From argument diagrams to argumentation mining in texts: a survey

Andreas Peldszus and Manfred Stede
Universit¨at Potsdam
EB Cognitive Science

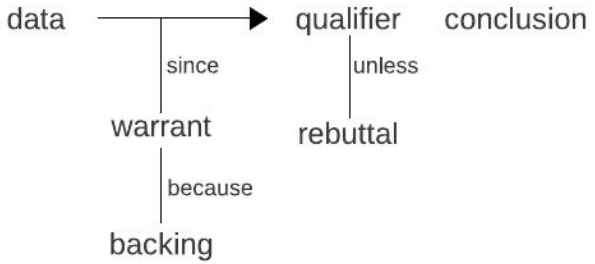
Table of contents

1 Introduction: analyzing argumentative text

- **Argumentation**: the process of conveying inclinations, attitudes or opinions, and trying to make the partner accept them - or even adopt them.
  - “real” text is often not particularly crisp and clean
- **Argument mining**: the automatic discovery of an argumentative text portion and the identification of the relevant components of the argument presented.
- Motivation ?
  - enrich the retrieval capabilities of legal databases ;
  - opinion mining : *reasons* those users give for their opinions, inclinations, or decisions ;
  - public deliberation : assessing public opinion on political questions
  - automatic persuasive text generation

2 Annotation schemes for argumentation

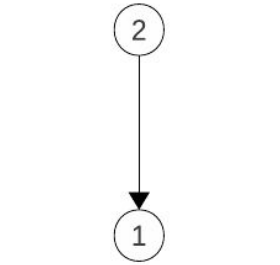
1. Toulimen(1958) : 6 functional roles, tree structure
2. Freeman(1991) :
    a. goal : integrate Toulmin's ideas into the argument diagraming
    b. new version 2011



(a) Toulmin

- CC central claim - macrostructure of argumentats
- basic dialectical structure : finding the corresponding critical question of the challenger that is answered by a particular segment of the text
- arguments : process(philosophy, rhetoric) and product(proposition)

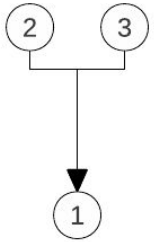
2 Annotation scheme - support



(a) basic argument

1 conclusion - 1 premise

[We should tear the building down.]1 [It is full of asbestos.]2

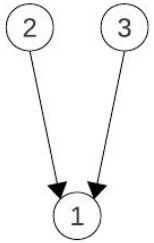


(b) linked support

1 conclusion - 2 related premises

[We should tear the building down.]1 [It is full of asbestos.]2 [All buildings with hazardous materials should be demolished.]3

2 Annotation scheme - support



(c) multiple support

- semantically independent, **additional argument** for the same conclusion
  - dialectical level : ”Can you give me an additional argument for that conclusion?”

1 conclusion - 2 separate pises

[We should tear the building down.]1 [It is full of asbestos.]2 [Also, people in the neighborhood have always hated it.]3

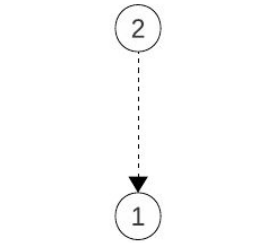
(d) serial support

- a new argument to convince the reader of the **acceptability** of a premise
  - dialectical level : ”Why should I accept that premise?”
  - premise evaluation : if one is unacceptable, the other one too (semantically interconnected)

1 conclusion - 2 independent premises

[We should tear the building down.]1 [It is full of asbestos.]2 [The commission reported a significant contamination.]3

2 Annotation scheme - support



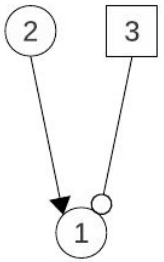
(e) example support

- inductive reasoning
  - dialectical level : ”Do you have an(other) example?”

1 conclusion - 1 evidence

[A citizens’ initiative can force the mayor to tear the building down.]1 [In Munich such a group forced the local authorities to tear down an old office building!]2

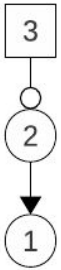
2 Annotation scheme - attack



(a) rebut a conclusion

- present an argument against the conclusion irrespective of the support for it
  - dialectical level : ”What makes you sure about your claim in the light of the following counterevidence?”

