PREPRINT VERSION
please consider & only cite the official publication in IJCINI, 7(1), p.1-31, 2013

From argument diagrams to argumentation mining in texts: a survey

Andreas Peldszus and Manfred Stede
Universität Potsdam
EB Cognitive Science
peldszus|stede@uni-potsdam.de

Abstract

In this paper, we consider argument mining as the task of building a formal representation for an argumentative piece of text. Our goal is to provide a critical survey of the literature on both the resulting representations (i.e., argument diagramming techniques) and on the various aspects of the automatic analysis process. For representation, we also provide a synthesized proposal of a scheme that combines advantages from several of the earlier approaches; in addition, we discuss the relationship between representing argument structure and the rhetorical structure of texts in the sense of Mann and Thompsons (1988) RST. Then, for the argument mining problem, we also cover the literature on closely-related tasks that have been tackled in Computational Linguistics, because we think that these can contribute to more powerful argument mining systems than the first prototypes that were built in recent years. The paper concludes with our suggestions for the major challenges that should be addressed in the field of argument mining.

1 Introduction: Analyzing argumentative text

One of the central aspects of human communication is *argumentation*: the process of conveying inclinations, attitudes or opinions, and trying to make the partner accept them - or even adopt them. Cognitive agents, when they team up to solve a complex task, are in a similar position when the "division of labor" is to be negotiated. For human beings, the medium for arguing is natural language, whereas software agents do so in some suitable formal language. Nonetheless, with the progress that text mining techniques and applications have achieved in recent years, argumentation becomes increasingly relevant also for automatic processes, and hence also for cognitive computing.

Among the considerable body of research in computational argument, surprisingly little attention has so far been devoted to the issue of locating and analyzing argumentation in naturally-occurring text. A presumable reason is the fact that argumentation in "real" text is often not particularly crisp and clean – the argument proper is being infiltrated with the full range of problems of linguistic expression that humans have at their disposal. On the other hand, especially with the growing importance of social media communication, the amount of written argumentative discourse is rapidly growing, and interesting practical applications of finding arguments in text become visible on the horizon.

In correspondence with the popular notion of *text mining* (nowadays a cover term for many classification, search, and information extraction tasks) and the thriving research field of *opinion mining* (the detection of sentiment or opinions on products, people, organizations, issues) we thus see *argument* *mining* as the automatic discovery of an argumentative text portion, and the identification of the relevant components of the argument presented there. For the purposes of this paper, we concentrate on the second of these two steps. That is, we assume that an argumentative text (or an argumentative portion of a text) is already at hand, and the goal is to produce an analysis of the underlying structure of the argument that is being presented — in other words, a labeling of text portions with their argumentative roles and their relations to one another.

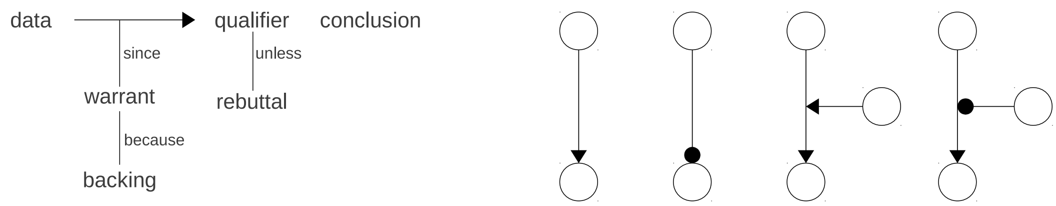
Why would this endeavor be of any practical relevance? One example is given by (Palau and Moens, 2011), who work with legal texts and aim at tracing the argumentation put forward by the parties involved, which may significantly enrich the retrieval capabilities of legal databases. Then there is the aforementioned area of opinion mining, for which the identification of arguments presents a natural extension: In addition to finding out whether internet users like or dislike a particular product (or any other entity), one might very well be interested in the *reasons* those users give for their opinions, inclinations, or decisions. Beyond the obvious commercial perspective on this task, there is a (potentially more exciting) prospect of perusing argument mining for public deliberation, e.g., as a tool for assessing public opinion on political questions.

In this paper, we provide a survey of research that is relevant to the goals we just formulated. Our contribution consists of two parts. The first begins with a critical discussion of various proposals for argument annotation schemes (diagramming techniques) that have been made in the argumentation community. This includes some diagramming techniques whose primary use was the illustration of the "essence" of argument, but which (at least in our opinion) can also be applied to the analysis of naturally occurring text. Our analysis of these approaches in Section 2.1 leads us to a proposal for a new schema that combines some advantages of different earlier schemes (Section 2.2). As an add-on, Section 3 discusses the relation between argument structure annotation and more general text structure annotation, in particular as exemplified by Rhetorical Structure Theory (Mann and Thompson, 1988), which has been suggested by some researchers to be suitable for argument representation as well. — The second part of the paper then surveys work on automatic argument mining, which includes results on subtasks from neighboring disciplines in text analysis (Section 4). Finally, Section 5 offers conclusions and our view of the major challenges that work on argument mining should address in the near future.

2 Annotation schemes for argumentation

In this section, we develop an annotation scheme for argumentation that draws on different ideas from the literature and our practical experiences with analyzing texts in the Potsdam Commentary Corpus (Stede, 2004). Naturally, the literature on argumentation is vast, and here we consider only work that has a clear focus on proposals for formal *notations*. A very useful earlier overview of the use of argument diagramming techniques to represent the structure of arguments has been given by Reed et al. (2007). They review the theories and diagramming schemes in logic, law and artificial intelligence and cover many important aspects relevant to (automatic) evaluation of arguments. In contrast, our aim is of a rather descriptive nature and our second focus is on the linguistic realization of argumentation, especially on the representation of argumentative opposition.

Therefore, some important and influential research will not be included here. For example, van Eemeren and Grootendorst (2004) discuss many interesting aspects of argumentation, but their goals are somewhat different: Their interest in a normative characterization of rational argumentation is not



(a) Toulmin (b) Grewendorf

Figure 1: Diagramming techniques in theories of argumentation of Toulmin and Grewendorf

directly applicable from a text-analytical and modeling point of view. Similarly, the argumentation graphs proposed by Dung (1995), while superficially similar to the kinds of graphs we discuss here, apply to a different level: Dung is interested in a "deep" representation of arguments that allows for formally modeling the reasoning processes; our concern, on the other hand, is with representation geared toward modeling the *textual presentation* of arguments.

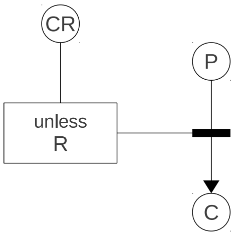
What follows is, first, a discussion of literature on theories of argumentation structure, where we intend to show that the approach of Freeman (1991, 2011) is a good starting point for the development of an annotation scheme. We then describe in detail the structures assumed in our annotation scheme and finally compare it with other influential approaches.

2.1 Theories of argumentation structure

An important step in the younger history of the development of a theory of argumentation is Stephen E. Toulmin's influential analysis of argument (Toulmin, 1958). Dissatisfied with the simple analysis of an argument into premises and conclusion, he investigated the actual use of arguments with the aim to identify different roles that utterances can play in arguments, i.e. the way they contribute to its persuasive force. Toulmin proposed a scheme with six functional roles (see Figure 1a): On the grounds of some evidence ('data') and a possibly implicit but defeasible generalization ('warrant') a conclusion is derived. The conclusion can be 'qualified' by a modal operator, indicating the strength of the inferential link. Furthermore, a 'rebuttal' can specify an exceptional condition that would undermine the inference if it holds. Finally, the warrant can be further supported ('backing').

Of the immense amount of literature on Toulmin's theory in different disciplines (ranging from philosophy, pedagogy to legal sciences, linguistics, artificial intelligence and others) we want to focus on a number of critiques that have been formulated, addressing problems of the application of the theory on complex, authentic argumentations, of the distinction of the functional roles and of the representation of the opponent.

Both Öhlschläger (1979) and Kienpointner (1983) suggest that the backing of the warrant should better be represented as a new, connected argument. In this new argument, the backing serves as the data and the warrant as the conclusion. Since a combination of several arguments is necessary for the representation of complex argumentation anyway, Öhlschläger presents a scheme that allows to recursively chain arguments together, thus building a serial structure, or 'multi-level argumentation' in the terms of Kopperschmidt (1989). Also, multiple arguments can be presented in favor of the same conclusion, which Kopperschmidt calls 'multi-threaded' argumentation. However, Öhlschläger's scheme does neither integrate Toulmin's rebuttal, nor the qualifier.



(a)

Proponent:	*We should tear the building down.*
Challenger:	*Why?*
Proponent:	*It is full of asbestos.*
Challenger:	*What makes you sure about your claim, given that it is in principle possible to clean it up?*
Proponent:	*According to the mayor, this would be forbiddingly expensive.*

(b)

Figure 2: Freeman's representation of a rebuttal and counter-rebuttal and the corresponding dialectal exchange

Similarly, Klein (1980) argued for a recursively applicable argumentation scheme. Furthermore, he claimed that the distinction between Toulmin's data and warrant cannot always be drawn precisely. He proposed a representation of argument that can be conceived basically as a support tree, with the root node as the main claim and supporting arguments in the unfolding tree structure.

However, all of the schemes discussed so far lack a proper representation of the opponent. Due to its dialectical nature, an argument always refers to an explicitly mentioned or at least supposed opponent, as for instance in the rebutting of possible objections. Wunderlich (1980) thus interpreted Klein's support-tree as a 'decision'-tree, where the root node is the 'quaestio', i.e. the question to be decided on. From there, not only arguments *for* and but also *against* the decision unfold recursively. Since there can be pro and contra for every node in the tree, the opponents role is integral to this representation.

Grewendorf (1980) then offered a dialog-oriented diagram method that also demonstrates the origin of arguments: It is possible to distinguish between counterarguments that are brought up by the opponent as attack from those that the proponent himself presents in order to refute them. In addition, Grewendorf replaces the tree structure with a graph, so that nodes can participate in multiple support or attack relations. In the diagram, they are no longer attributes of nodes but represented with different arrows (see Figure 1b). Those with an arrowhead denote support, those with a circle an attack. Finally, Grewendorf makes the important move to allow support and attack not only for statements (nodes) but also (recursively) for support and attack relations. Hence it is in principle possible to also represent metacommunicative disputes. However, Grewendorf provides only a rough outline of his diagram method and no formal specification. One of the things missing is a specification for conditions of a node having multiple support by a series of nodes. As a consequence, authors who took up his proposal sometimes produced ambiguous graphs that are difficult to interpret (e.g., Adachi-Bähr 2006).

A detailed examination of Toulmin's theory has been presented by Freeman (1991), whose goal was to integrate Toulmin's ideas into the argument diagraming techniques of the informal logic tradition (see Beardsley (1950) and its refinement by Thomas (1974)). Recently, an updated but compatible version of the theory has been presented in (Freeman, 2011). If necessary, we will distinguish between both versions in the following discussion, but otherwise simply speak of Freeman's theory.

