Towards Explaining Natural Language Arguments with Background Knowledge

Ioana Hulpuş[1], Jonathan Kobbe[1], Maria Becker[2], Juri Opitz[2], Graeme Hirst[3], Christian Meilicke[1], Vivi Nastase[2], Heiner Stuckenschmidt[1], and Anette Frank[2]

[1] Data and Web Science Group, University of Mannheim, Germany
`{firstname}@informatik.uni-mannheim.de`
[2] Institute for Computational Linguistics, Heidelberg University, Germany
`{mbecker, opitz, nastase, frank}@cl.uni-heidelberg.de`
[3] University of Toronto, Toronto, Canada
`gh@cs.toronto.edu`

Abstract. In this paper, we propose the task of argument explicitation, a task that makes the structure of a natural language argument explicit, as well as the background knowledge the argument is built on, in the form of implicit premises or contextual knowledge. The purpose of argument explicitation is to support the understanding of an argument by providing users with an end-to-end analysis that offers a critical assessment of arguments including identification of argument weaknesses. Besides, the results of the argument explicitation process can be used by machines to retrieve similar arguments as well as counter-arguments. We propose a framework for argument explicitation that joins a variety of AI and NLP-based argumentation mining sub-tasks that by now have mostly been treated separately in the literature. We identify the challenges this task entails, while at the same time highlighting the opportunities brought by the recent development of structured, external knowledge sources.

1 Introduction

The analysis and use of Argumentation in natural language texts is an active field of research in Artificial Intelligence. Common lines of work include the identification of argumentation units [32, 44, 50, 52] and relations [11, 36, 40, 50], the measurement of argument quality [24,57] and the synthesis of argumentative texts [56]. While many tasks in natural language processing (NLP) can be solved with surprising accuracy using only surface features, tasks relating to argumentation often require a deeper understanding of the reasoning behind a line of argumentation.

In this paper, we discuss the *problem of providing explanations for arguments*, giving an account of the opportunities and challenges this involves. We define the task of *explicitation of arguments* whose purpose is to support the understanding of a given argument by providing either end users or a computational system that tries to interpret an argument, with a structured and semantically founded analysis of the argument and to enrich it, if necessary, with explanations of otherwise implicit information that is crucial for the logics and understanding of the argument. This task brings together multiple research directions, some of which have already been investigated in the literature – however mostly in theoretical, as opposed to computational approaches. Indeed, we emphasize that while many of the challenges have been long debated in philosophy and logics communities, there are very few accounts of them in the NLP and modern AI communities, where these questions are now only starting to be addressed.

Argument explicitation is important in order to support end-users to critically judge natural language arguments. The need for systems that are able to perform argument explicitation has become particularly critical in the light of the current wave of references to "fake news". Explicitation of how the stated premises support or attack a given conclusion, as well as the provision of a full-fledged argument structure can shed light on both *validity* (does the conclusion follow logically from the premises?) and *soundness* (are the premises true?) of arguments. Beyond a purely logical account of argumentation, as one end of the extreme, or recourse to fact checking to corroborate the truth of premises on the other, argument explicitation aims at making explicit any background knowledge relevant for the understanding of the argument, either in the form of implicit premises, or relevant facts, states of affairs, or relations that connect explicitly stated argument components in a meaningful way.

In this paper, we discuss notions of explanations known in other contexts and motivate a new kind of explanation that is targeted to the explicitation of natural language arguments that makes the knowledge and the mechanisms of an argument explicit (Section II). We will distinguish different facets of argument explicitation and what specific kinds of knowledge are required for them (Section III). In Section IV, we discuss different types of argument explicitation and what kinds of explanations we can expect from them, in view of a content-based assessment of the validity, strength and quality of an argument. Section V summarizes our findings and concludes with an outlook on promising first steps towards a computational account of argument explicitation.

2 Explaining Arguments

2.1 Explaining Arguments with Deductive Reasoning

Researchers in the field of Logics consider arguments as logical formulas: the truth of the conclusion is a logical consequence of the truth of the premises. In this setting, the logical proof that establishes the entailment or inconsistency serves as an explanation of the respective relation. Consider the following example inspired from Walton and Reed (2005) [59]:

Example 1. Drastic decline in natural life is cause for alarm. Toads and frogs are forms of natural life and there is a drastic decline in toads and frogs. Hence, there is a cause for alarm.