1 conclusion - 1 support premise - 1 attack on claim (*rebutter*)
[We should tear the building down.]1 [It is full of asbestos.]2 [On the other hand, many people liked the view from the roof.]3

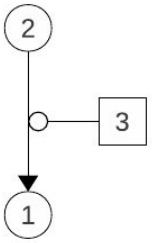


(b) rebut a premise

- attack the cogency of the given argument by attacking its premises

1 conclusion - 1 support premise - 1 attack on premise (*rebutter*)
[We should tear the building down.]1 [It is full of asbestos.]2 [Yet, nobody really made a precise assessment of the degree of contamination.]3

2 Annotation scheme - attack



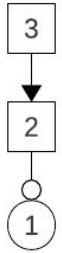
(c) undercut an argument

- attack the cogency of the given argument by diminishing their supporting force
  - dialectical level : ”Why do your premises make you so sure in light of the following condition?”
  - the invalidity of the inferential step from premises to conclusion by pointing to a possible exception
  - rebutting neither premise nor conclusion, but **acceptablility.** Thus, pointed to the relation

1 conclusion - 1 support premise - 1 attack on acceptability (*undercutter*)
[We should tear the building down.]1 [It is full of asbestos.]2 [The building might be cleaned up, though.]3

- Rebutting vs Undercutting
  - undercutters must be semantically related to the premise in some way, contrary to rebutters
  - rebutter: anticipated argument - undercutter : exception

2 Annotation scheme - attack



(d) supporting a rebutter

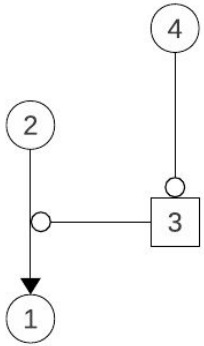
- present an anticipated argument against his conclusion or an anticipated exception to his argument, but also to strengthen it by explaining why it is worth taking this objection into account

1 claim - 1 attack on claim (*rebutter*) - 1 support to attacking premise

[We should tear the building down.]1 [On the other hand, many people liked the view from the roof.]2 [On weekends in summer, the roof is usually crowded with sunset partygoers.]3

2 Annotation scheme - counterattack

(a) rebut a rebutter



(b) rebut an undercutter

- If the attack itself was a rebutter, then the counter-rebutter is an argument for the negation of the rebutter.

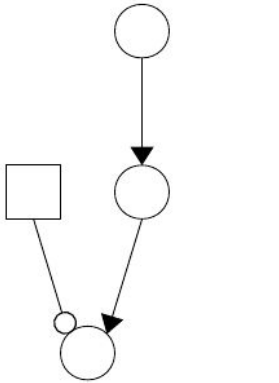
1 conclusion - 1 attack on claim (*rebutter*) - 1 attack to the contra premise
[We should tear the building down,]1 [even though it's supposed to be some touristic attraction.]2 [But, I've never seen any visitor groups there!]3

- If the attack itself was an undercutter, then the counter-rebutter is an argument for the negation of the undercutter.

1 conclusion - 1 support - 1 attack on acceptability (*undercutter*) - 1 attack to undercutter
[We should tear the building down.]1 [It is full of asbestos.]2 [Some new scientific study reportedly considers asbestos harmless,]3 [but that is probably only a hoax.]4

2 Annotation scheme - counterattack

if the rebutter is further supported



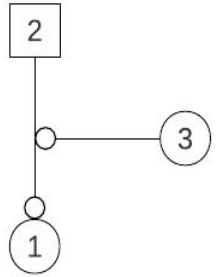
(c) counterconsideration

- leaving a rebutter uncountered, the author assumes that the arguments presented in favor of the claim will outbalance the arguments against the claim
  - the rebutting attack is of minor strength
  - the pro arguments are especially important
  - seems counterproductive but appears frequently in commentary texts

1 conclusion - 1 *rebutter* on claim - 1 support to claim - 1 support to the first support
[We should tear the building down.]1 [On one hand, many people liked the view from the roof.]2 [On the other hand, it is full of asbestos.]3 [The commission reported a significant contamination.]4

Freeman : uncountered attacks (or counterconsiderations in the terminology of Govier (1985) into his theory). While Freeman (1991, p. 173) argued that such counterconsiderations need not to be represented in argument structure, because they could be seen as rhetorical accessory, logically not effecting the case for the claim, they are now represented as a special 'even though' rebuttal in (Freeman, 2011, p. 29).