The central claim of Freeman's theory is that the so called macrostructure of arguments, i.e. the different ways in which premises and conclusions combine to form larger complexes, can be modeled as a hypothetical dialectical exchange between a proponent, who presents and defends claims, and a challenger, who critically questions them in a regimented fashion. Every move in such a *basic di-*

alectical situation corresponds to a structural element in the argument diagram. The analysis of an argumentative text is thus conceived as finding the corresponding critical question of the challenger that is answered by a particular segment of the text. Freeman's theory thus makes an explanatory valuable connection between the focus on arguments *as process*, such as found in the study of dialectical dialogues in philosophy or in rhetorics, or even as a special form in judicial proceedings, and arguments *as product*, such as found in the study of persuasive text in radio or newspaper commentaries, in scientific writing or even advertising. The dialectical process serves as a model for the studied product. This orientation makes the theory accurately fitting our goal of argument structure recognition for authentic text.

Freeman presents critical questions a challenger would ask in a basic dialectical situation for the structural complexes typically assumed in the informal logic tradition: for serial, linked, convergent and divergent structures.[1] In his reception of Toulmin's theory, Freeman rejects the distinction between data and warrant: Although they are obviously discernible in arguments as process, he argues in an extensive discussion that the distinction is not applicable for arguments as product (see ch. 3 in both Freeman (1991, 2011)). Freeman thus assumes one category of premises subsuming both data and warrant. Also, the category of backing is dropped in favor of an analysis of serially connected arguments. The qualifier is integrated into Freeman's theory as a property of the inferential links between premises and conclusion.

Especially interesting for us is Freeman's integration of Toulmin's rebuttal. As described above, the *rebuttal* is an exception of the generalization presented as the warrant. It specifies a condition under that the claim would not hold. Typically, an author mentions a possible exception to preempt his critics and then in turn rebuts that anticipated objection, in Freeman's terms a *counter-rebuttal*. In the basic dialectical situation, the challenger asks a critical question with a possible objection, and thereby forces the proponent to defend his argument accordingly. A diagram featuring a simple sequence of claim, premise, rebuttal and counter-rebuttal and the corresponding hypothetical dialectical exchange is shown in Figure 2a. Additionally, Freeman (1991) allows the challenger to make his possible objection stronger by supporting it with further premises, which he terms "defended rebuttals". Although Freeman identifies different ways to attack and defend an argument, his use of the terms "rebuttal" and "counter-rebuttal" is rather general and corresponds to argumentative attack and counter-attack.

Freeman's diagramming technique is not perfect, though, especially for certain complex combinations of features: For some reason, Freeman lists all rebuttals of one argument in a single rebuttal-box. In order to relate the counter-rebuttals to their target, co-indexation is introduced instead of representing the relation by links in the diagram. Also, to represent a defended rebuttal that is countered by attacking its support, Freeman (1991, p. 199f) introduces further diagram features such as crossed-out boxes additionally containing the defeated rebuttal. These artificially increase the structural complexity, even though a more simple representation with links, analogous to the rebuttal, would be possible.

Another issue concerns the representation of the distinction between what Pollock (1995) calls *rebutting* and *undercutting defeaters* in defeasible reasoning. Even though Freeman makes equivalent distinctions, in acknowledging both the denial of the conclusion and the undercutting of its support (i.e. exceptional 'rebutting' in Toulmin's sense) as ways to attack an argument, or both denying the exception's holding and denying the exception's undercutting force as ways to counter an argument, these differences are not reflected in the argument diagram. Rebutting and undercutting attacks of the challenger are represented uniformly in Freeman's 'rebuttal'-box (although Freeman (2011, p. 23) lists them with different prefixes). Rebutting and undercutting counter-attacks of the proponent are

[1] A more detailed description of those complexes follows in Section 2.2

represented uniformly as nodes attached to the 'rebuttal'-box. Only (Freeman, 2011, p. 57) chose to diagram undercutting counter-attacks in a new fashion in order to visualize the difference.

Finally, it is worth noting that Freeman now integrates uncountered attacks, or *counterconsiderations* in the terminology of Govier (1985) into his theory. Here, instead of rebutting or undercutting a possible objection, it is left uncommented because it is understood as being outbalanced by more weighty reasons in favor of the claim. 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). This extension of the theory was in our view an advantageous move, as this argumentative strategy appears frequently in argumentative text and could not be adequately represented before. There are many more noteworthy features of Freeman's approach that are beyond the scope of this paper, as for instance the elaborate and ongoing discussion of the linked-convergent distinction or the representation of suppositions. However, this recapitulation of what is most relevant to our goals should be a sufficient basis for the following section.

Further theories are of interest, as for instance the classification of objections of Apothéloz et al. (1993) or of Walton (2011). We will discuss them in more detail in Section 2.3 in comparison with our synthesized scheme.

2.2 A synthesized scheme

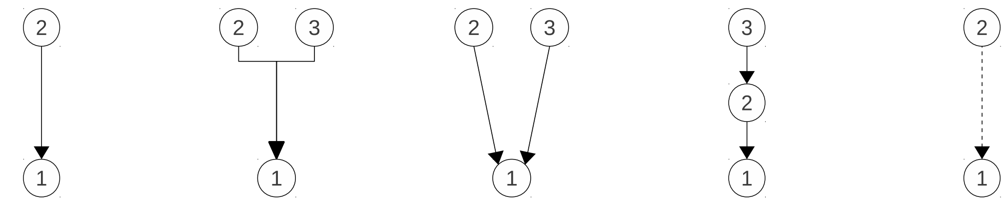
We now present our proposal of an annotation scheme, on the basis of which we annotate argumentation structure. It follows Freeman's idea of using the moves of proponent and challenger in a basic dialectical situation as a model of the structure of argumentation in texts, but represents the rebutting/undercutting distinction and complex attack- and counter-attack constellations more elegantly.

2.2.1 Basics

We define an argument to consist of a non-empty set of premises supporting some conclusion. We thus use the term 'argument' not for premises, but for the complex of one or more premises put forward in favor of a claim. Premises and conclusions are propositions expressed in the text segments. We can graphically present an argument as an argument diagram, with propositions as nodes and the support relation as an arrow linking the premise nodes to the conclusion node. The most simple configuration of an argument would consist of two propositions, one conclusion that is supported by exactly one premise, as in this simple example: *[We should tear the building down.]*$_1$ *[It is full of asbestos.]*$_2$. The corresponding structure is shown in Figure 3a.

If an argument involves multiple premises that support the conclusion only if they are taken together, we have a *linked* structure in Freeman's terminology. On its own none of the linked premises would be able to support the conclusion. In the basic dialectical situation, a linked structure is induced by the challenger's question as to why a premise is relevant to the claim. The proponent then answers by presenting another premise explicating the connection. Building linked structure is thus to be conceived as completing an argument.[2] As an example, consider the following continuation of our

[2]As discussed in Section 2.1, both Toulmin's data and warrants are represented as linked premises by Freeman. One might argue that data and warrant should not be linked according to this definition, for an argument might be fully functional without a premise corresponding to warrant. However, in this case the warrant would simply be implicitly assumed by the



(a) basic argument (b) linked support (c) multiple support (d) serial support (e) example support

Figure 3: Basic support relations and complex formation

first example: … *[All buildings with hazardous materials should be demolished.]*$_3$. Linked support is shown in the diagram by connecting the premises before they link to the conclusion, see Figure 3b.

When we speak of 'argumentation', we mean the structure that emerges when multiple arguments are related to each other and form larger complexes. The manner in which arguments combine to larger complexes can be generally described as either supporting, attacking or counter-attacking. In the following subsections we will describe each of them.

2.2.2 Support

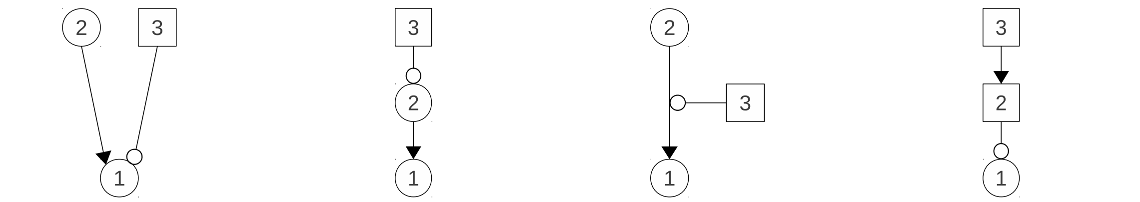
There are different ways to provide further support to the conclusion. One is to bring up a separate argument for the same conclusion, the other is to further develop the argument already given.

Let us start with the strategy where the author puts forward a separate, new argument for the same conclusion. For instance, consider this continuation of the original example: … *[Also, people in the neighborhood have always hated it.]*$_3$. Both arguments stand for themselves and each of them could be put forward by the author without the other. Both arguments are independent from another in the sense that the supporting force of one argument would not be impaired if the supporting force of the other is undercut.[3] On the dialectical level, the challenger asks: "Can you give me an additional argument for that conclusion?" and the proponent answers by offering a new argument accordingly. We call this structure *multiple* support, in order to prevent confusion with Freeman's convergent structure. In our case, we could say there are two *arguments* converging on the same conclusion. In contrast, what Freeman identifies as convergent structures are two *reasons* converging in one and the same argument. A discussion of that difference can be found in (Freeman, 2011, ch. 5). Bringing forward a new argument for the same conclusion is graphically represented as a separate arrow linking the premises of the new argument to the common conclusion, as in Figure 3c.

Another way to provide further support to the conclusion is to further develop the argument already given, by supporting one of the argument's premises. This would be the case if our original example would continue as follows: … *[The commission reported a significant contamination.]*$_3$. The author presents a new argument to convince the reader of the acceptability of a premise. By directly supporting the premise, he is indirectly giving support to the conclusion. The role of the supported text segment is then twofold, on the one hand serving as a premise in the original argument, on the other serving as the conclusion of the following argument. On the dialectical level the challenger

author. Since we aim to describe the author-relative argument as product, we postpone the issue of representing implicit premises for now. See also the discussion in Section 2.3.

[3]However, the arguments are not required to be independent in the sense of premise acceptability: If both arguments share a premise or have semantically interconnected premises, it may turn out that evaluating a premise in one argument as unacceptable also renders one in the other unacceptable.



(a) rebut a conclusion (b) rebut a premise (c) undercut an argument (d) supporting a rebutter

Figure 4: Challenger's attacks of the proponent's argument.

asks: "Why should I accept that premise?" and the proponent answers by offering a new argument accordingly. Following the terminology of Freeman and others, we call the resulting structure *serial*. Such serial structure is straightforwardly presented in the argument diagram by a new arrow linking the premises of the new argument to their conclusion, which is one of the premises of the original argument (see Figure 3d).

