each of the 6 moods	12 numerical features
RCm	Relative frequency of each tense/mood combination (only for those that actually appear).	9 numerical features
Bin	Appearance of each of the 6 tenses and each of the 6 moods	12 binary features
Dom	Most frequent tense, mood, and tense/mood combination	3 string features
	TOTAL	41 features

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Shell language

Memanfaatkan <u>rule-based</u> bersama dengan <u>probabilistic sequence model</u>

MODAL → do | don't | can | cannot | will | would | . . .

ADVERB → strongly | totally | fundamentally | vehemently | . . .

AGREEVERB → disagree | agree | concur | . . .

AUTHORNOUN → writer | author | speaker | . . .

SHELL → I [MODAL] [ADVERB] AGREEVERB with the AUTHORNOUN

Figure 1: An example pattern that recognizes shell language describing the author's position with respect to an opponent's, e.g., *I totally agree with the author* or *I will strongly disagree with the speaker.*

- Rule-based : 25 pola hand-written regular expression.
- Manual annotation pada 170 esai
- Sequence model : Conditional Random Field (CRF) menggunakan sejumlah general feature berdasarkan frekuensi leksikal.

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Argument Ontology

- **Rule 1**: If the sentence begins with a Comparison discourse connective, or if the sentence contains any string prefixes from {conflict, oppose} and a four-digit number (intended as a year for a citation), then tag with **Opposes**.
- **Rule 2**: If the sentence begins with a Contingency connective and does not contain a four-digit number, then tag with **Supports**.
- **Rule 3**: If the sentence contains a four-digit number, then tag with **Citation**.
- **Rule 4**: If the sentence contains string prefixes from {suggest, evidence, shows, Essentially, indicate} (case-sensitive), then tag with **Claim**.
- **Rule 5**: .... Tag with **Hypothesis**
- **Rule 6**: .... Tag with **Hypothesis**
- **Rule 7**: .... Tag with **Current Study**
- **Rule 8**: .... Tag with **Opposes**

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Klasifikasi Argumentasi : Claim dan Evidence

- **Context Dependent Claim (CDC)**
  kalimat yang secara langsung mendukung atau menentang topik

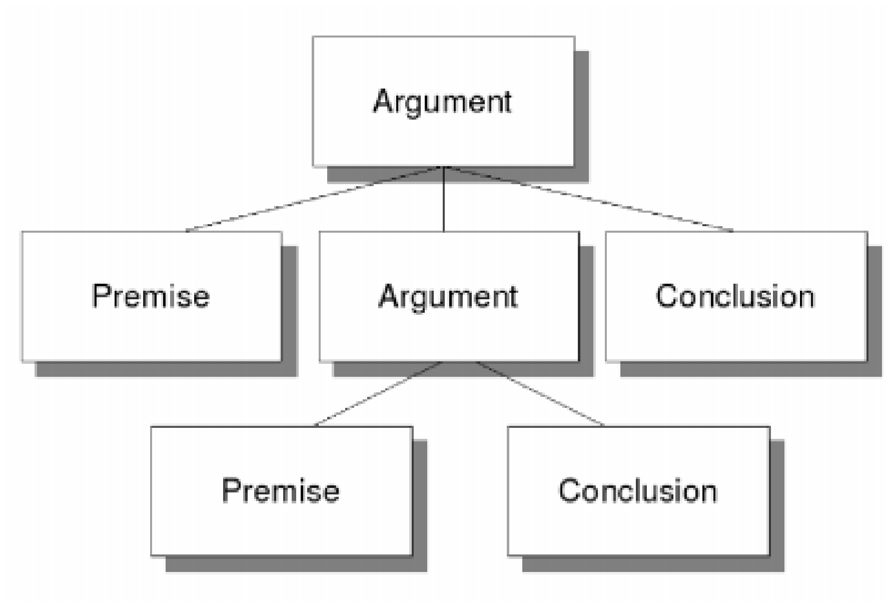
- **Context Dependent Evidence (CDE)**
  bagian dari teks yang secara langsung mendukung CDC dalam konteks topik yang diberikan

Contoh :
Topik : The sale of violent video games to minors should be banned
CDC : Violent video games increase youth violence
CDE :
The most recent large scale meta-anlysis-- examining 130 studies with over 130,000 subjects worldwide -- concluded that exposure to violent video games causes both short term and long term aggression in players

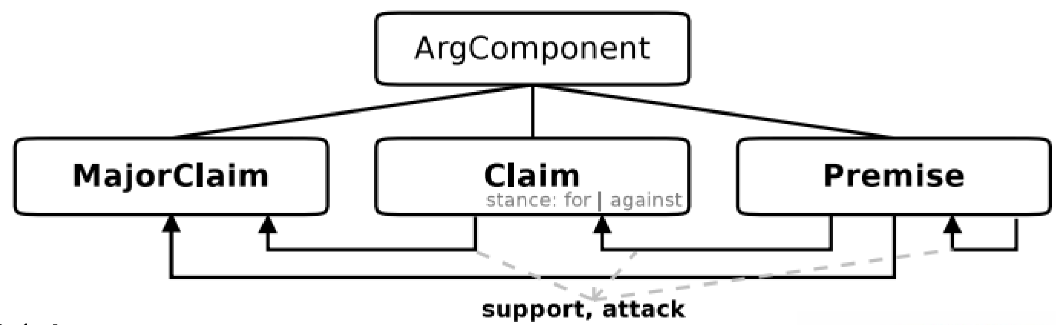
BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation