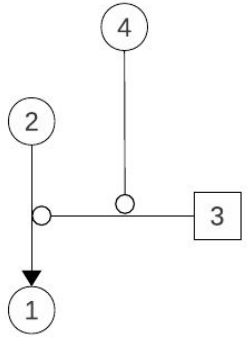
2 Annotation scheme - counterattack



(c) undercut a rebutter

- present an exception to the argument for the negated conclusion

1 conclusion - 1 attack on claim (*rebutter*) - 1 attack on acceptability of the rebutter
[We should tear the building down,]1 [even though it's supposed to be some touristic attraction.]2 [They'll surely build something more attractive on the site.]3



(d) undercut an undercutter

- present an exception to an exception

1 conclusion - 1 support - 1 *undercutter* - 1 attack the acceptability of undercutter
[We should tear the building down.]1 [It is full of asbestos.]2 [In principle it is possible to clean it up,]3 [but according to the mayor that would be forbiddingly expensive.]4

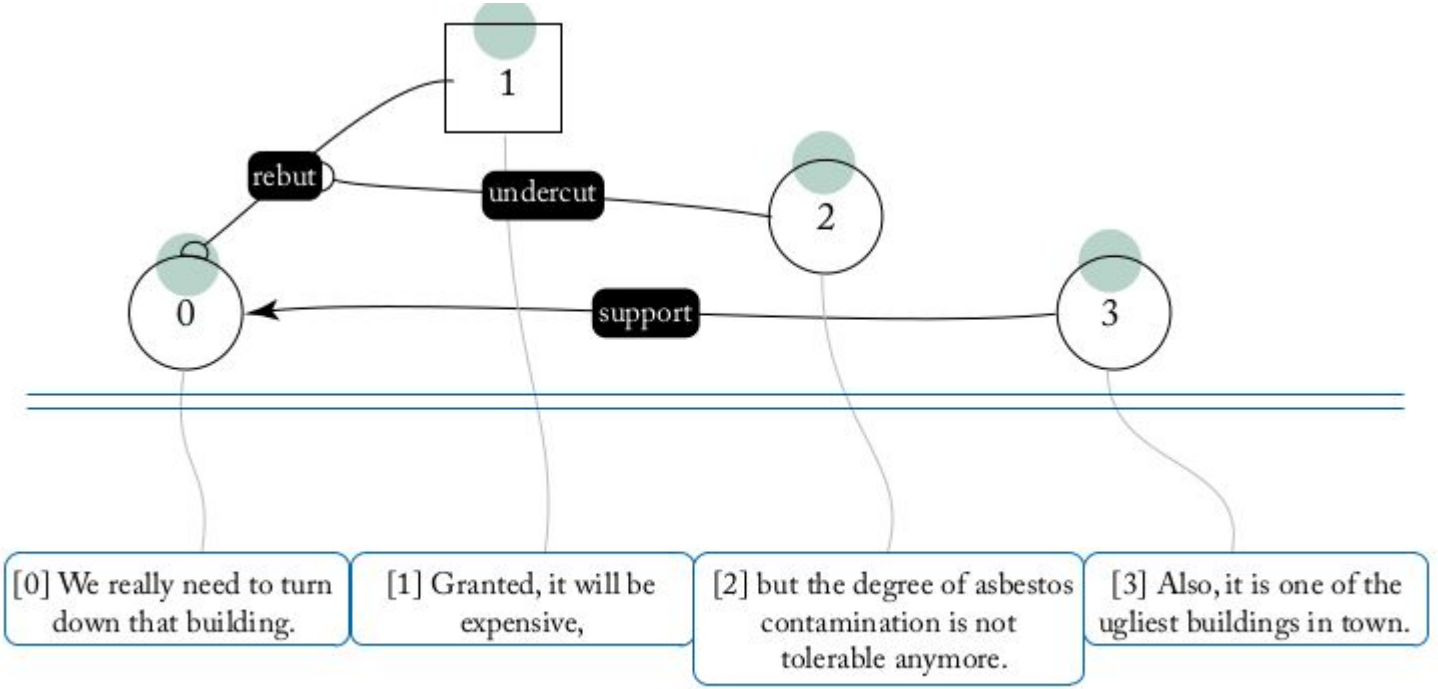
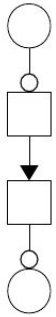


Figure 1.1: Graph representation of argumentation structure for Example 1.6.