Premise 1 $\forall x, natural\_life(x) \wedge drastic\_decline(x) \Rightarrow alarm(x)$
Premise 2 $natural\_life(toads\_and\_frogs)$
Premise 3 $drastic\_decline(toads\_and\_frogs)$
Conclusion $alarm(toads\_and\_frogs)$

Fig. 1. Example of formal logics-based explicitation of the argument in Example 1.

The example shows a syllogistic argument whose formalization is available in Figure 1. Given the formalization, an automated reasoner such as a Prolog reasoner can validate the argument. However, looking at this argument from the perspective of an everyday argument, it is straightforward to recognize several problems that reach beyond its deductive validity.

First, the text of the exemplified argument is rather unnatural, as the statement *toads and frogs are forms of natural life* is very unlikely mentioned in an everyday argument but it is most often implied. However, without it, the argument becomes deductively invalid, since it would miss *Premise 2* in Fig. 1. Most everyday arguments would face this problem. Arguments with unstated premises are called enthymemes [60] and we get back to them in the following sections.

Second, the argument's soundness is not beyond doubt. While the second premise would appear to be true to the majority of people, the truth of the first and third premises pertains to a higher level of subjectivity (when is decline *drastic*?). Indeed, in informal reasoning, counter-arguments question the validity of arguments as well as their soundness.

Thus, everyday arguments cannot be modeled in a deductive framework [59]. These arguments, whose conclusion can be defeated by either defeating the premises, or by adding new premises, are called *defeasible arguments*. In the following, we focus particularly on types of explicitations suitable for them.

2.2 Explaining Arguments with Informal Reasoning

In the informal reasoning literature, we identify several types of explanations each fulfilling a particular role, in different contexts:

Explanation as a discursive act has the function of providing reasons in support of an accepted conclusion [9,34,38]. In this regard, an explanation differs from an argument, as the explanation does not aim to prove the validity of the conclusion (which is the role of an argument), but rather considers the conclusion as being valid, and tries to provide the reasons for the occurrence of the event or state of affairs expressed by the conclusion.

Explanation as hypothesis occurs particularly in the context of abduction - the method of creating new hypotheses to explain unexpected observations, e.g. in the context of scientific literature [28, 53, 55].

Explanation for transparency is applied to enrich automatic systems with an output functionality that aims to inform the end-user with all the knowledge and processes used by the system for producing its primary output. This is the most common type of explanation in artificial intelligence [2, 43, 46].

In this paper, we discuss a new type of explanations, called *argument explicitation*: the explanation of an argument with the specific purpose of making the knowledge and mechanism of the argument explicit. The recent advances in natural language understanding and the availability of structured knowledge bases bear many opportunities to tackle some of the hard problems that this task entails.

3 Argument Explicitation

Broadly, the task of argument explicitation that we address consists of two subtasks. The first task – *argument analysis* – is concerned with analyzing the text in order to identify the *argument components* (e.g., premises and conclusion) and the overall structure of the argument. The second task – *argument reconstruction* – is concerned with making explicit any unstated, but implicit premises, as well as implicit connections between concepts mentioned in argument components, in terms of background knowledge. Most of the AI and particularly computational linguistics research in argumentation focuses on the first subtask [35], [39], [18], [32, 50], [3]. The second sub-task has by now been mainly addressed from a theoretical, or philosophical perspective by Walton and Reed (2005) [59], who reconstruct enthymemes (arguments with unstated premises) with argumentation schemes.

In the area of the argument analysis task, three very recent contributions outline the need for understanding argumentation on a deeper level. One investigation [37] shows that predictions of a state-of-the-art argumentative relation classification system are mostly driven by contextual shallow discourse features, while the model pays only little attention to the actual content of an argument. The need for deeper understanding of the content of the argumentative text has also been acknowledged with respect to the argumentative reading comprehension task (ARC)[4] [8]. The approach of Kobbe et al. (2019) [27], takes a step in this direction, but their knowledge-augmented model only marginally outperforms the linguistic baseline. Deeper understanding of arguments is even more crucial for the task of argument reconstruction, and as long as argument analysis is only achieved at a shallow level, there is very little hope for successful argument reconstruction on top of it. In light of these observations, we point out the kind of knowledge that such a system must access, model and integrate.