Palau dan Moens, 2009



Stab & Gurevych, 2014

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Klasifikasi Argumentasi:
Major Claim, Claim, Premise & Non-Argumen

- **Major Claim (MC)**
  "Newspapers have lost their competitive advantage to sustain their prolonged existence"
- **Claim (C)**
  "The print media has failed to keep its important role in the provision of information"
- **Premise (P)**
  "The internet has been more and more popular for recent years, providing people with a huge source of information"
- **None (N)**
  "As a result of this, print media such as newspapers have experienced a dramatic decline in the number of readers"

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Annotation

Klasifikasi Argumentasi:
Major Claim, Claim, Premise & Non-Argumen

- **Structural Features**
  "Newspapers have lost their competitive advantage to sustain their prolonged existence"
- **Claim (C)**
  "The print media has failed to keep its important role in the provision of information"
- **Premise (P)**
  "The internet has been more and more popular for recent years, providing people with a huge source of information"
- **None (N)**
  "As a result of this, print media such as newspapers have experienced a dramatic decline in the number of readers"

Argumentation Annotation

Klasifikasi Argumentasi:
Major Claim, Claim, Premise & Non-Argumen

- Structural Features
- Lexical Features
- Indicator Features
- Syntactic Features
- Prompt Similarity Features
  - kesamaan kalimat dengan topik dan dengan beberapa kalimat lainnya
- Word Embedding Features
  - menggunakan Glove sebagai feature
- Discourse Features
  - hubungan implisit dan eksplisit

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

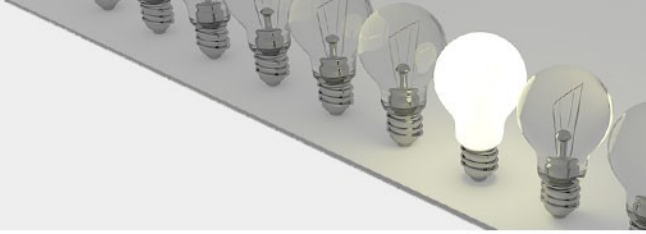
Argumentation Analysis

- Untuk menilai kualitas sebuah argumentasi, kita perlu melihat pada hal yang **secara intrinsik termuat** pada sebuah argumentasi.
- Hal ini tidak mudah untuk dilakukan, tidak seperti kategorisasi yang secara umum melihat langsung lebih banyak pada teksnya (ekstrinsik).
- **Discourse marker** yang bisa menjadi komponen utama dalam melihat argumentasi tidak bisa digunakan lagi pada penilaian kualitas argumen jika yang dilihat adalah hal-hal yang tidak terlihat dari luar.
- **Argumentasi yang baik** adalah argumentasi yang bisa meyakinkan pembaca bahwa argumen yang disampaikan adalah valid.

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis



Persuasiveness Level

Feature Category	Feature Name	Feature Description
Surface Text Features	length	# of the words, sentences and paragraphs in c.
	url	# of urls contained in c.
	unique # of words	# of unique words in c.
	punctuation	# of punctuation marks in c.
	unique # of POS	# of unique POS tags in c.
Social Interaction Features	tree_size	The tree size generated by c and rc.
	reply_num	The number of replies obtained by c and rc.
	tree_height	The height of the tree generated by by c and rc.
	Is_root_reply	Is c a root reply of the post?
	Is_leaf	Is c a leaf of the tree generated by rc?
	location	The position of c in the tree generated by rc.
Argumentation Related Features	connective words	Number of connective words in c.
	modal verbs	Number of modal verbs included in c.
	argumentative sentence	Number and percentage of argumentative sentences.
	argument relevance	Similarity with the original post and parent comment.
	argument originality	Maximum similarity with comments published earlier.

Table 2: Feature list (c: the comment; rc: the root comment of c.)

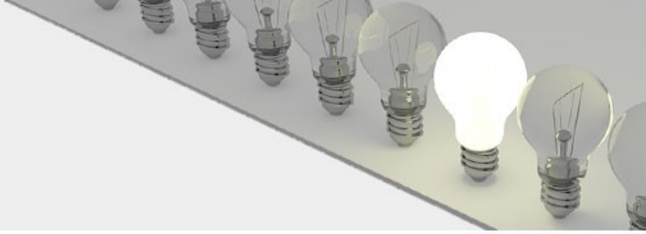
Metadata yang diperhatikan:

- Waktu posting
- Reputasi penulis

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis



Argument Strength

Melibatkan 1000 esai argumentatif untuk dilabelkan oleh human annotator

Anotator dipilih dari 30 applicant yang sudah familiar dengan pemberian skor rubrik. Mereka diberikan beberapa esai sample untuk di anotasi.

6 terbaik yang cukup konsisten dengan expected score dilibatkan dalam anotasi.

Score	Description of Argument Strength
4	essay makes a **strong argument** for its thesis and would convince most readers
3	essay makes a **decent argument** for its thesis and could convince some readers
2	essay makes a **weak argument** for its thesis or sometimes even **argues against it**
1	essay **does not make an argument** or it is often **unclear what the argument is**

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis

Argument Strength

Feature set yang digunakan:

1. POS N-grams (POS)
2. Semantic Frames (SFR)
3. Transitional Phrases (TRP)
4. Coreference (COR)
5. Prompt Agreement (PRA)
6. Argument Component Predictions (ACP)
7. Argument Errors (ARE)

http://www.hlt.utdallas.edu/~persingq/ICLE/acl15.pdf

BINUS UNIVERSITY School of Computer Science

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis

Argument Acceptability

Kombinasi antara **textual entailment** dan teori argumentasi

Example 1.
T1: Research shows that drivers speaking on
a mobile phone have much slower reactions i
n braking tests than non-users, and are wors
e even than if they have been drinking.
H:The use of cell-phones while driving is a p
ublic hazard.