2 Annotation scheme - counterattack

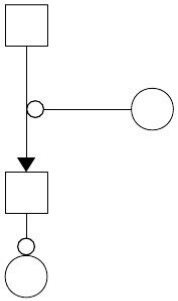
if the rebutter is further supported



(a) rebut an attack-support

- rebutting the premise in favor of the objection

1 conclusion - 1 *rebutter* on claim - 1 support to rebutter - 1 attack on rebutter's support
[We should tear the building down.]1 [On the other hand, many people liked the view from the roof.]2 [On weekends in summer, the roof is usually crowded with sunset partygoers.]3 [Yet, nobody never saw anybody partying on that roof.]4



(b) undercut an attack-support

- undercutting the support of the premise for the objection

1 conclusion - 1 *rebutter* on claim - 1 support to rebutter - 1 undercut the acceptability
[We should tear the building down.]1 [On the other hand, many people liked the view from the roof.]2 [On weekends in summer, the roof is usually crowded with sunset partygoers.]3 [The view might be blocked by the taller building in front.]4

- Using the terms ‘objection’ or ‘attack’ for these cases may be a bit misleading, because authors rarely intend to attack themselves or argue with themselves; instead, they ‘imagine’ an opponent and the confrontation.

- balance of consideration or pro and con arguments

- Frequent phenomenon : student essay (avoid *myside bias*), court decision, etc.

- Conductive arguments in which both positively relevant (pro) and negatively relevant (con) factors are put forward by an arguer. In such arguments, an arguer seeks to support a conclusion by citing a number of premises that count, or are taken to count, in its favor. He or she also acknowledges negatively relevant factors that count against, or are taken to count against, the conclusion.
- Counter-considerations are claims that count against, or are taken to count against, the conclusion claim, but despite that fact, are incorporated within an argument in which supporting considerations are also present. Conductive arguments that include counter-considerations as well as supporting premises may be called ‘balance of consideration’ arguments. They have also been called pro and con arguments.

3 ARG vs RST

- **Discourse structure** : explain the **coherence** of a text in general
- some theories : systematic connection to existing linguistic theories of syntax and semantics - SDRT, LDM, D-LATG
- **Rhetorical Structure Theory** (RST) : geared towards pragmatic description and doesn't worry about syntax or semantics
    - coherence relations : the idea that adjacent spans of text stand in a semantic or pragmatic relationship to one another, such as causality or contrast
        - make reference to the underlying intentions of the speaker or writer
    - Example : adversative, causal, temporal, and additive relations.

3 ARG vs RST : RST in a nutshell

- segments do not have equal status for realizing the underlying intention
  - **nucleus** (central : the main message can still be reconstructed) ≠ conclusion

    satellite (supporting : without nuclei, simply incoherent) ≠ premise
  - Marcu's compositionality criterion : large segments holds most important units (max nuclearity degree)

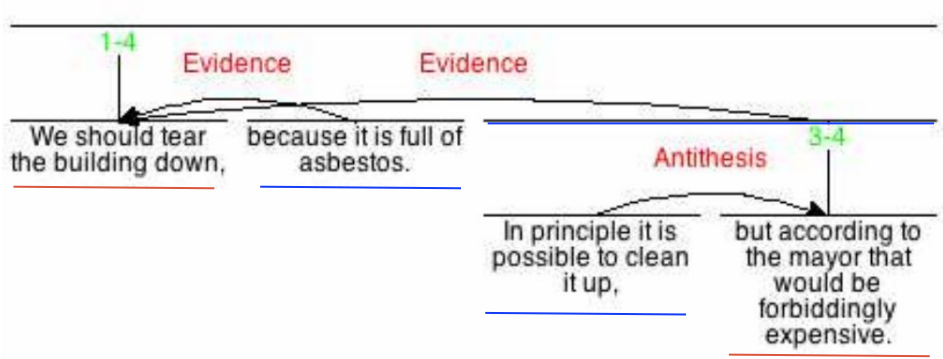


Figure 9: RST analysis for a short text

3 ARG vs RST : RST in a nutshell

- **Rhetorical Structure Theory** (RST) **definition of the relation**
  - constraints on the nucleus, e.g.: “reader might not believe nucleus to a degree satisfactory to the writer” or “nucleus is an action in which the reader is the actor”;
  - likewise, constraints on the satellite;
  - constraints on the combination of nucleus and satellite, e.g.: “reader’s comprehending the satellite increases his/her belief in the nucleus”
  - the intention of the writer, e.g.: “Reader’s positive regard for the nucleus is increased” or “reader recognizes that the satellite provides the framework for interpreting the nucleus”