A special form of lending support to a claim is that of giving *example*s. If the author claimed a generalization, he can provide evidence that it proved to apply correctly at least in the given example. Consider the following example: *[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!]*$_2$. Those arguments are based on inductive reasoning. Since Freeman represents inductive reasoning as convergent premises, there is no special type of question for the challenger in his conception. To make this more fine-grained distinction, it is reasonable to assume the challenger would simply be asking: "Do you have an(other) example?" As with all supporting arguments, we represent the example arguments with an acute arrowhead, though with a dashed instead of a solid line; see Figure 3e.

2.2.3 Attacks

Now that we have presented the different ways to support an argument, we focus on ways to attack it. One is to present an argument against the conclusion irrespective of the support for it; the other is to attack the cogency of the given argument by attacking its premises or by diminishing their supporting force. Both of these strategies can be used by the challenger to attack the proponent's arguments and by the proponent to counter the challenger's attacks. In the argument diagram, attacks will be indicated by solid arrows with a round arrowhead. Furthermore, we think it is useful to be able to clearly distinguish between the challenger's and the proponent's attacks. Thus, in allusion to Freeman's rebuttal box, segments corresponding to attacks of the challenger will be represented as box nodes, while those corresponding to counter-attacks of the proponent as circle nodes. We now describe attacks of the challenger, and consider the proponent's counter-attacks in the section 2.2.4.

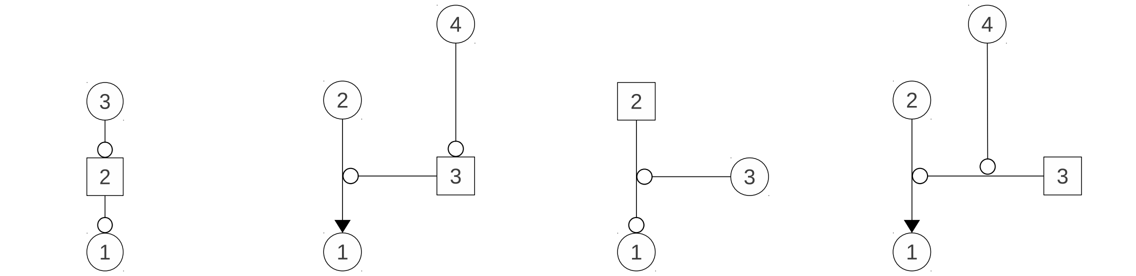
Let us start with the strategy to provide a new argument against the conclusion. For instance, let our original example be continued with the following statement: … *[On the other hand, many people liked the view from the roof.]*$_3$. The author anticipates that there are premises supporting the negation of the conclusion. In accordance with Pollock and Freeman, we call this type of attacking argument *rebutter*s directed against the conclusion. As mentioned in Section 2.1, Freeman does distinguish between rebutting and undercutting attacks by the challenger. However, he still represents both by the same structure in the argument diagram and does not present a challenger's question specific to rebutting attacks. Since we want to represent this distinction structurally, the corresponding challenger's question would be: "What makes you sure about your claim in the light of the following counterevidence?" Such rebutters are depicted in the argument diagram as arrows with round arrowhead from the challenger's premise to the proponent's conclusion. An example is shown in Figure 4a.

Instead of rebutting the conclusion, the challenger could also attack the given argument by rebutting one of its premises. As an example, consider this alternative continuation: . . . *[Yet, nobody really made a precise assessment of the degree of contamination.]*$_3$. Technically, this is not a new structure: Whether the attacked claim serves as a premise or as a conclusion in some argument is irrelevant for its being rebutted. However, this can be seen as a different strategy. By rebutting the argument's premises, the challenger argues against the argument's cogency. For a corresponding argument diagram see Figure 4b.

Another way to attack the arguments cogency is by questioning the supporting force of the premises for the conclusion. On the text level, the author is anticipating a possible exception (to an implicit or explicitly state rule) that could defeat his argument if it would hold. For instance, see the following continuation: . . . *[The building might be cleaned up, though.]*$_3$. On the dialectical level, the challenger argues for the invalidity of the inferential step from premises to conclusion by pointing to a possible exception. In doing so, he is neither rebutting the premise nor the conclusion, but restricting the applicability of the argument. The challenger's question presented by Freeman is: "Why do your premises make you so sure in light of the following condition?" With Pollock and Freeman, we call this type of attacking arguments *undercutters*. They are represented diagrammatically as arrows with round arrowhead directed to the body of the arrow representing the attacked support relation. Figure 4c shows an example.

Rebutting and undercutting attacks can sometimes be hard to distinguish on the challenger side: Is the given segment to be understood as an exception of the inferential move from premises to conclusion, or as an argument in favor of the conclusion's negation? A convenient way to tell them apart is to focus on the attacker's commitment to the conclusion. If the attacker claims that the negation of the conclusion actually holds, this is a clean indicator for a rebutting attack. However, the challenger is not allowed to assert a proposition in the basic dialectical situation. His role is defined very restrictively as that of a constructive partner testing the proponent's argumentation by asking critical questions. His goal is to wrench the best possible argument for the main claim from the proponent. He will thus never argue out of his own interest to convince the proponent of some claim. Consequently, he can neither claim that the negation of the conclusion holds, nor that some exception holds. He can only present *possible* arguments in favor of the conclusion's negation or *possible* exceptions to some inference from premises to conclusion, in order to provoke a corresponding reaction of the proponent. Nevertheless, we are hopeful that annotators will be able to discriminate between both cases, for undercutters must be semantically related to the premise in some way, contrary to rebutters. A possible test would then be to see how felicitous the attack is if the premise turns out to be false, is suspended, or is omitted. The rebutter will presumably be unaffected, while an exception without inference seems questionable.

Freeman (1991) permits the challenger to provide support to his attacks and so do we. As an example take the following segments: *[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.]*$_3$. On the text level this means that the author not only has the chance to 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. All sorts of supporting relations described in the previous subsection are available for that purpose. Dialectically, this support of an attack is modeled by a temporal role switch between challenger and proponent. In our argument diagrams these temporal role switches are already resolved, so that all supporting and attacking argu-



(a) rebut a rebutter (b) rebut an undercutter (c) undercut a rebutter (d) undercut an undercutter

Figure 5: Proponent's counter-attacks of the challenger's attack.

ments are related to proponent and challenger according to the main claim. An example is shown in Figure 4d, where a rebutting argument is supported by an additional premise.

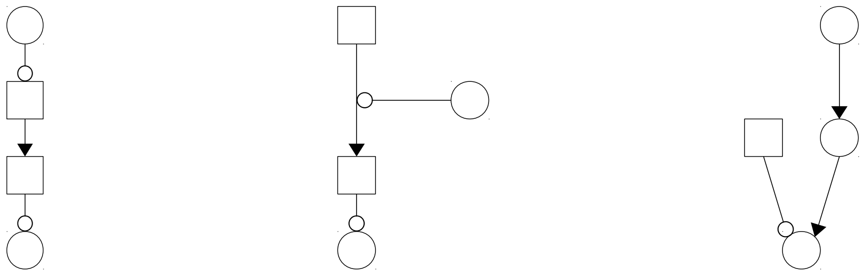
2.2.4 Counter-Attacks

How can the proponent answer to these challenges? Which possibilities are available to the author to counter the anticipated attacks? Freeman identified several ways to defend an argument. We will present what we regard as the most important ones and then discuss an additional possibility.

One straightforward way to counter an attack is to rebut it. Depending on the type of attack to counter, we can distinguish two cases: If the attack itself was a rebutter, then the counter-rebutter is an argument for the negation of the rebutter, i.e. the author is for some reason denying the anticipated argument against his original claim. For instance in: *[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!]*$_3$. For the corresponding structure see Figure 5a. If the attack itself was an undercutter, then the counter-rebutter is an argument for the negation of the undercutter, i.e. the author is denying that the exception holds. This is the case in the following continuation of our original example: …*[Some new scientific study reportedly considers asbestos harmless,]*$_3$ *[but that is probably only a hoax.]*$_4$ It may be that the exception would undercut his argument if it were true, but it is not. An example diagram is shown in Figure 5b.

An attack can also be countered by undercutting it. Again, this depends on the type of attack to counter: Undercutting a rebutter means to present an exception to the argument for the negated conclusion. The author not only shows that the anticipated argument against his claim needs to be restricted, but also that the argument is irrelevant for his claim, because the exception holds. An example would be: *[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.]*$_3$. Figure 5c illustrates this structure. Undercutting an undercutter correspondingly means to present an exception to an exception. The author does not even need to address whether the anticipated exception to his argument holds or not, because he can show that the anticipated exception itself is rendered irrelevant due to an exception. This constellation is shown in Figure 5d. For instance, consider the following continuation of the example introduced at the beginning of this subsection: …*[In principle it is possible to clean it up,]*$_3$ *[but according to the mayor that would be forbiddingly expensive.]*$_4$.

While distinguishing rebutters from undercutters seemed possible, though not trivial for the challenger's attacks, we expect it to be an easier task for the proponent's counter-attacks. Since the basic



(a) rebut an attack-support (b) undercut an attack-support (c) counterconsideration

Figure 6: Further features of the argumentation scheme

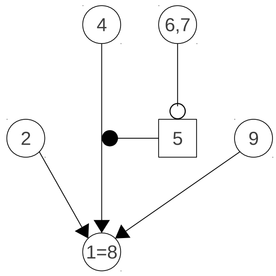


Figure 7: Supplemental features

dialectical situation does only forbid the challenger to assert, but not the proponent, it is likely that strong linguistic signals are found when (in rebutting) the negation of the target is actually claimed or when (in undercutting) the exception is actually claimed to be holding.

Given that the challenger provided support for his objection by additional arguments, another strategy to counter his objection is in attacking those supporting arguments. In this case, the proponent is arguing against the cogency of the argument in favor of the objection and thus diminishing its strength. The argument can be attacked either by undercutting the support of the premise for the objection (as shown if Figure 6b) or by rebutting the premise in favor of the objection (Figure 6a). For the sake of brevity we will not present further full examples. The interested reader is invited to extend the given examples accordingly.

The last possibility to react to an attack is to leave it uncountered. At first glance this seems counterproductive to the author's goal to convince the reader of his main claim. However, this appears frequently in commentary text. By leaving a rebutter uncountered, the author assumes that the arguments presented in favor of the claim will outbalance the arguments against the claim, either because the rebutting attack is conceived to be of only minor strength, or because the pro arguments are seen as especially important. This had been discussed under the term *counterconsiderations* in Section 2.1. Since the observation that a rebutter is for some reason not countered, can only be made retrospectively, no additional structure is required to represent counterconsiderations in our scheme. This is trivially shown in Figure 6c, where a rebutting attack is simply followed by premises directly and indirectly supporting the main claim, leaving the rebutter uncountered.