Example 2 (Continued).
T2: Regulation could negate the safety benef
its of having a phone in the car. When you're
stuck in traffic, calling to say you'll be late ca
n reduce stress and make you less inclined to
drive aggressively to make up lost time.
H:The use of cell-phones while driving is a p
ublic hazard

Sistem yang hendak mengetahui TE harus bisa mendeteksi hubungan entailment antara T1 dan H (Example 1), dan kontradiksi antara T2 dan H (Example 2)

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis

Argument Acceptability

Kombinasi antara textual entailment dan **teori argumentasi**

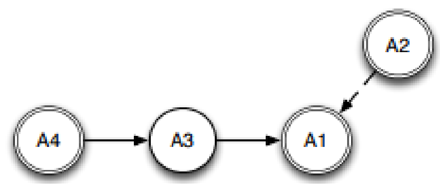


Figure 1: The AF built from the results of the TE module for Example 1, 2 and 3, without introducing additional attacks. Plain arrows represent *attacks*, dashed arrows represent *supports*.

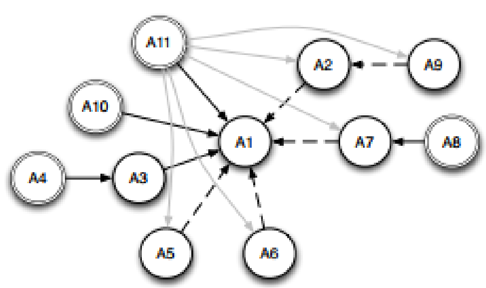


Figure 2: The AF built from the results of the TE module for the entire debate. Grey attacks are of type 1. For picture clarity, we introduce type 1 attacks only from A_{11}. The same attacks hold from A_{10} and A_3.

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis

Argumentation Sufficiency

Kriteria ini memisahkan antara argumen yang di *support* secara *sufficient* dari yang di *support* namun tidak *sufficient*.
Pengukurannya dilakukan dari kontribusi yang diberikan *premise* kepada *claim* pada argumen

"A full time job with a steady income was enough for a potential employee to support himself and his family. But with the recent recession, **almost all the countries all over the world are seeing significant economic downfall**. This overwhelming crisis is prompting delayering and reduction of workers. As a result, **people are focusing on doing several jobs** or gaining further qualification to secure their economic condition."

evidence yang diberikan dianggap cukup untuk memberikan korelasi diantara *recession* dengan *economic downfall*

Sufficient argument

BINUS UNIVERSITY School of Computer Science

INACL Indonesian Association for Computational Linguistics

Argumentation Analysis

Argumentation Sufficiency

"However, **using computer can make people easier to complete their work**. Businessmen, for example, use computer device to presentation or communication with their colleagues. Communication by feature-computer such as email and internet help busy businessmen to make deal with relation even in different country. This is due to using computer more efficient and, many workers cannot do more in work without computer."

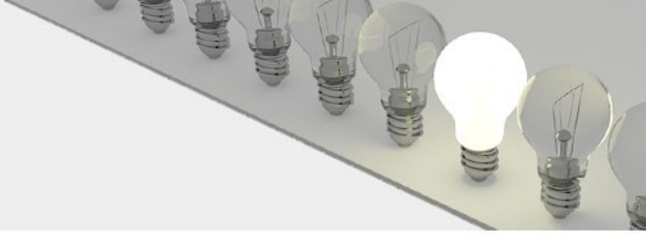
Argumen ini dianggap insufficient dikarenakan support yang diberikan oleh premis terlalu spesifik. Poin utama dari argumennya adalah "using computer can make people easier to complete their work".
Akan tetapi, businessman hanyalah bagian kecil dari society.

Insufficient argument

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some **Experiments**



1. Predefined Features vs. Word Vector Representation

1st Experiment

- Features which are extracted from the corpus consists of 4 (four) general categories; they are **structural, lexical, syntactic, and indicator**.
- **Contextual feature** is not yet implemented like Stab & Gurevych (2014b) did.
- Previously, they use 55 discourse markers from the Penn Discourse Treebank 2.0 Annotation Manual (Prasad et al., 2007) yet **we use 286 discourse markers** (Knott & Dale, 1993).

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some Experiments

1. Predefined Features vs. Word Vector Representation

2nd Experiment

- Almost of the features presented in the first experiment are implemented.
- N-gram feature is removed and it is replaced by **Glove word vector representation**. The vector is placed alongside with the other features.
- This experiment is made to measure how well the pre-trained word vector works as the features to classify argument components.

some Experiments

1. Predefined Features vs. Word Vector Representation

Settings:

- **Weka data mining software** is used to quantify the performan ce of the features.
- Testing category uses **10-fold** cross validation.
- **Support Vector Machine (SVM)** is used as the classifier.
- We have 3 different testing scenario for utilizing word vector representation. Each scenario differs in the **pre-trained word vectors dimensionality**; they are 50, 100 and 200.

some Experiments

1. Predefined Features vs. Word Vector Representation

TABLE 1 Comparison Accuracy using N-gram Versus Pre-Trained Word Vector

Features	Attributes	Accuracy
N-gram	313	73.9311
Glove 50 dim.	63	73.8717
Glove 100 dim.	113	73.9311
Glove 200 dim.	213	73.5154