➢ a variety of semantic and pragmatic aspects

3 ARG vs RST : RST in a nutshell

another important characteristic of **rhetorical Structure Theory** (RST)

- "subject-matter" relations hold in "the world" and are merely being reported in the text
  - writer intention : "Reader recognizes that X"
  - example : causal relation *Tom's train was delayed, and therefore he didn't make it to the meeting.*

- "presentational" relations are employed by the writer to actually change the beliefs or attitude of the reader
  - writer intention : "Reader's desire to perform the action described in the nucleus is increased"
  - example : *Motivational* or *Evidence*
  - particularly relevant to representing argument structure

3 ARG vs RST : RST for argumentative text

- **Rhetorical Structure Theory** (RST)
  - Goal : explaining the coherence of the actual linear sequence of textual units as its object of study
  - Limitation :
    i. long-distance dependencies of various kinds
      - arguments in text are not linearized in a straightforward way
    ii. rebuttal/counter-rebuttal configurations

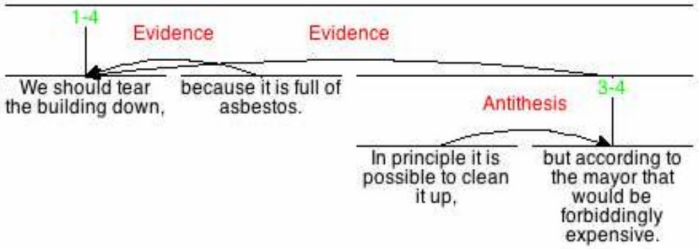
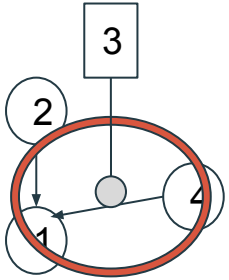


Figure 9: RST analysis for a short text



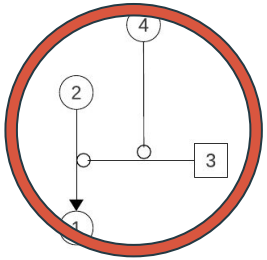


Figure 11: An analysis of the text from Figure 9 according to our proposed annotation scheme.

4 Toward automatic argumentation mining

1. Segmentation : Segment the text into argumentative discourse units (ADUs)
   - similar to find the ‘elementary discourse units’ (EDUs) in discourse parsing (phrases, clauses)
   - it is not clear that ADUs should always be as small as EDUs
   - approach : run a clause-based EDU segmentation, then establish relation holding between neighboring EDUs (in which combining EDUs into ADUs if necessary.
2. Segment classification : Identify the role/function of ADUs
   - a minimal argumentation analysis merely detects ‘premises’ and ‘conclusion’
3. Relation identification : Identify relations between ADUs
   - detect complex configurations and set in correspondence with one another
4. Argument completion : Build the overall structural representation
   - identify and instantiate an argument schema
     - postulation of implicit ADUs

1. Identify argumentative text
2. Segmentation : Segment the text into argumentative discourse units (ADUs)
3. Identify the central claim CC
4. Segment classification : Identify the role/function of ADUs
5. Relation identification : Identify relations between ADUs
6. Argument completion : Build the overall structural representation
7. Identify the type and the quality of the argumentation

4 Toward automatic argumentation mining

results neighboring disciplines

- Classifying segment status : analyze a text portion contribution to the overall text function
  - Teufel and Moens [2002] content zone analysis : automatic classification of zones in conference papers
  - 'Argumentative zoning' technique - genre = plausible notation/scheme
- Identifying coherence relations : Causal(support) vs Contrastive(attack, counterattack)
  - Linguistic realization : *explicit*(lexemes, syntactic constructions) and *implicit*(world knowledge, inference) relations
  - Approach:
    - Knowledge-based inferencing in the automatic detection of relations
    - Word-pair learning in causal relations from large corpora : fall, hurt (physical world cause-effect chains)

(a) The book never appeared, because the publisher had gone bankrupt.