2.2.5 Extensions: applying the scheme to authentic text

So far, we have presented the ‘pure’ scheme, arguing from the need to represent abstract configurations of argument. When it comes to annotating authentic text, a few extensions are in order. For one thing, we frequently find that an author, when presenting an anticipated attack, may or may not signal that he regards this attack as *insufficient*. This evaluation can be done on two different grounds: On the one hand he can mark an attack as insufficient, because he attributes only little strength to it in the first place (as for instance in counterconsiderations). On the other hand, he could mark it insufficient in advance, because he will (from his point of view successfully) counter it in his later argumentation. Notice that in absence of a signal of insufficiency we cannot assume that the author regards the argument to be sufficient. Since the presence of such a signal is a valuable feature in the study of the principles of argument presentation, we highlight those attacks marked as insufficient by the author with a filled arrowhead.

Another important issue concerns the role of segmentation. The argumentation scheme presented above applies to a text that is already segmented into elementary units. While argumentation theories often assume a clean extracted list of the arguments found in a real text, a practical, segment-based annotation has to cope with the linguistic style of the author and the peculiarities of the segmentation process. We propose three operations that enable the annotator to handle some typical problems: The *glue* operation combines multiple adjacent segments into one argumentative unit, if the segment can be understood as one proposition. Consider the following three segments as an example: *The building is contaminated with asbestos. In every single corner. From the first to the last floor!*. We depict these in the argument diagram as one single node with a comma-separated list of the involved segments. Also, we sometimes find *restatements*, usually of the main claim in the argument. Therefore, if two (typically non-adjacent) segments appear to express the same proposition, both segments can be represented as one node with an equation of the involved segments. Finally, the *skip* operation allows the annotator to skip text segments that serve no argumentative purpose but contain, for example, only text-organizational material or background information.

As an illustration of those extensions see Figure 7, that shows the diagram of an imaginary text consisting of nine segments. The first represents the main claim, and a simple argument follows. The third segment has been skipped, because it served no argumentative purpose. Another argument follows and is undercut by the fifth segment. However, the author regards this undercutting attack as insufficient, because he rebuts it in the following two glued segments. The eighth segment is a restatement of the main claim, followed by a last supporting argument.

2.3 Comparisons with other classifications of argumentative opposition

Now that we introduced the various ways of attacking and countering attacks in our annotation scheme, we want to compare it to other classifications of objections found in the literature.

The first study that we consider here is that of Apothéloz et al. (1993), who investigate different strategies of objection in argumentation. Even though no formal notation of argumentation is provided, they aim to formalize the proposed operations of objection. Their approach shares some general assumptions with Freeman’s theory: They acknowledge the necessity of anticipating disagreement or possible objections in monologue argumentation. Furthermore, the difference between arguments as product and arguments as process (’as operation’ in their terminology) is reflected as well as the need to reconstruct the process in order to analyze the product (even though they do not propose an explicit model of this relation as Freeman does with the basic dialectical situation).

Their approach first distinguishes between two sorts of negative argumentation: that on the general discourse level (as in claims of misunderstanding or acts of discrediting the speaker), and counter-argumentations that attack reasons on a more factual level. Those counter-argumentations are then divided into four subcategories that attack different aspects of the stated reason: (i) its plausibility, (ii) its completeness, (iii) its relevance, and (iv) its argumentative orientation. They provide the following examples for those types of attacks:

(i)
1. Mary was in a very bad mood.
2. She didn't smile all evening.
3. Mary? She didn't stop laughing.

(ii)
1. You should buy the same car as Peter.
2. It is extremely comfortable.
3. It is much too expensive for me.

(iii)
1. I am not going to take the exam.
2. I didn't prepare for all the questions.
3. Just because you didn't prepare all the questions is no reason not to take the exam.

(iv)
1. ”A World Apart” is not a very good film.
2. It doesn't teach us anything new about apartheid.
3. That's precisely what makes it good.

To us, (i) seems equivalent to rebutting a premise (see Figure 4b). The distinction between completeness and relevance, however, seems difficult. Recall that in Freeman's linked structures, an argument is made more complete by providing another reason that makes the relevance of the first argument to the conclusion clear (see Section 2.2.1). The difference to Apothéloz et al. might be that in case of objections to completeness, the opponent already has some relevant reason in mind that was not considered in the proponent's argument; while for objections to relevance, the opponent has not yet understood the relevance of the proponent's premise to the conclusion. Both cases can be represented with our scheme: (ii) in general is equivalent either to undercutting the inference (as exemplified in Figure 4c) or to rebutting the premise (see Figure 4a), depending on the semantic relation between the attacking segment and the premise. In this particular example, since being availability of funds is a necessary precondition of buying, it appears to be an undercutter. In (iii) no reason is yet put forward by the opponent as to why the conclusion does not follow from the premise. According to our experience, it is quite unlikely that an author of a persuasive text would leave such a rejection without reason, especially for an anticipated rejection. If he continues to provide a reason, it would need to be integrated as usual, parallel to the case in (ii). Otherwise, if we allow rejections without reason, an 'empty' undercutter (i.e. without reference to a text segment) could straightforwardly represent the structure. Finally, objections to the argumentative orientation are an interesting special case, where one side uses a reason provided by the other side for their own ends, and thus claims an opposite orientation. Since our scheme represents claims of the proponent differently to those of the opponent (circle vs. box nodes), recurrence of the same proposition for another argumentative role is a convenient way of representation, as shown in Figure 8.

The other theory we want to compare our annotation scheme with, is the theory of argumentation schemes, which has been used extensively for the analysis of argumentative text. Here, arguments are

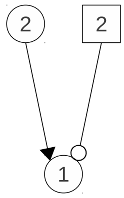


Figure 8: Recurrence of a proposition in objections concerning argumentative orientation

understood as instances of abstract argumentation schemes and a large and detailed catalogue of those schemes is provided. Each scheme specifies the required premises, the assumptions implicitly holding (unless they are questioned), and the exceptions that may undercut the argument. Every premise, assumption or exception corresponds to a critical question specific to the argumentation scheme. The theory is integrated as a diagramming technique in argument visualization tools such as Carneades (Gordon, 2010). Since our annotation scheme for argumentation shares a significant amount of the structural possibilities with Walton's argumentation schemes, it is reasonable to motivate why we chose not to use this theory as the basis for our annotation efforts.

Walton (2011) presented an analysis of objections that is similar to our proposal here. All premises in one argument are understood as being linked. Serial arguments and multiple arguments are basic complexes. What we understand as examples are in Walton's theory premises of a specific argumentation scheme. Rebutting attacks are expressed as arguments of opposite polarity. Undercutting attacks are expressed by claiming exceptions. Those attacks can be supported or rebutted again. However, one constellation that (at least technically) does not directly translate into argumentation scheme diagrams is the undercutting of an undercutter. In the argumentation schemes theory, exceptions are treated as a specific kind of premise and are thus part of the argument. An argument is undercut by giving evidence to show that the exception applies. While it is possible to undercut the argument presenting evidence for the exception, there seems to be no way to undercut the exception itself. This restriction may be intended by Walton though, since exceptions are licensed by the argumentation scheme: If there were an exception to the exception, then either the argumentation scheme was flawed, or wrongfully applied to the text.

This points to the obvious and more fundamental difference: Our annotation scheme is not based on argumentation schemes. We focus on the initial distinction between supporting, attacking and counter-attacking arguments and assume the structures needed to describe authentic argumentative texts in these general terms. This does not mean that a more fine-grained analysis of the involved arguments and the representation of implicit premises necessary to this end, as it is presented in the theory of argumentation schemes, is out of the question. On the contrary, our analyses might serve as starting point for a deeper analysis of argumentation.

Finally, it is important to be aware of the different aims. Our analysis is intended as a general, descriptive basis for corpus-linguistic studies and ultimately for computational applications such as argument mining. While argumentation as a process serves as a model for us, argumentation as product is the main focal point of our endeavour. Research around the theory of argumentation schemes on the other hand has shown increased interest in building tools to support argumentation as process for instructional, legal or decision-making ends and in advancing techniques of argument evaluation. In light of emerging standards, resources may hopefully be shared across the borders of those research goals.

3 Argumentation structure versus "rhetorical structure"

When aiming at identifying the structure of argumentation in authentic text, one should take the range of theories on *discourse structure* into consideration. Their goal is to explain the coherence of a text in general: How come that a text appears to the reader as a unified whole, and how do its parts relate to one another? Several such discourse structure theories have been proposed, and their design decisions reflect the specific sub-discipline in which they originated. Thus, some theories place their emphasis on the systematic connection to existing linguistic theories of syntax and semantics (e.g., SDRT Asher and Lascarides (2003); LDM Polanyi and Scha (1984); D-LTAG Webber et al. (2003)), but due to this orientation toward linguistic theory, applying them to authentic text is not straightforward. The approach that obviously is a candidate for representing argument structure is Rhetorical Structure Theory (RST; Mann and Thompson (1988)), which has been conceived as an empirical tool for practical text analysis, and the developers originally justified their design decisions with the claim that a fairly large number of texts from different genres have been successfully analyzed. Moreover, RST is geared towards pragmatic description (and does not worry much about syntax or semantics), since the definitions of coherence relations make reference to the underlying intentions of the speaker or writer. For these reasons, RST can be argued to constitute an adequate framework for the task of argument mining as set forth in this paper. In the following, we first provide a brief outline of the main ideas of RST, and then critically discuss work by researchers who proposed to use RST for representing argumentation structure. Approaches to the automatic derivation of RST trees will be described later in Section 4.4.

3.1 RST in a nutshell

The central notion for basically any theory of discourse structure is that of *coherence relation*, i.e., the idea that adjacent spans of text stand in a semantic or pragmatic relationship to one another, such as causality or contrast. This plausible intuition then needs to be operationalized, and this is one point where theories differ: Is there a finite, and reasonably small, set of such relations so that "any" text can be analyzed in this way? How are coherence relations to be defined? Mann and Thompson (1988) proposed a specific set of 25 relations, but they pointed out that the set should be regarded as in principle open to extension. However, the manifold practical uses of RST over the past 25 years (see (Taboada and Mann, 2006)) have shown that the relation set can in fact be regarded as relatively stable. It contains, for example, several adversative, causal, temporal, and additive relations.

The main criterion for judging whether a relation holds between two segments is the reconstruction of the writer's intention: Why did he state the two segments, why did he place them in adjacency, what did he want to achieve with this combination? In this regard, a major claim of RST is that for the vast majority of relations, the two segments do not have equal status for realizing the underlying intention. Instead, one segment (called 'nucleus') is the central one, whereas the other ('satellite') plays only a supportive role. Mann and Thompson point out that when all the satellites are being removed from a text, the main message can still be reconstructed (notwithstanding certain problems in cohesion, i.e. information flow). When nuclei are removed, on the other hand, the text will be simply incoherent. As an exception to the general rule, there are a few multi-nuclear relations that do not make this distinction – an example is the temporal relation Sequence, which applies to segments where one event is described to take place after another event.