TABLE 2 Comparison Accuracy using N-gram Versus Pre-Trained Word Vector After Filtered by Attribute Selection

Features	Attributes	Accuracy
N-gram	10	75.2969
Glove 50 dim.	8	74.7031
Glove 100 dim.	9	74.7031
Glove 200 dim.	12	74.9406

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some **Experiments**

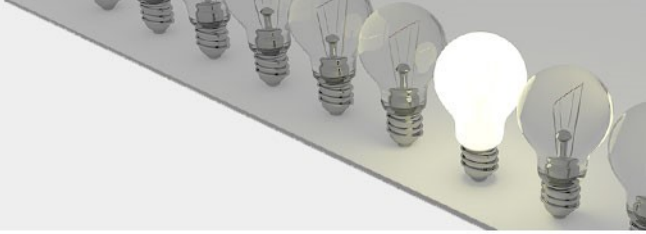
1. Predefined Features vs. Word Vector Representation

TABLE 3 Comparison Accuracy using N-gram Versus Pre-Trained Word Vector without Other Predefined Features

Features	Attributes	Accuracy
N-gram	300	52.9691
Glove 50 dim.	50	53.3254
Glove 100 dim.	100	53.2067
Glove 200 dim.	200	52.9691

TABLE 4 Comparison Accuracy using N-gram Versus Pre-Trained Word Vector without Other Predefined Features after filtered by Attribute Selection

Features	Attributes	Accuracy
N-gram	9	54.2162
Glove 50 dim.	23	53.4442
Glove 100 dim.	37	53.2067
Glove 200 dim.	57	52.9691



some Experiments

1. Predefined Features vs. Word Vector Representation

- If the confusion matrix is observed, we find out that the main issue is **to detect major claim (MC) properly**.
- From the whole scenarios above, **none of the major claim is well detected**. The majority of correct classification is the premise (P).
- On the other hand, **the accuracy to identify claim (C) is still very low.** Therefore, we still need to look for definitive features in detecting major claim (MC) and claim (C).

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some Experiments

2. Using LSTM (Long Short Term Memory)

- **Keras** (http://keras.io/) is used as the neural network library to implement our experiment
- In this experiment, we used **two (2) layer LSTM** (Long Short Term Memory) as one of variants in RNNs (Recurrent Neural Network).
- **50 dimensional word vectors** from Glove was used as the data in the input layer.
- Due to four (4) categories that we have, we use **categorical crossentropy** for compiling the model.
- We utilized two (2) kind of activation functions; they are **tanh and sigmoid** with additional dense layer.

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some Experiments

2. Using LSTM (Long Short Term Memory)

TABLE 5 Accuracy of Argument Component Classification using Categorical Crossentropy

LSTM Configuration	Activation Functions	
	Sigmoid	Tanh
1. using Dropout	Loss value: 2.6971 Accuracy: 54.80%	Val loss: 4.0191 Accuracy: 62.60%
2. without Dropout	Loss value: 3.95 Accuracy: 55.58%	Loss value: 13.1043 Accuracy: 12.47%
3. without Dropout, and initialize LSTM with Uniform	Loss value: 3.9253 Accuracy: 54.54%	Loss value: 12.987 Accuracy: 12.99%
4. without Dropout, and initialize dense layer with Uniform	Loss value: 4.6681 Accuracy: 57.40%	Loss value: 10.9687 Accuracy: 12.47%
5. without Droput, initialize LSTM with Uniform, and initialize dense layer with Uniform	Loss value: 4.882 Accuracy: 56.36%	Loss value: 10.9687 Accuracy: 19.48%

some Experiments

2. Using LSTM (Long Short Term Memory)

- As the additional scenario number 4, we changed the dense layer to be `glorotuniform` which got 1.1871 as the loss value and 57.66% as the accuracy.
- Compared with another result, we conclude **this is the best setting** so far. The progress of each iteration shows a good learning process from the model.
- We find out that **the experiment did not really carry out a good result**. For deep learning experiment, we guess that 90 essays are too small as the dataset. A larger size of data is recommended.

some Experiments

3. Combining All Features for Argument Component Detection

- Implementing 8 categories of features
  (68 sub-features in total)
- Support Vector Machine (SVM) as classifier
- 10-folds cross validation
- Utilizing corpus of 402 annotated persuasive essays by Stab and Gurevych (2016).

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some Experiments

3. Combining All Features for Argument Component Detection

Table 1. Previous works performance

Related Work	Accuracy
Palau and Moens (2007)	74%
Palau and Moens (2009)	74.04%
Stab and Gurevych (2014b)	77.3%
Lippi and Torroni (2015)	71.4%
Stab and Gurevych (2016)	77.3%
Wachsmuth, Al-Khatib, and Stein (2016)	74.5%
Habernal and Gurevych (2016)	75.4%
Al-Khatib et al. (2016)	66.8%

akurasi 79.96%

Table 2. Confusion matrix of the system accuracy results (SVM) for argument component classification

	MC	Cl	Pr	No
MC	**578**	130	43	0
Cl	226	**309**	970	1
Pr	28	147	**3656**	1
No	0	0	0	**1638**

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

some Experiments

3. Combining All Features for Argument Component Detection

- **Performance of each group of features**

Table 3. Previous works performance

Feature Name	Accuracy	Feature Name	Accuracy
Structural	77.83%	Syntactic	51.35%
Indicator	54.73%	Prompt Similarity	54.79%
Contextual	63.10%	Word Embedding	49.46%
Lexical	61.06%	Discourse	49.41%
All Features	**79.96%**		