Knowledge about natural language is by far the most exploited type of knowledge in the literature with respect to argument mining. However, such knowledge has many facets, but it is by now only captured by relatively shallow features, such as discourse markers that indicate argument components (see e.g. [40]), or implicitly captured through training feature-based classifiers and recently, neural models (cf. [33, 49, 51]).

Knowledge about argumentation has been extensively researched, mostly in the philosophical literature. Here, multiple ways of modelling arguments have been proposed, including patterns of defeasible reasoning [14].

[4] SemEval-2018: https://competitions.codalab.org/competitions/17327

Background knowledge has probably been the most neglected type of knowledge in the current state of the art of argument analysis. Early argument comprehension systems [1, 5] made heavy use of hard coded, very precise domain knowledge. At the same time, in philosophy we encounter Schank's scripts [45] as the most referenced representation of domain knowledge for both argument comprehension and reconstruction [60]. Nonetheless, apart from very recent work of Botschen et al. [8] and Kobbe et al. [27], little progress has been made in using domain knowledge for argument comprehension and reconstruction. Recent work investigated the reconstruction of implicit knowledge in argumentative texts by way of manual annotation [4,7], but computational reconstruction approaches are still out of sight.

We claim that automated argument explicitation must model and reason with all of these complementary types of knowledge. In the following, we detail some of the sub-tasks of argument explicitation, focusing particularly on the challenges that can be addressed by, or that require exploiting background knowledge. We think that advances in the availability of large-scale knowledge bases bring significant opportunities in this direction.

3.1 Model-based Explicitation

In order to understand how and why defeasible arguments work, multiple argument models have been proposed. Generally these models aim to classify argumentative units on a more granular level than the generic *premise/conclusion* classification. In the following, we describe two of the most popular such models, and illustrate how we envisage argument explicitation based on them. However, we do not exclude the explicitation based on other models, such as the seven-step argument analysis approach of Scriven [47].

Toulmin Model-based Explicitation In research on argument analysis, one of the most well-known models for arguments is the Toulmin model [54]. It was defined particularly for legal arguments, but has since proven its suitability for a wide range of arguments [26]. This model defines five types of argument components, whose identification facilitates argument understanding.

claim is the statement that the argument intends to prove, and is analogous to the conclusion in other argumentation models;
datum is a statement of a fact, or evidence that supports the claim;
warrant is a statement that provides the connection between claim and datum, facilitating the datum to support the claim;
backing is a statement that justifies why the warrant holds;
qualifier is a statement that indicates the strength of the warrant;
rebuttal is a statement of an exceptional case whose occurrence would remove the authority of the warrant.

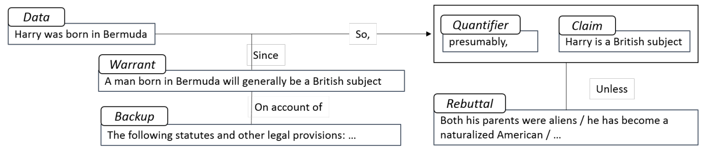

Fig. 2. Example of an argument structure following Toulmin's model

Figure 2 shows a classical example [54] of an argument modelled with the Toulmin scheme. An important challenge for explicitating natural language arguments with the Toulmin model is that most often, not all the components are present in the text. Consequently, a legitimate goal of argumentation explicitation can be to (i) *signal the lack* of specific argument components to the end user, to support her judgment of the validity of the argument, or (ii) to *identify and provide* such missing argument components from Toulmin's scheme, such as Data, Warrants or Backups, to complete the full understanding of the argument. We will come back to this discussion in Section 3.2.