For illustration, Figure 9 shows an RST analysis of a short (constructed) text (equivalent to the exam-

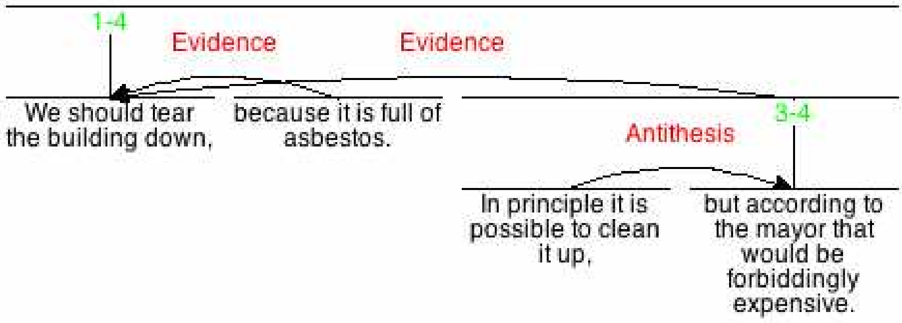


Figure 9: RST analysis for a short text

ple of Figure 2). It follows the notation suggested by Mann and Thompson and was produced with the RSTTool software[4]. Curved lines connect a satellite to a nucleus, with the arrowhead pointing to the nucleus, which is also indicated by a vertical line. Horizontal lines demarcate a larger segment arising from the fusion of smaller segments. Despite the unusual notation, the structure is in effect a tree. In RST, the main structural claims are that only adjacent segments can be connected by a relation, that the complete text needs to be covered by the analysis (there are no gaps) and that no crossing edges arise. (For a more thorough formal analysis of RST's predictions, see Marcu (2000)). The same set of coherence relations, therefore, is used to describe the relationships between the small segments (minimal units of the analysis) and recursively of the larger text segments. The definition of a relation consists of

- 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"

As an example, Figure 10 shows the definition of the Evidence relation as provided by Mann and Thompson. In general, the descriptions of constraints and intentions can refer to a variety of semantic and pragmatic aspects and in this sense do not constitute a very systematic framework. Clearly, they appeal to the intuitions of the analyst, and in the end the RST tree for a text will be the representation of one possible interpretation of a text. Nonetheless, when sufficiently-strict annotation guidelines are being formulated, it is possible to achieve acceptable agreement among human analysts (e.g., Carlson et al. (2003)).

Another important characteristic of RST to be mentioned here is the distinction between "subject-matter" and "presentational" relations. The former refer to relationships that hold in "the world" and are merely being reported in the text; in these cases the intention of the writer is of the form "Reader recognizes that X". An example is a causal relation that describes the link between two events: *Tom's train was delayed, and therefore he didn't make it to the meeting.* Presentational relations, on the other

[4] `http://wagsoft.com/RSTTool`

Relation: EVIDENCE
Constraints on N: R might not believe N to a degree satisfactory to W
Constraints on S: R believes S or will find it credible
Constraints on the N+S combination: R's comprehending S increases R's belief on N
Effect: R's belief on N is increased
Locus of the effect: N

Figure 10: RST-Definition 'Evidence' from Mann and Thompson (1988)

hand, are employed by the writer to actually change the beliefs or attitude of the reader. The Evidence relation, shown above, is an example; another one is Motivation, where the intention is "Reader's desire to perform the action described in the nucleus is increased". Obviously, it is this family of relations that is particularly relevant to representing argument structure.

Finally, we point out an observation made by Marcu (2000), which is important for interpreting RST trees, and which has gained wide acceptance in the community. The "compositionality criterion" states that when a relation holds between large segments, it also holds between its "most important units", where importance is defined by a maximum degree of nuclearity. When starting at the root node of the tree representing the segment and following only nucleus lines downward, one ends up at the most important elementary units of that segment. As a consequence, when applying this to the whole text, one is supposed to find the central statement/s of the text.[5]

3.2 RST for argument representation?

With its focus on speaker intentions and changes in reader attitudes, RST is by its design well-suited for studying argumentative text. While purely descriptive texts (e.g., encyclopedia entries) or narrations, including news reports, often have relatively "boring" RST analyses, the description of argumentative text can reflect the way this text works in an interesting way. Thus, even though the tasks of explaining the coherence of a text (the goal of RST) and capturing the argumentation found in a text are not identical, it is tempting to employ RST for representing the argumentation structure of texts and thereby to eliminate the need for a distinct diagram notation as we have discussed them in Section 2.

In this vein, Azar (1999) argued that the nucleus-satellite distinction is crucial to distinguish the two roles needed in an argumentative relationship, and that, in particular, five RST relations should be regarded as 'argumentative' in the sense that one segment is a conclusion (or an opinion), and the other segment is an argument brought forward to, for instance, increase the reader's belief in the conclusion. Specifically, Azar sees four relevant scales on which "increasing" can work, and they correspond to five RST relations: desire to act (Motivation) / positive regard (Antithesis and Concession) / belief (Evidence) / readiness to accept (Justify). Azar illustrates his idea with a few short sample texts, which he analyzes in terms of RST trees using these relations, and he claims that the argumentation found in those texts is represented adequately. However, in one of the examples, Azar uses an interesting twist in the tree representation: the 14 minimal units are labeled '1-6' (indicating that the substructure among these units is not relevant for the argumentation), '7+12', '8', '9', '10', '11', '12', '13', and '14'. The interesting one is '7+12', which is the central claim of the text, supported by Evidence relations from '8' to '14'. The claim thus is split between two non-adjacent text units, and Azar

[5] This will be a single elementary unit when no multinuclear relation is involved on the path from root to leaf.

simply fuses them into a single node. This is clearly in conflict with a central principle of RST, whose object of study is the actual linear sequence of textual units and their coherence. Azar therefore seems to regard RST more as a notation (which can be adapted to one's purposes) rather than as a theoretical framework.

In a somewhat similar fashion, Green (2010) borrows certain aspects from RST (several relation definitions and the nucleus/satellite distinction) for her 'hybrid' representation that is supposed to capture both the argument structure and the text structure. Green studies medical patient letters that explain a diagnosis and provide reasons for recommendations on the patient's behavior. The ultimate goal of the project is the automatic generation[6] of these letters, but text representation is a significant part of it. The tree structure suggested by Green consists of text segments participating in the argumentation, which are linked by RST relations, but this information is supplemented by decidedly argumentative annotations:

- *Implicit* statements are to be reconstructed by the analyst and added to the tree (as leaf nodes, on a par with the minimal text segments).
- Links from RST relation to the segments are labeled not only as nucleus and satellite but also with the role the segment is playing in Toulmin's scheme (data, warrant, claim, backing).
- Further, the links can be labeled with names of argumentation schemes (Walton et al., 2008).

When the analyst has identified the Toulmin-roles of the segments and the argumentation schemes that apply in that context, it becomes clear to her or him which implicit propositions need to be reconstructed. The tree structure in the end consists of text segments and formulations of implicit statements, and in addition, a leaf may contain a text segment copied from elsewhere, to cover cases where old information from the text is needed to complete an argument. On the whole, the representation thus takes some inspiration from RST but serves a different purpose than that intended by Mann and Thompson (i.e., to reflect the coherence of a text, with segments taken in the order of appearance).

We conclude this section by returning to the question whether the original, *bona fide* Rhetorical Structure Theory tree can be an appropriate device for representing argument structure. This is a relevant issue not only for theoretical considerations but also from the practical viewpoint of argument mining, since automatic RST parsing is nowadays an active research area (see Section 4.4). On the one hand, the interesting parallels between RST's presentational relations and argumentative moves, which had already been noted by Azar, make RST a promising candidate for a text-oriented argumentation representation. On the other hand, several observations from our own text analysis work with the Potsdam Commentary Corpus Stede (2004) indicate that there are (at least) two principal limitations.

The first is long-distance dependencies of various kinds. This is the problem that Azar circumvented by creating an artificial node covering two non-adjacent text segments. We found that quite often arguments in text are not linearized in a straightforward way: Pro- and contra arguments can be dispersed across the text and need to be linked to their common conclusions, which can violate RST's ruling out of crossing edges. Similarly, the end of a text often repeats, or slightly extends, the main thesis that has been stated earlier, so that the two segments have to be brought together.[7]

[6]RST has figured prominently in various other projects of the generation of persuasive text, which we do not cover here.

[7]We do not imply here that those texts are "bad" – they are perfectly easy to understand and have straightforward RST analyses. But from that analysis, the argumentation structure cannot be read off without adding "deeper understanding" and rewriting the representation.

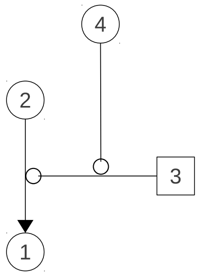


Figure 11: An analysis of the text from Figure 9 according to our proposed annotation scheme.

The other observation concerns the structural configuration of rebuttal and counter-rebuttal. Consider the text given in Figure 9. An analysis in terms of our scheme (Section 2.2) would regard segment 3 as a rebuttal to the support relation between 2 and 1, and 4 as the corresponding counter-rebuttal. See Figure 11. Our RST analysis in Figure 9 does not reflect the rebuttal-configuration at all. An RST tree that better reflects it would need to link segment 2 first to 3-4, for example via Antithesis with 2 as nucleus, and then 2-4 would be the Evidence for 1. From the text-descriptive viewpoint, this would be a bit unfortunate (the main "break" in the representation is in the middle of the first sentence rather than at a sentence boundary), but still it would be a plausible analysis. If the text were somewhat more complex, the overall configuration still makes sense if Marcu's compositionality criterion is applied: The most-nuclear segment in 2-4 is 2, which is the major Evidence for 1. Segment 3, however, is only a satellite to 4 (and rightly so), and therefore the compositionality principle cannot yield the information that 3 is closely linked to 1-2 (it is the rebuttal for this link) – this information is simply lost in the RST tree.

Examples of this kind are not uncommon, and unfortunately, the problem can get even worse. Imagine a linearization variant of the text: *[The building is full of asbestos,]*$_1$ *[so we should tear it down.]*$_2$ *[In principle it is possible to clean it up,]*$_3$ *[but according to the mayor that would be forbiddingly expensive.]*$_4$ Segment 1 can only be analyzed as Evidence for 2. The Antithesis between 3 and 4 stays the same, and it is then another Evidence for 2. But we have no way of capturing the rebuttal configuration between 1 and 3-4, as long as we adhere to RST's principles. Again, this is not a criticism of RST – the analysis for *its* purposes is perfectly plausible – but an observation on the limitations of its accounting for "deeper" structural configurations in argumentative text (assuming that one is interested in the rebuttal/counter-rebuttal configurations that we have discussed in the previous section).