Contextual and **lexical** features were the next significant features among all

some Experiments

4. Deep Learning Architecture (Argument Components Classification)

Model No.	Model Name	Units	Embedding	Batch Size	Accuracy (%)	Precision (%)	Recall (%)	F1 Macro (%)
1	**Baseline***	-	-	-	55.00	13.70	25.00	17.70
2	**Human***	-	-	-	87.70	86.40	87.90	87.10
3	**SVM***	-	-	-	77.30	77.30	68.40	72.60
4	**1D CNN**	-	Glove	64	56.88 ± 0.99	48.76 ± 1.47	47.84 ± 2.24	47.80 ± 1.73
5	**LSTM**	128	Glove	64	59.61 ± 1.30	52.43 ± 2.96	47.88 ± 1.72	49.20 ± 1.92
6	**GRU**	128	Glove	64	58.72 ± 1.50	52.06 ± 2.37	47.91 ± 1.91	49.05 ± 1.94
7	**Bidirectional LSTM**	128	Glove	64	59.26 ± 1.18	53.02 ± 2.37	48.14 ± 1.72	49.54 ± 1.65
8	**Bidirectional GRU**	128	Glove	64	58.28 ± 1.03	50.82 ± 2.01	47.58 ± 1.46	48.42 ± 1.36
9	**LSTM + Attention**	128	Glove	64	58.86 ± 1.60	52.46 ± 1.96	49.07 ± 1.95	50.02 ± 1.83
10	**GRU + Attention**	128	Glove	64	59.75 ± 1.79	53.40 ± 1.53	49.77 ± 2.38	50.78 ± 1.94
11	**Bidirectional LSTM + Attention**	128	Glove	64	59.98 ± 1.28	53.33 ± 2.13	50.19 ± 1.49	51.11 ± 1.39
12	**Bidirectional GRU + Attention**	128	Glove	64	59.69 ± 1.03	53.16 ± 1.90	49.83 ± 1.75	50.80 ± 1.35
13	**HAN**	64	Glove	64	57.26 ± 2.33	47.93 ± 13.15	37.46 ± 4.99	37.02 ± 7.65

some Experiments

4. Deep Learning Architecture combined with XGBoost (Argument Components Classification)

- Hasil yang terbaik untuk arsitektur *deep learning* di capai oleh Bidirectional LSTM dengan keikutsertaan mekanisme atensi (*attention mechanism*) di dalamnya.

- Dengan melihat pada tabel, dapat disimpulkan juga bahwa hampir semua model *Recurrent Neural Net work* (RNN) melebihi performa dari *Convolutional Neural Network* (CNN).

some **Experiments**

4. Deep Learning Architecture combined with XGBoost (Argument Components Classification)

Model No.	Feature Extractor	Units	Embedding	Accuracy (%)	Precision (%)	Recall (%)	F1 Macro (%)
1	1D CNN	-	Glove	77.08 ± 1.07	79.27 ± 1.06	66.51 ± 1.08	70.91 ± 1.01
2	LSTM	128	Glove	65.33 ± 1.53	61.46 ± 2.49	53.6 ± 2.02	56.23 ± 2.26
3	GRU	128	Glove	65.58 ± 1.34	61.43 ± 1.91	53.3 ± 1.53	55.86 ± 1.69
4	Bidirectional LSTM	128	Glove	64.91 ± 1.67	61.12 ± 2.58	52.88 ± 1.96	55.48 ± 2.25
5	Bidirectional GRU	128	Glove	65.21 ± 1.79	60.76 ± 2.7	52.99 ± 2	55.44 ± 2.18
6	LSTM + Attention	128	Glove	63.95 ± 1.37	58.43 ± 2.02	51.63 ± 2.15	53.79 ± 2.12
7	GRU + Attention	128	Glove	64.57 ± 1.25	59.65 ± 1.62	52.4 ± 1.93	54.75 ± 1.89
8	Bidirectional LSTM + Attention	128	Glove	65.85 ± 1.73	61.92 ± 2.45	53.88 ± 2.47	56.45 ± 2.47
9	Bidirectional GRU + Attention	128	Glove	66.47 ± 1.77	62.6 ± 2.7	54.89 ± 2.36	57.36 ± 2.46
10	HAN	64	Glove	69.52 ± 0.92	66.06 ± 1.42	59.92 ± 1.29	62.21 ± 1.25

Hasil terbaik diperoleh dengan menggunakan model ID CNN yang dikombinasikan dengan XGBoost