As it can be seen in the example of Figure 2, and as discussed in detail by Freeman [19], warrants often take the form of generalization rules, that are often not explicitly stated. For instance, the argument in Figure 2 would most likely be encountered in everyday argumentation as *Harry is presumably a British subject because he was born in Bermuda*. The availability of background knowledge such as encyclopedic knowledge (i.e., DBpedia) can be exploited in order to suggest such potential warrants. For example, even if omitted from the text, the warrant that *A man born in Bermuda will generally be a British subject* can be reconstructed by noticing (for instance, in DBpedia) that a big fraction of people born in Bermuda are British citizens. The bigger challenge is how to deal with commonsense knowledge, or more specifically, what Feeman (2008) [19] names *empirical warrants* which *ordinarily* hold, for example *Given that x has mailed the letter, one may take it that x wants the addressee to receive it* or *If X ignited a fuse connected to a bomb, X intended to explode the bomb.*

Walton Schemes-based Explicitation Walton proposed about 50 argumentation schemes [59] organized in a hierarchy. These schemes represent common patterns of everyday reasoning, and Fig. 3 shows two of them.

There have already been a number of attempts to classify natural language arguments into Walton argumentation schemes, as well as their components, some of which are purely theoretical [10, 58] while others implement feature-based supervised classification models [17, 29, 30]. An example of an argument from verbal classification, originally published by Lawrence and Reed (2016) [30] is shown in Fig. 4. In this example, the argument text is annotated with the two premises and the conclusion. We use this example to pinpoint two important challenges, besides the actual classification of the arguments based on their Walton scheme.

Argument from Analogy
Premise 1: Generally, case C1 is similar to case C2.
Premise 2: A is true (or false) in case C1.
Conclusion: A is true (or false) in case C2.

Argument from Verbal Classification
Premise 1: a has property P
Premise 2: for all x, if x has property P, x can be classified as having property G .
Conclusion: a has property G.

Fig. 3. Example of two Walton schemes: *Argument from Analogy* and *Argument from Verbal Classification*

First, the logical conclusion following from the two premises, is *The PowerShot SX510 has great image stabilization.* For the conclusion in Fig. 4 to be logically entailed, we must assume the further premise *Cameras with great image stabilization are fantastic.*, which is implied by the text, but is not stated.

Example. *The PowerShot SX510 is a fantastic camera. It is made by Canon and all Canon cameras have great image stabilisation.*

Premise 1: It is made by Canon
Premise 2: all Canon cameras have great image stabilisation
Conclusion: The PowerShot SX510 is a fantastic camera.

Fig. 4. Example of an instance of Argument from Verbal Classification.

Second, in real life, the above argument would likely omit ***Premise 1:*** *It is made by Canon*, and the text would sound closer to *The PowerShot SX510 is a fantastic camera as Canon cameras have great image stabilization.* This adds another level of complexity and challenge to correctly classify the argument as an Argument from Verbal Classification.

A thorough explicitation of this argument that addresses both challenges is illustrated in Fig. 5. As illustrated, the argumentative text that contains only two explicit statements (Premise 2 and the Conclusion), actually implies a chain of two arguments, where the conclusion of the first serves as a premise to the second. In order to obtain such explicitations automatically, it is not sufficient to classify arguments into their corresponding Walton scheme. In addition, the classification of the components (premises and conclusions) is required, and even more challenging, the classification of the schema variables. Given the Argument from Verbal Classification scheme in Fig. 3, the classification of variables for the text *The PowerShot XS510 is a fantastic camera as all Canon cameras have great image stabilization* would be: {*The PowerShot XS510*: **a**, *fantastic camera*: **G**, *Canon cameras*: **P**, *great image stabilization*: **G**}. This classification, would then clarify which are the bits of knowledge that are needed for reconstructing the argument in such a way that it follows the Argument from Verbal

Example. *The PowerShot SX510 is a fantastic camera as all Canon cameras have great image stabilisation.*

First Argument
Unstated Premise 1: *The PowerShot SX510 is a Canon camera*
Premise 2: all Canon cameras have great image stabilisation
Unstated Conclusion: *The PowerShot SX510 has great image stabilisation.*

Second Argument
Unstated Premise 1: *The PowerShot SX510 is a camera and has great image stabilisation.*
Unstated Premise 2: *Cameras with great image stabilisation are fantastic.*
Conclusion: The PowerShot SX510 is a fantastic camera.