4 Toward automatic argumentation mining

Finding arguments in text automatically is a relatively new research area. It is potentially relevant to any kind of text mining application that is directed at argumentative text; in particular, promising applications have been suggested for the following types:

- Legal cases: Mochales Palau and Moens (2009), amongst others, discuss the importance of finding argumentations and their structure in legal cases, as a subtask of the more general problem of finding *precedents* for a case that is currently under investigation.
- Scientific text: While the work in all scientific disciplines is clearly related to argumentation, the

texts in some disciplines are more amenable to automatic analysis than others. In the biomedical domain, for example, we often find concise argumentation why a new experiment is more successful or relevant than another one. (References to work in this vein will be given in Section 4.3.1.)

- User-generated content: The widely-popular task of *opinion mining* aims at detecting users' appreciations or disappointments with products or services, as found in various kinds of social media. A natural extension is to also find automatically the *reasons* they provide for their evaluations. Several researchers have explored this direction, including Wyner et al. (2012).

- Political deliberation: An extension of product-oriented mining for *reasons* is to be found in discussions of political issues, to be found in letters to the editor, or, again, in social media. The problem is more difficult because the issues tend to be more complex – it does no longer to suffice to look for sentences including the target term and the evaluation ("I *don't like* that new *cheeseburger*) and then for the reason in the immediate context ("...*because it's much too spicy.*"). Instead, it can take several sentences to mention the target issue, or some aspect of it, and to formulate an attitude toward it.

While this work is mostly in its early stages, there are several related tasks where interesting results have been achieved already. In this section, we first give a definition of the argumentation mining problem and then provide a brief survey of related research: at first on classifying the status of segments and on finding coherence relations, then on RST-style discourse parsing, and finally on "argument mining proper".

4.1 Argumentation mining: The problem

In a wide sense, mining a document collection for arguments involves the first step of finding relevant texts or text portions, i.e., those that provide a thesis (or "conclusion") and argumentation supporting it. As stated in Section 1, for the purposes of this paper, we ignore this step and thus assume that the relevant texts[8] have already been made available. Then, the following subtasks (or, depending on the concrete goal and the theoretical orientation, a subset of them) are to be addressed:

1. Segmentation: Break the text down into minimal units of analysis, henceforth called 'argumentative discourse units' (ADUs).

2. Segment classification: Determine the role that each ADU is playing for the argumentation.

3. Relation identification: Establish relations between individual ADUs, possibly leading to a complete tree or graph structure, or to an instantiated schema of sorts.

4. Argument completion: Steps 2 and 3 may involve the postulation of 'implicit' ADUs, which the analyzer constructs in order to achieve a complete structural description.

The segmentation task is similar in nature to that of finding 'elementary discourse units' (EDUs) in discourse parsing (which can be inspired by RST or any other relation-based discourse theory). Sentence boundaries are always considered to be EDU boundaries; but, in addition, complex sentences

[8]Henceforth, we use 'text' to refer both to complete (argumentative) texts and to extracted parts of text that, in some application, have been identified as relevant for the argumentation analysis.

may be broken into several EDUs, which generally correspond to clauses. This is relatively easy when clauses are combined by coordinating or subordinating conjunctions, whereas other clause combinations (e.g., gerunds) are more difficult to identify. Sometimes, certain types of prepositional phrases are also considered as individual EDUs. For argument mining, one can borrow the techniques that have been developed for EDU segmentation (see (Stede, 2011, Sct. 4.2)). However, it is not clear that ADUs should always be as small as EDUs: When two EDUs are joined by some coherence relation that is irrelevant for argumentation, the resulting complex might be the better ADU, when it collectively plays some specific role in the argumentation. One way to approach the task is to run a clause-based EDU segmentation, then try to establish the kind of relation holding between neighboring EDUs, and where necessary to combine EDUs into ADUs.

The way that steps 2-4 are being spelled out obviously depends on the type of structural description to be built, and thus on the selected theory. A minimal argumentation analysis merely detects 'premises' and 'conclusion' in step 2 (i.e., it performs a partial labeling of the ADUs). For certain applications (such as finding reasons for opinions in sentiment analysis), this may be sufficient. When more complex configurations are to be detected (cf. our discussion in Section 2), ADUs need to be set into correspondence with one another, which is done in step 3. Finally, in case the goal is to identify and instantiate an argument schema in the sense of Walton et al. (2008), this can involve the postulation of implicit ADUs in step 4.

The following two subsections discuss related work on steps 2 and 3 as it is done in the discourse processing community and can be carried over to argument mining. Likewise, subsection 4.4 discusses work on RST-inspired discourse parsing. Finally, work that is directly targeting argument mining (steps 2-4) will be presented in Section 4.5.

4.2 Classifying segment status

Given the knowledge that a text belongs to a particular *genre* (e.g., instruction manual, film review, newspaper article), portions of this text can be analyzed in terms of their contribution to the overall text function. In other words, a text can be broken down into 'content zones' that play a particular role for what the text is trying to do. Different genres have different content zones, and an individual genre can be characterized by specifying

- which zones are mandatory, and which are optional; and
- constraints and preferences on the linear order of zones.

In Computational Linguistics, one genre that has received quite a bit of attention in terms of content zone analysis is scientific papers. While there are some fine-grained differences depending on the domain or the particular scientific community, papers share essentially the same function (present a new result, compare it to the state of the art, argue why it is important) and the zone inventory and ordering is relatively standardized. Teufel and Moens (2002) aimed at the automatic classification of zones in conference papers and took a relatively fine-grained approach by treating individual sentences as the minimal segments (as opposed to paragraphs or sections). They called their approach 'argumentative zoning' and used the following zones:

- Aim: research goal of the paper
- Textual: statements about section structure

- Own: description of the authors' work (methodology, results, discussion)
- Background: generally accepted scientific background
- Contrast: comparison with other work
- Basis: statements of agreement with other work
- Other: description of other researchers' work

These 'argumentative zones' are somewhat different from what we are interested in here, but the problems are closely related: The segment status we need for argument mining is less genre-dependent and instead tailored to the argument notation or schema we want to use. Still, the status of segments needs to be identified, and the features used by Teufel and Moens for this purpose can be taken as inspiration. They include, amongst others: the position of a sentence within its section and paragraph; length of the sentence; formulaic expressions (e.g., *when compared to*, *following the argument in*) and the presence of 'significant' terms according to a *tf.idf* measure; three syntactic features (voice, tense, the presence of modal auxiliaries). Further, the context was modeled with a feature giving the most likely category of the preceding sentence. Finally, as an approximation of a semantic analysis, a hand-crafted verb lexicon was used to determine the action type expressed in the sentence (from an inventory of 20 classes), as well as the agent performing it. This is meant to distinguish between the authors themselves and other researchers, in 13 different verb categories.

Teufel and Moens implemented classification using a naive-Bayes approach. The performance differs widely between the zones: F-measure ranges from 26% for Contrast to 86% for Own. One zone that is of relevance to argument mining is that of criticism/contrast sentences where a precision of 57% and recall of 42% was achieved.

4.3 Identifying coherence relations

There are two groups of coherence relations that are especially relevant in argumentation analysis: Causal relations cover argumentative support, whereas Contrastive relations are used for Rebuttal and Counter-Rebuttal configurations. Obviously, 'Causal' here is meant in a wide sense, comprising not only causality between events in the physical world, but also relations (sometimes called Reason or, as seen above, Evidence) that speakers or writers establish between their assertions of various kinds, such as inclinations and their motivations.

As far as the linguistic realization of such relations is concerned, a distinction commonly made is that between *explicit* and *implicit* relations: The former (see (a) below) are overtly marked by certain lexemes (such as connectives) or syntactic constructions, whereas the latter (see (b)) can be recognized only on the basis of world knowledge and inference.

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

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

In (b), although the verb's tenses serve as a clue to the temporal ordering of the events, the causal relationship can only be inferred by the reader on the basis of his knowledge on the origin of books and the consequences of insolvencies. Knowledge-based inferencing was used in AI experiments in the 1980s, but lately it has not played a significant role in the automatic detection of relations. Only very recently, robust inferencing has once again been proposed for this task (see the next subsection). Besides, there have been some experiments on learning word pairs that regularly appear in causal relations from large corpora, but this is a matter of modeling typical cause-effect chains in the physical world (e.g., *fall – hurt*), and this method is unlikely to be helpful for argument mining, where such "typicality" can hardly be expected. Therefore, in the remainder of this section, we will focus our attention on detecting explicit relations.

There have been several studies on the fraction of coherence relations that are explicit in text, and they all arrived at roughly 40% (cf. (Stede, 2011, p. 110)). While there are no specific results for argumentative text, we can expect that the argumentative support ('causal') relations will often go unsignalled, whereas the rebuttal/counter-rebuttal configurations usually require a lexical signal so that the reader can identify the contrastive argumentative move. (The Concession coherence relation, for instance, is known to require a connective, such as *although*, *nonetheless*, etc.) In addition to connectives, which form a relatively well-demarcated class of lexical items, some researchers work with the slightly more general notion of *cue phrase*. It also covers phrasal items such as *for this reason*, which are open to modification and extension and thus are more difficult to enumerate. For English, a comprehensive study had been presented by Knott (1996). Later, Marcu (2000) worked with a similar inventory and operationalized it for the automatic detection of realizations of coherence relations and their EDUs. His approach operated on the text surface; patterns involved the lexical cues in tandem with search instructions such as "up to the next punctuation symbol". More recent work usually involved a certain amount of syntactic analysis, ranging from part-of-speech tagging, via chunking, to full sentence parsing.

4.3.1 Causal relations

In Computational Linguistics, the identification of causal relations has become a research topic in the question-answering (QA) community, where the handling of *Why* ... questions has received quite a bit of attention, and in bioinformatics, where the idea is to find descriptions of connections such as genetic mutations causing some particular symptom. To a large extent, this research has targeted intra-clausal patterns such as subject-verb-object triples involving causative (*to lead to*, *to bring about*, ...) and certain other verbs (see, e.g., Girju (2003)). In argumentation, the relevant units (in our terminology: ADUs) are usually larger, so that one is interested in inter-clausal patterns, with connectives or cue phrases playing the central role.

An influential early study in the medical domain was undertaken on Medline abstracts by Khoo et al. (2000), who implemented a pattern matcher that operates on the output of a dependency parser. By means of a corpus study, the authors handcrafted a set of 68 causality patterns, which included different types of connectives (conjunctions, adverbials, prepositions) as well as causative verbs. For the task of extracting causes and effects, they achieved F-measures between 41% and 59%.

More recently, Blanco et al. (2008) focused on intra-sentential cause-effect constructions and worked with less-domain specific corpora. Analyzing the TREC corpus, they identified four general syntactic patterns, and their pattern matcher achieved an accuracy of 51%. The authors then redefined the task as one of classification, and used features including the connective and its modifiers, the tense and the semantic class of cause and effect verbs. By decision-tree learning, they arrived at F-measures around 90% in their evaluation on the semantically-annotated SemCor corpus.