some Experiments

4. Deep Learning Architecture with and without XGBoost (Argumentation Sufficiency)

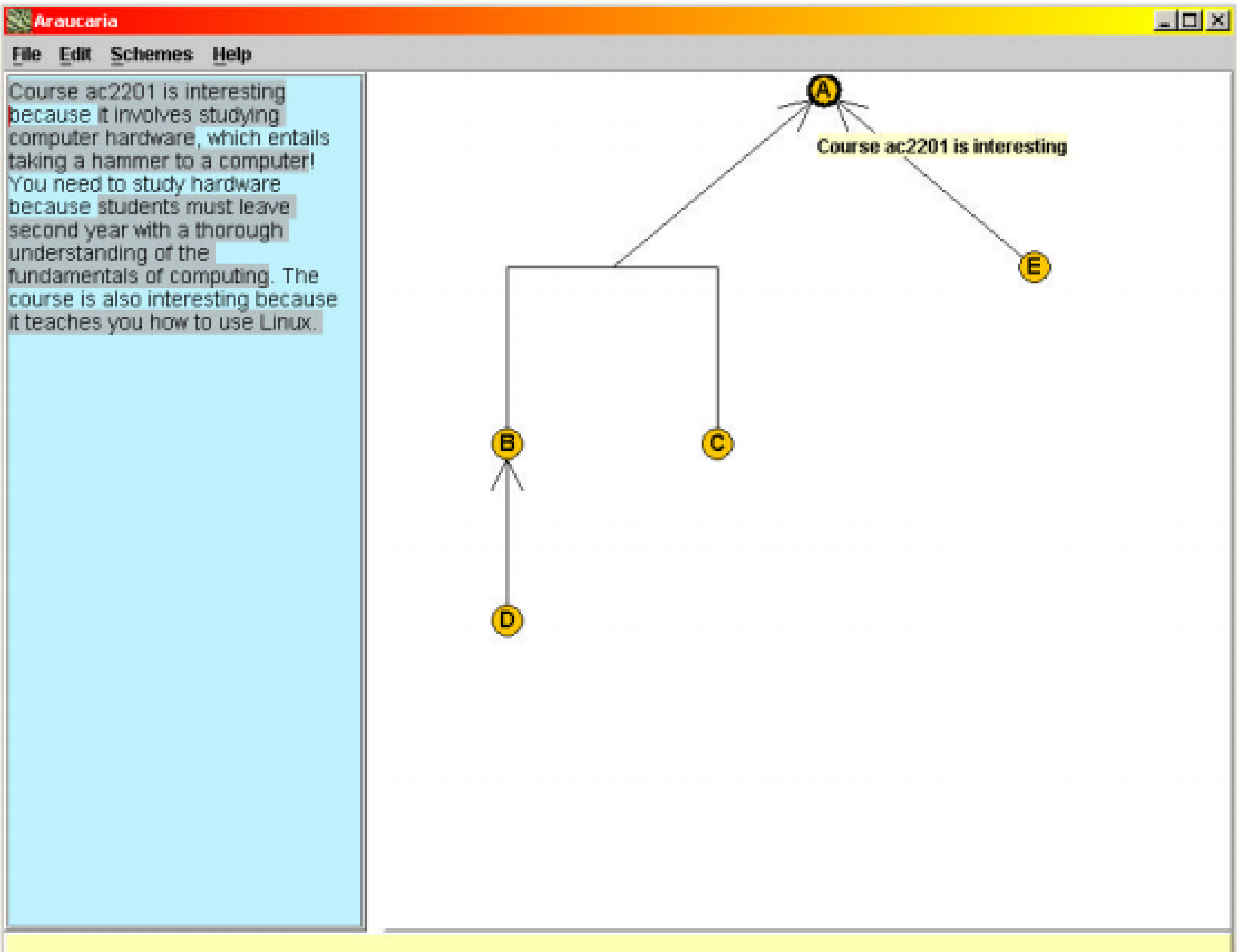
- Will be delivered today after this presentation by my student

13:00-14:00	**Keynote Speaker #4: Derwin Suhartono, S.Kom.,M.T.I.** Argumentation Mining	
14:00-15:00	**Presentasi Makalah #5**	**Indonesian Question Answering for Questions Containing Comparison with Structured Answer Source** Oleh: Athia Saelan, Ayu Purwarianti dan Dwi Hendratmo Widyantoro
		Hierarchical Attention Network with XGBoost for Recognizing Insufficiently Supported Argument Oleh: Derwin Suhartono, Aryo Pradipta Gema, Suhendro Winton, Theodorus David, Mohamad Ivan Fanany dan Aniati Murni Arymurthy
		Automatic Debate Text Summarization in Online Debate Forum Oleh: Alan Darmasaputra Chowanda, Albert Richard Sanyoto, Derwin Suhartono, Criscentia Jessica Setiadi

Tools untuk Pengolahan Data Argumentasi

REED & ROWE, ARAUCARIA:ARGUMENT DIAGRAMMING AND XML

diagram can be inverted at any time with a single key-press.



Araucaria (2004)

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Tools untuk Pengolahan Data Argumentasi

2 *D. Khartabil, S. Wells & J. Kennedy / Large-scale Argument Visualization (LSAV)*