(b) The book never appeared. The publisher had gone bankrupt.

  - Causal : intra-clausal pattern SUBJECT-VERB-OBJECT triple, causative or certain verbs (QA, bioinformatic)
  - Contrastive : manually-identified patterns of contrastive conjunctions and parts-of-speech

5 Major challenges for argumentation mining research

- Problem : lack of data, gap in annotation and analysis of argumentative text
- Reduced version on annotation scheme proposed in section 2:
  - Enable automatic system produces a coarse-grained annotation which subsumes a fine-grained one produced by human
  - Pragmatics-based annotation involves subjective judgment → IAA shouldn't be defined as strict identity
- **RST** : a theory of general text coherence
  - 2 limitations : long-dependencies and rebuttal/counter-rebuttal configurations (as presented in section 3)
  - Good first step, but needs an additional argumentation-oriented description

Questions asked previously

On mail of 20 April

Q1 : An undercutting is defined as a relation denies the relevance of a statement in supporting another. Is this counter-consideration relation only directed to an Attacking-type relation? Can it be directed to a Supporting relation or a statement(an ADU)?

Answer :

1. First, the term counter-consideration is wrongly used here, this term (originated from Gover [1985]) means the phenomena of the presence of both support- and attack-type premise in a claim.
2. Secondly, an undercutter relation can be used to attack the **acceptability** of either a support or a rebut.
3. The undercutter is meant to questions the relevance of a statement, thus only directed toward a relation not ADU itself. The relations being questioned regarding theirs relevance can be either *undercutter*, *rebutter* or *support* relations.

Questions asked previously

On mail of 20 April

Q2 : In Arg-micro-text corpus, there are 5 relations: CC, support, rebut, undercut and linked? And linked premises are usually found in supporting statements, rather than the attack ones?

Answer:

1. In this annotation scheme, there are 5 relations : 2 attacking-types - *rebut* and *undercut*, 2 supporting-types - *standard/normal* and *example*, and CC.
2. Linked is not a relation but one of the basic common basic argument structures.
3. Linked premises are usually found in *supporting* arguments, and only rarely in attacks. [Peldszus and Stede 2016]

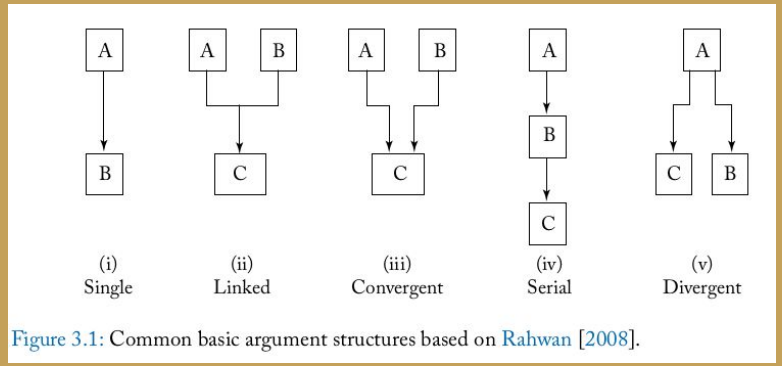
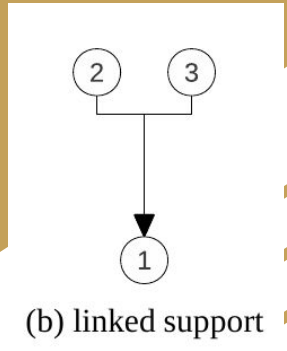


Figure 3.1: Common basic argument structures based on Rahwan [2008].



(b) linked support

More on RST and ARG

Argumentation Mining , Stede and Schneider, p.33

Table 3.1: Bentahar, Moulin, and Bélanger's taxonomy of argumentation models, modified from Bentahar et al. [2010, p. 215]

Model Type	Argument Evaluation Based On	What is Linked	How They Are Linked	Structure
Monological	Tentative proofs	Premises, claims	Internal inference structure	Microstructure
Rhetorical	Audience's perception	Whole arguments	Persuasion structure	Rhetorical structure
Dialogical	Defeasible reasoning	Whole arguments	Dialogical structure	Macrostructure