An interesting comparative study was recently presented by Mulkar-Mehta et al. (2011), who worked on several semantic relations, including causality, and used both surface-based approaches similar to those mentioned above, and deep semantic reasoning (abduction) with the Mini-TACITUS engine. They found on the one hand, that surface patterns for causality are highly domain-dependent (they compared biomedicine with football) and that a surface-level approach, which pays attention to such differences, is somewhat more successful than the deep reasoning approach.

Since connectives play an important role for identifying causal relations, for argument mining it can be helpful to attend to the difference between 'subject-matter' and 'presentational' relations, which we pointed out in Section 3.1. At least for some languages, this distinction corresponds to differences in connective usage. For example, in German, *denn* and *da* are generally assumed to be more argumentative in nature, whereas the frequent *weil* is regarded as more 'semantic'. For a detailed study on the contrast between 'semantic' and 'pragmatic' connectives in English and Dutch, see Knott and Sanders (1998).

4.3.2 Contrastive and other relations

Contrastive relations have received less attention so far, but some work has been carried out in the Textual Entailment (RTE) and, again, in the biomedical domain. For example, Kim and Park (2006) aim at finding statements of contrasts between objects or observations, and they use a set of manually-identified patterns of contrastive conjunctions and parts-of-speech in their context. In particular, they are interested in detecting syntactically-parallel structures involving the connectives *whereas, although, while, and, despite, but, since, however* and *yet* – targeting descriptions of phenomena where substances interact with one another while related ones do not. Clearly, for argument mining, this is too narrow a view, because rebuttals and counter-rebuttals need not employ syntactic parallelism. Still, the set of frequently-used contrastive connectives in a language is relatively small, and as pointed out above, the use of such a connective (or a cue phrase) is basically required in such constructions, so that their identification by surface-oriented methods is easier than for the causal relations.

Finally, notice that causal and contrastive relations are not the only ones of interest for argument mining. In particular, when several arguments are presented sequentially, additive signals such as *in addition* or *moreover* can play a role. This is why the broader topic of RST-style discourse parsing is relevant for designing approaches to argument mining.

4.4 Rhetorical parsing: Deriving RST trees automatically

While we noted some problems with using RST for argumentation representation in Section 3.2, it is clear that the task of RST parsing is quite similar to that of detecting argument structure, and therefore we briefly discuss some approaches to RST parsing here. An early implementation was presented by Marcu (1999), who applied the technique of shift-reduce parsing and demonstrated that the sequence of parsing operations can be learned from an annotated corpus. Marcu worked with 17 coherence relations and combined these with the three nuclearity assignments nuc-sat, sat-nuc, and nuc-nuc, yielding a total of 102 different Reduce operations, plus the Shift operation, that the parser needs to choose from.[9]

Marcu manually annotated a corpus of texts with RST trees and then automatically converted them to

[9] Hence, in contrast to standard RST, Marcu allowed all relations in both a nucleus/satellite and a multinuclear configuration.

the sequence of parsing operations that would construct them. For his machine learning experiment, he used the C4.5 decision tree inducer. The features were restricted to evaluate the top three items on the stack and the first item in the input sequence. In addition to relation-specific surface features, Marcu used the last five parsing operations as well as a few structural features (number of trees on the stack; types of textual units in those trees, i.e., paragraph, sentence, or headline; number of immediate children of root nodes; relations linking those children to the root; etc.) The parse-action classifier in isolation achieved an accuracy of 61%. When adding the segmentation task and thus building complete trees automatically, the performance of Marcu's system ranged about 15-20% below the performance of humans on three different datasets.

Several variants of the reduction scheme have been explored since that early work, e.g., by Subba and Di Eugenio (2009). They do not learn the Shift operation but use it as default when no Reduce rule can be applied with a sufficiently high confidence. Reduce rules are implemented as relation classifiers that are being tested at all valid attachment points of the tree built so far, and the highest-scoring combination of relation and attachment point is selected. The rules also evaluate some structural features such as the hierarchical representation of the segments built. An interesting aspect of the approach is the use of semantic representations. The accuracy of the trees, using a relatively large set of 26 coherence relations, is around 35%.

Another bottom-up search strategy was recently proposed by Hernault et al. (2010). These authors base their supervised learning on the RST Discourse Treebank (Carlson et al., 2003) and employ two different classifiers:

- A binary classifier STRUCT that yields a score for the decision whether *any* coherence relation holds between a pair of adjacent segments,
- A multi-class classifier LABEL that decides on the relation label (using the 18 relations from the RST Discourse Treebank) and nuclearity assignment for a pair of adjacent segments.

Both classifiers are implemented by means of support vector machines (SVM). Given the sequence of EDUs, the algorithm runs STRUCT on each pair of adjacent EDUs, selects the highest-ranking pair, applies LABEL to it, and builds the corresponding partial tree. This tree is then twice given to STRUCT (once for each neighbor), and again LABEL is applied to the highest-scoring pair of adjacent segments overall, resulting in the next tree combination, and so forth.

In contrast to the shift-reduce approaches, the strategy of Hernault et al. can be called a radically-greedy one, since the entire sequence of EDUs and partial trees is subject to choosing the maximally-likely next combination, which is then fixed (instead of taken as a hypothesis that may later be over-written).

4.5 Argument detection

Our discussion so far has dealt with tasks that have been carried out in discourse processing in general, and that are relevant to argument mining in accordance with our definition. Next, we describe what has so far been accomplished for "argument mining proper". This is the task we defined at the beginning of Subsection 4.1, which in its complete form would map a text (portion) to a graph representation such as the one we proposed in Section 2.2. For illustration, a complete argument mining system would build the graph shown above in Figure 11 for the text given (together with its RST analysis) in Figure 9. In practice, however, current systems build significantly simpler structures – which for certain purposes is adequate.

In one line of research, which drew inspiration from RST parsing but added rule-based syntactic and semantic analysis, Saint-Dizier (2012) presents the 'TextCoop' platform for analyzing text on the discourse level, and with emphasis on argumentation. The first application was tailored to procedural text and the arguments given therein, but gradually the implementation has been extended to cover other types of texts. TextCoop thus aims at fusing many of the tasks we have discussed above, and it does so on the basis of a logic-based grammar language ('Dislog'), which was designed to extend linguistic analysis from the sentence to the discourse level. While the implementation of TextCoop is still ongoing, experiments with other argumentation genres and phenomena have been conducted. Bal and Saint-Dizier (2009) proposed an annotation scheme for (so far, manual) argumentation analysis in newspaper editorials. Somewhat similar in spirit to the 'hybrid' representation of Green (2010), mentioned above, it combines the labeling of argumentative segments as 'conclusion' or 'support' on the one hand with RST-style coherence relations between segments. A set of additional attributes is used to characterize the strength of arguments and the persuasion effect, the type of the segment (fact or opinion) and in case of opinion its orientation (positive, negative, neutral), and several others. – Another application targeted by TextCoop is the analysis of evaluative expressions (opinions) and its connection with coherence relations, which is investigated by Villalba and Saint-Dizier (2012).

A different line of work, which has reached a good level of maturity, is that by (Palau and Moens, 2011), who focus on the domain of legal texts and aim at automatically detecting argumentation therein. Their data comes from the European Court of Human Rights (ECHR), and from the Araucaria[10] corpus of annotated arguments. In contrast to the more strongly linguistics-oriented work of the TextCoop group, Mochales Palau and Moens perform machine learning, using a variety of features on different levels of description. Their implementation proceeds in two steps: First, sentences are being classified as either 'argumentative' or 'non-argumentative'. In their implementation with a multinomial naive Bayes classifier and a maximum-entropy model, their best average accuracy was almost 74%. Features included word n-grams (n ranging from 1 to 3), combinations of any two words in the sentence, adverbs and modal auxiliaries as detected by a POS tagger, measures of parse tree complexity, punctuation symbols, and certain connectives from the list of Knott (1996) (mentioned above).

The second step tries to further classify the 'argumentative' sentences into the categories of 'premise' and 'conclusion', and here they achieved F-measures of 68% and 74%, respectively. The features include the position of the sentence in the document, sentence length, verb tense, reference to law articles, several domain-specific argumentative phrases, and the type of subject (applicant, defendant, court, other), amongst others.[11]

In addition, Mochales Palau and Moens ran some initial experiments on detecting the structure of the argumentation, again taking inspiration from RST parsing. However, they chose a rather different approach and devised a context-free grammar over terminal symbols that capture several of the features mentioned above plus some additional ones; examples are 'sentence with conclusive meaning' or 'support rhetorical marker (moreover, furthermore, also, ...)'. The goal then is to link premises and conclusions to one another, and with a small set of texts from their ECHR corpus, the authors obtained

[10] `http://araucaria.computing.dundee.ac.uk/`

[11] Recently, Rooney et al. (2012) addressed this classification problem, also working on Araucaria data, by kernel methods; the performance was slighlty lower than that achieved by Mochales and Moens, but the authors argue that no heuristic feature selection is required, thus making it easier to move to a new domain.

60% accuracy in detecting those structures.

Moving a conceptual step beyond this work, Feng and Hirst (2011) describe an approach to classifying argument schemes in the sense of Walton et al. (2008). Their basis is the set of five "most frequent" schemes, where different configurations of premises lead to a conclusion: Argument from example, from cause to effect, from consequences, from verbal classification, and Practical Reasoning. Their data also comes from the Araucaria corpus, and one central goal beyond the implemented schema classification is the reconstruction of enthymemes (unstated premises). Feng and Hirst assume that their input text segments are already classified in terms of *premise* and *conclusion*; then, they run a scheme classifier. It uses some general features modeling the absolute and relative position of premises and conclusion, the ratio of length of premise and conclusion, and the number of premises. In addition, the type of argumentation structure – whether it is linked or convergent (see Section 2) – needs to be supplied to the classifier (it cannot be detected automatically). Also, scheme-specific features are added, which consist mainly of particular cue phrases, and a few other measures. In one-against-others classification, the system yields best average accuracies of over 90% for two schemes, while for the other three schemes the results are between 63% and 70%.

5 Discussion: Major challenges for argumentation mining research

Argument mining in text is a new area of research with the potential of many interesting and useful applications. The fact that not too many results have been reported yet is, in our view, largely due to the lack of data: Essentially, there is only the Araucaria corpus mentioned in the last section, which offers a set of readily-analyzed arguments in a reduced form – it is not a *text* corpus that would offer annotations of argumentative moves in authentic textual environments. Corpus creation, however, requires annotation schemes and technical formats. While some formats for the abstract representation and exchange of arguments exist, we see a gap in the annotation and analysis of argumentative text.

AML, the markup language used in Araucaria, and the argument interchange format AIF (Rahwan and Reed, 2009) provide a good basis for representing arguments but were not intended specifically for purposes of text annotation. The elaborate annotation scheme for newspaper editorials suggested by Bal and Saint-Dizier (2009) (mentioned in Section 4.5) illustrates the wide range of phenomena that should be covered when authentic text is to be handled thoroughly.