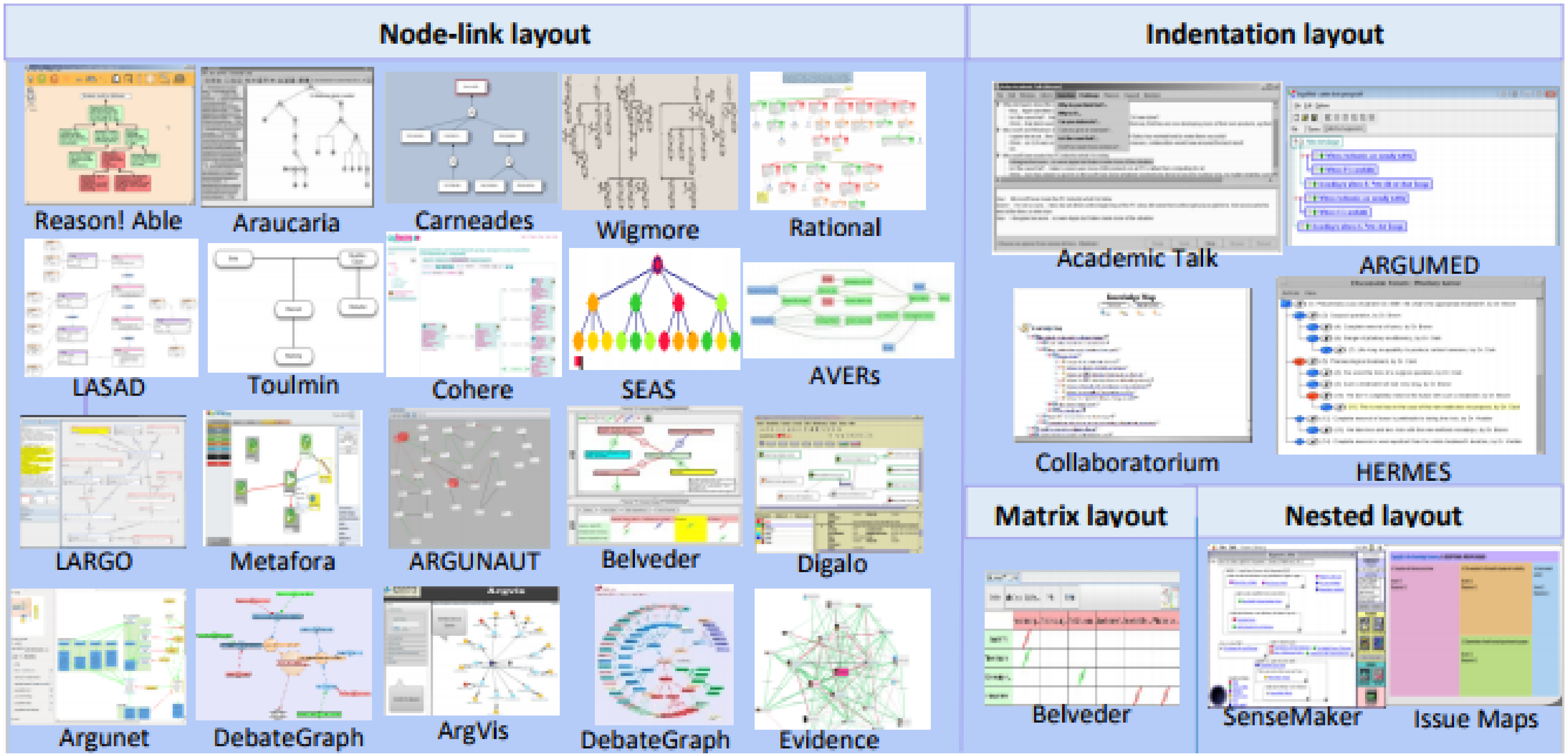


Figure 1: *Snapshots of existing argument visualization tools.*

BINUS UNIVERSITY School of Computer Scienc

Indonesian Association for Computational Linguistics

Tools untuk Pengolahan Data Argumentasi

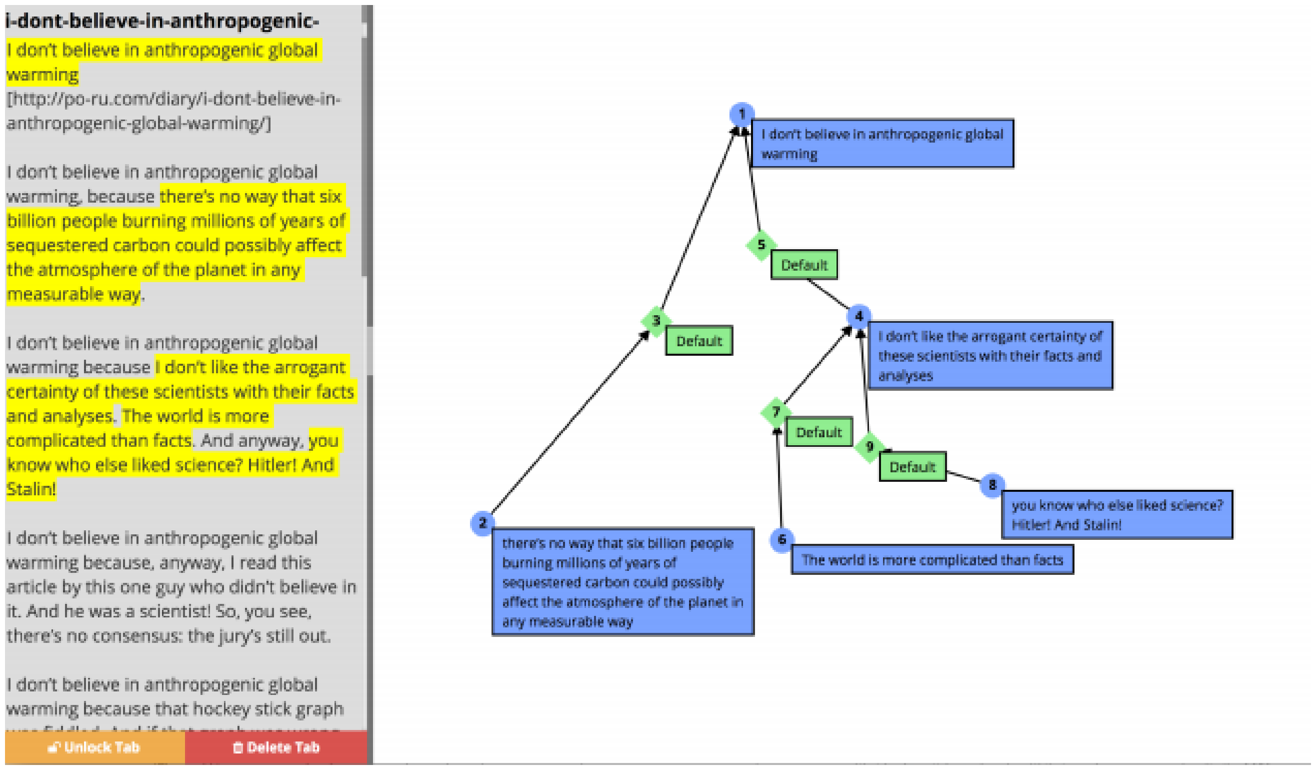


Figure 1: The default Monkeypuzzle User Interface showing the standard, two-pane UI popularised by Araucaria. The left-hand pane is the source pane, a tabbed collection of textual resources for analysis. The right-hand pane is the visualisation pane. The source pane can be completely collapsed to give a user more room to freely create an argument diagram independent of any specific source text allowing the app to be used for argument construction and exploration as well as argument analysis.

MonkeyPuzzle (2017)

BINUS UNIVERSITY School of Computer Scienc

INACL Indonesian Association for Computational Linguistics

Argumentation Mining Forum

Workshop on Argumentation Mining
- 2017 (4th) → Copenhagen, Denmark
- 2016 (3rd) → Berlin, Germany
- 2015 (2nd) → Denver, US
- 2014 (1st) → Baltimore, US

Workshop on Computational Models on Natural Argument (http://www.cmna.info/)
- 2017 (17th) → London, UK
- 2016 (16th) → New York, US
- 2015, 2014, … 2001

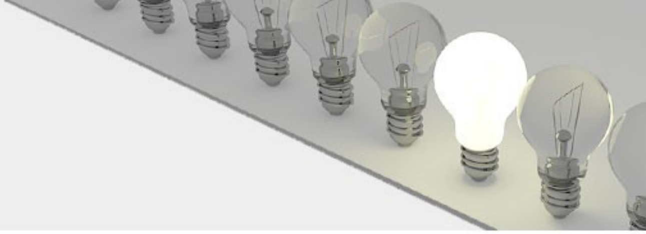
BINUS UNIVERSITY School of Computer Science

INACL Indonesian Association for Computational Linguistics

Pustaka

1. Florou, E., Konstantopoulos, S., Kukurikos, A., dan Karampiperis, P. (2013). Argument Extraction for Supporting Public Policy Formulation. Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities, pages 49-54, Sofia, Bulgaria, August 8.
2. Madnani, N., Heilman, M., Tetreault, J., dan Chodorow, M. (2012). Identifying High-Level Organizational Elements in Argumentative Discourse. Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 20-28, Montreal, Canada, June 3-8.
3. Ong, N., Litman, D., dan Brusilovsky, A. (2014). Ontology-Based Argument Mining and Automatic Essay Scoring. Proceedings of the First Workshop on Argumentation Mining, pages 24-28, Baltimore, Maryland USA, June 26.
4. Palau, R.M., dan M.F. Moens (2009). Argumentation Mining: The Detection, Classification and Structure of Arguments in Text. The 12th International Conf erence on Artificial Intelligence and Law, Barcelona.
5. Stab, C., dan Gurevych, I. (2014a). Annotating Argument Components and Relations in Persuasive Essays. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pages 1501-1510, Dublin, Ireland, August 23-29.

Pustaka



6. Stab, C., dan Gurevych, I. (2014b). Identifying Argumentative Discourse Structures in Persuasive Essays. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 46-56, October 25-29, Doha, Qatar.
7. Stab, C., & Gurevych, I. (2017, April). Recognizing Insufficiently Supported Arguments in Argumentative Essays. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2017), p. 980-990.
8. Reed, C., & Rowe, G. (2004). Araucaria: Software for argument analysis, diagramming and representation. International Journal on Artificial Intelligence Tools, 13(04), 961-979.
9. Wei, Z., Liu, Y., & Li, Y. (2016, August). Is This Post Persuasive? Ranking Argumentative Comments in the Online Forum. The 54th Annual Meeting of the Association for Computational Linguistics (p. 195).
10. Persing, I., & Ng, V. (2015, September). Modeling Argument Strength in Student Essay s. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp. 543-552.
11. Cabrio, E., & Villata, S. (2012, July). Combining textual entailment and argumentation theory for supporting online debates interactions. In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2 (pp. 208-212). Association for Computational Linguistics.

Pustaka

12. Derwin Suhartono, Afif Akbar Iskandar, M. Ivan Fanany, Ruli Manurung. Utilizing Word Vector Representation for Classifying Argument Components in Persuasive Essays. 3rd International Conference on Science, Engineering, Built Environment, and Social Science (ICSEBS), Bandung, Indonesia, 28 November – 1 December 2016. (ongoing to be published in Pertanika Journal of Science and Technology, July 2017)
13. Yunda Desilia, Velizya Thasya Utami, Cecilia Arta, Derwin Suhartono. An attempt to combine features in classifying argument components in persuasive essays. 17th Workshop on Computational Models of Natural Argument (CMNA) in conjunction with 16th International Conference on Artificial Intelligence and Law (ICAIL), London, United Kingdom, 16 June 2017
14. Derwin Suhartono, Aryo Pradipta Gema, Suhendro Winton, Theodorus David, Mohamad Ivan Fanany, Aniati Murni Arymurthy. Hierarchical Attention Network with XGBoost for Recognizing Insufficiently Supported Argument. (to be submitted)
15. Derwin Suhartono, Aryo Pradipta Gema, Suhendro Winton, Theodorus David, Mohamad Ivan Fanany, Aniati Murni Arymurthy. Comparative Analysis of Deep Learning Techniques in Argumentation Mining Tasks. (to be submitted)

Terima kasih

Derwin Suhartono
dsuhartono@binus.edu
+6281288004495