As discussed in Section 3, Rhetorical Structure Theory has occasionally been proposed as the basis for an annotation scheme, but we pointed out that both Azar (1999) and Green (2010) saw the need to extend it for their respective purposes. And we gave examples illustrating that RST, as a theory of general text coherence, is from a structural viewpoint not quite fit to represent the deeper dependencies of argumentation; furthermore, the RST relations are of course meant to be general and thus do not precisely reflect the specific argumentative configuration. Thus we rather see RST-style analysis as a useful first step, but an additional text-type specific level of description – here: for argumentation – is needed.

Drawing on proposals for argument diagramming in the literature, in Section 2 we suggested an inventory of relations that covers the relevant forms of supporting, attacking, and counter-attacking. In addition, some extensions were listed that are necessary when dealing with the analysis of authentic texts. We have applied our scheme to texts from the Potsdam Commentary Corpus (Stede, 2004), which contains a range of phenomena that we could not represent adequately with the earlier schemes from the literature. An empirical evaluation of our new scheme, involving inter-annotator agreement studies, is currently under way.

Obviously, some of the fairly fine-grained distinctions in annotation schemes are intended for human annotation, but we do not expect automatic argumentation mining to be able to capture all of them. An important issue for future work therefore is to work out a reduced version of the scheme that is technically compatible with the original one but contains less information: the more coarse-grained annotations (as produced by automatic systems) would technically subsume the more fine-grained ones (as produced by human annotators). For example, the coarse scheme would not distinguish between a filled and an empty arrow, marking whether the author signalled a counterargument as insufficient. The accuracy of automatically-produced 'coarse' analyses can then be measured against the 'fine' annotations without the need to have two separate annotations. Furthermore, defining the subsumption relations between annotation constructs also allows for more elaborate computation of inter-annotator agreement between human annotators; this is important because any pragmatics-based annotation involves subjective judgment, and "agreement" should not be defined as strict identity.

When substantial annotated corpora covering different text genres are available, rapid progress can be expected for machine-learning approaches to automatic analysis – as other areas of computational linguistics have demonstrated. The first important challenge for an analysis in terms of a scheme along the lines we have discussed is the identification of the central thesis – because the other analysis steps (support, attacks) depend on it. The difficulty of finding the thesis differs widely between genres, which is one reason for the goal of ensuring genre-diversity in the corpus. In certain texts, human annotators will have difficulty agreeing on the central thesis, because it is conveyed only implicitly by the text, rather than being explicitly spelled out. If such cases are to be handled, more flexibility in the annotation scheme is required.

As a final remark on genre, it is in our view advisable to distinguish carefully between genre-specific methods and features (e.g., word n-grams, phrases, position in document) on the one hand, and genre-neutral features on the other. The behavior of causal and contrastive connectives, for example, will be very much identical across genres, so that models should be applicable to quit diverse applications. To this end, it will be helpful to exploit results from linguistics (for connectives, in particular, there is a rich literature analyzing their semantic and pragmatic features) and transfer them to computational implementation. Surface features alone are not likely to be sufficient, though, because identifying the structure of an argument to some extent depends on world knowledge (which is hard to make available and thus is often replaced by lexical models) and on the structure of the text document: At what point in the text particular argumentative zones are to be expected, depends (once again) on the genre, so that knowledge on typical text structures, represented in an effective way, will often be useful resource.

References

Adachi-Bähr, S. (2006). *Kontrastive Analyse von Gliederungsprinzipien in argumentativen schriftlichen Texten im Deutschen und Japanischen*, volume 20, Ausgabe 1 of *Arbeitspapiere und Materialien zur deutschen Sprache*. Institut für Deutsche Sprache, Mannheim.

Apothéloz, D., Brandt, P.-Y., and Quiroz, G. (1993). The function of negation in argumentation. *Journal of Pragmatics*, 19(1):23–38.

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press, Cambridge.

Azar, M. (1999). Argumentative text as rhetorical structure: An application of rhetorical structure theory. *Argumentation*, 13:97–114.

Bal, B. K. and Saint-Dizier, P. (2009). Towards an analysis of argumentation structure and the strength of arguments in news editorials. In *Proceedings of the AISB Symposium on Persuasive Technologies*, Edinburgh.

Beardsley, M. C. (1950). *Practical Logic*. Prentice-Hall, New York.

Blanco, E., Castell, N., and Moldovan, D. (2008). Causal Relation Extraction. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, pages 310–313, Marrakech, Morocco.

Carlson, L., Marcu, D., and Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. In van Kuppevelt, J. and Smith, R., editors, *Current Directions in Discourse and Dialogue*, pages 85–112. Kluwer, Dordrecht.

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357.

Feng, V. W. and Hirst, G. (2011). Classifying arguments by scheme. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 987–996, Portland/Oregon.

Freeman, J. B. (1991). *Dialectics and the Macrostructure of Argument*. Foris, Berlin.

Freeman, J. B. (2011). *Argument Structure: Representation and Theory*. Argumentation Library (18). Springer.

Girju, R. (2003). Automatic detection of causal relations for question answering. In *Proceedings of the ACL Workshop on Multilingual Summarization and Question Answering*, pages 76–83.

Gordon, T. F. (2010). An overview of the carneades argumentation support system. In Reed, C., editor, *Dialectics, dialogue and argumentation. An examination of Douglas Walton's theories of reasoning and argument*, pages 145–156. King's College London, London.

Govier, T. (1985). *A Practical Study of Argument*. Wadsworth, Belmont, CA, 1st edition.

Green, N. L. (2010). Representation of argumentation in text with rhetorical structure theory. *Argumentation*, 24:181–196.

Grewendorf, G. (1980). Argumentation in der Sprachwissenschaft. *Zeitschrift für Literaturwissenschaft und Linguistik*, 10(38/39):129–150.

Hernault, H., Prendinger, H., duVerle, D., and Ishizuka, M. (2010). HILDA: A discourse parser using support vector machine classification. *Dialogue and Discourse*, 1(3):1–33.

Khoo, C. S., Chan, S., and Niu, Y. (2000). Extracting causal knowledge from a medical database using graphical patterns. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL)*.

Kienpointner, M. (1983). *Argumentationsanalyse*, volume 56 of *Innsbrucker Beiträge zur Kulturwissenschaft*. Innsbruck.

Kim, J.-J. and Park, J. C. (2006). Extracting contrastive information from biomedical literature. *ACM Transactions on Asian Language Information Processing*, 5(1):44–60.

Klein, W. (1980). Argumentation und Argument. *Zeitschrift für Literaturwissenschaft und Linguistik*, 10(38/39):9–56.

Knott, A. (1996). *A Data-Driven Methodology for Motivating a Set of Coherence Relations*. PhD thesis, University of Edinburgh.

Knott, A. and Sanders, T. (1998). The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, 30:135–175.

Kopperschmidt, J. (1989). *Methodik der Argumentationsanalyse*. Frommann-Holzboog, Stuttgart.

Mann, W. and Thompson, S. (1988). Rhetorical structure theory: Towards a functional theory of text organization. *TEXT*, 8:243–281.

Marcu, D. (1999). A decision-based approach to rhetorical parsing. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pages 365–372, College Park/MD.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarization*. MIT Press, Cambridge/MA.

Mochales Palau, R. and Moens, M.-F. (2009). Argumentation mining: The detection, classification and structure of arguments in text. In *Proceedings of the ICAIL 2009*, pages 98–109. Barcelona, Spain.

Mulkar-Mehta, R., Welty, C. A., Hobbs, J. R., and Hovy, E. H. (2011). Using part-of relations for discovering causality. In *Proceedings of FLAIRS Conference*.

Öhlschläger, G. (1979). *Linguistische Überlegungen zu einer Theorie der Argumentation*. Niemeyer, Tübingen.

Palau, R. M. and Moens, M.-F. (2011). Argumentation mining. *Artificial Intelligence and Law*, 19(1):15–22.

Polanyi, L. and Scha, R. (1984). A syntactic approach to discourse semantics. In *10th International Conference on Computational Linguistics and 22nd Annual Meeting of the Association for Computational Linguistics, Proceedings of COLING '84*, pages 413–419. ACL.

Pollock, J. L. (1995). *Cognitive Carpentry: A Blueprint for How to Build a Person*. MIT Press, Cambridge, MA.

Rahwan, I. and Reed, C. (2009). The argument interchange format. In Rahwan, I. and Simari, G., editors, *Argumentation in Artificial Intelligence*, pages 383–402. Springer, New York.

Reed, C., Walton, D., and Macagno, F. (2007). Argument diagramming in logic, law and artificial intelligence. *The Knowledge Engineering Review*, 22(1):87–109.

Rooney, N., Wang, H., and Browne, F. (2012). Applying kernel methods to argumentation mining. In *Proceedings of the 25th FLAIRS Conference*.

Saint-Dizier, P. (2012). Processing natural language arguments with the TextCoop platform. *Journal of Argumentation and Computation*, 3(1):49–82.

Stede, M. (2004). The Potsdam Commentary Corpus. In *Proceedings of the ACL Workshop on Discourse Annotation*, pages 96–102, Barcelona.

Stede, M. (2011). *Discourse Processing*. Morgan and Claypool.

Subba, R. and Di Eugenio, B. (2009). An effective discourse parser that uses rich linguistic information. In *NAACL '09: Proceedings of Human Language Technologies – The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 566–574. Association for Computational Linguistics.

Taboada, M. and Mann, W. (2006). Applications of Rhetorical Structure Theory. *Discourse Studies*, 8(4):567–588.

Teufel, S. and Moens, M. (2002). Summarizing scientific articles – experiments with relevance and rhetorical status. *Computational Linguistics*, 28(4):409–445.

Thomas, S. N. (1974). *Practical Reasoning in Natural Language*. Prentice-Hall, New York.

Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press, Cambridge.

van Eemeren, F. H. and Grootendorst, R. (2004). *A Systematic Theory of Argumentation. The Pragma-Dialectic Approach.* Cambridge University Press.

Villalba, M. G. and Saint-Dizier, P. (2012). Some facets of argument mining for opinion analysis. In *COMMA 2012: Proceedings of the 4th International Conference on Computational Models of Argument*, Vienna. IOS Publishing.

Walton, D. (2011). How to refute an argument using artificial intelligence. *Studies in Logic, Grammar and Rhetoric*, 23(36):123–154.

Walton, D., Reed, C., and Macagno, F. (2008). *Argumentation Schemes*. Cambridge University Press.

Webber, B., Stone, M., Joshi, A., and Knott, A. (2003). Anaphora and discourse structure. *Computational Linguistics*, 29(4):545–587.

Wunderlich, D. (1980). Pro und Kontra. *Zeitschrift für Literaturwissenschaft und Linguistik*, 10(38/39):109–127.

Wyner, A., Schneider, J., Atkinson, K., and Bench-Capon, T. (2012). Semi-automated argumentative analysis of online product reviews. In *COMMA 2012: Proceedings of the 4th International Conference on Computational Models of Argument*, Vienna. IOS Publishing.