Fig. 5. Example of explicitation that includes analysis as well as reconstruction of an instance of Argument from Verbal Classification. The reconstruction makes explicit two arguments following the same scheme of Argument from Verbal Classification. One premise of the second argument is the conjunction of a premise and the conclusion of the first argument. The unstated components are written in Italics.

Classification scheme. Specifically, that **a** (*The PowerShot XS510*) must have property **P** (*Canon cameras*), resulting into Unstated Premise 1. We highlight here the opportunity for using structured knowledge bases that are available on the Web of Data to fill in such generalizing premises. Next, having two distinct strings serving the same role of **G** (*great image stabilization* and *fantastic camera*) can indicate that the author of the argument implies that there is a logical entailment between the two strings, leading to Unstated Premise 2. In the following, we discuss explicitations whose role is to fill in unstated premises.

3.2 Explicitation based on Enthymeme Reconstruction

Arguments with omitted premises are called *enthymemes*. They have been debated in philosophical literature since Aristotle [16,21,22,25,31,59,60]. Regarding our task of argument explicitation, dealing with enthymemes is one of the core challenges. Although explicitation based on Toulmin's model or Walton schemes may be regarded as a tangible aim as long as the problem of implied premises is ignored, we argue that most (informal) natural language arguments are enthymemes, and their explicitation, which includes reconstruction, should not be neglected. In Section 3.1, we provided some hints on how Walton schemes might be used to explicitate enthymemes, while in Section 3.1 we discussed Freeman's (2008) [19] claim that when modelling arguments with the Toulmin model, it is very common that the warrant is implied and omitted. We therefore consider explicitation based on enthymeme reconstruction as a form of explicitation that complements and deepens other types of explicitation proposed above.

The problem of enthymeme reconstruction is arguably an AI complete problem. Broadly, a system tackling enthymeme reconstruction – called an enthymeme machine [59] – must be able to answer three questions: (i) is the analyzed argument an enthymeme? (ii) which are the gaps that need to be filled? (iii) which are the missing premises? Approaches for addressing questions (i) and (ii) depend on the chosen argument model (e.g., Walton scheme or Toulmin model). Addressing question (iii) is more challenging and actually brings us to the question of the actual purpose or use cases of the task. If the purpose of enthymeme reconstruction is to support the user in judging arguments, we can relax the requirement of stating *the missing premise*. We may instead just ask the system to present a *possible* premise. For instance, reconsidering the example in Fig. 5, instead of generating Unstated Premise 1 *The PowerShot SX510 is a Canon camera* and Unstated Premise 2 *Cameras with great image stabilization are fantastic*, the system would draw the attention of the user to consider some highlighted piece of inserted information that *could* form a coherent argument, e.g., (i) **The Powershot SX510** has the property **Canon camera** and (ii) **great image stabilization** implies **fantastic camera**. This way, it is the user's responsibility to validate the argument, while the system guides this process.

If, however, the purpose of the system is to provide a *true and valid* missing premise, the system must be able to check whether these premises state true facts, e.g., they may be validated against a knowledge base, or they can be flagged as subjective statements. In our example from Fig. 5, the system would search for relations holding between *The Powershot SX510* and *Canon cameras* in a knowledge base, and judge whether the found relation is similar to the relation required by the argument scheme: ***The Powershot SX510*** *has the property* ***Canon camera***. Validating the second unstated premise in our example, by contrast, should be impossible, since it is a subjective statement, not a fact. In such a case, the system might reconstruct a possible premise (***great image stabilization*** *implies* ***fantastic camera***), and flag it as subjective.

We conclude that the system must be able to distinguish between missing premises that are *subjective* as opposed to those that are *facts*. While subjective ones can be flagged as such, using state of the art opinion detection tools, reconstructing facts involves fact checking. This can only be achieved with respect to real-world knowledge available to the system. Such real-world knowledge can be: (i) encyclopedic (e.g., *The Powershot SX510 is made by Canon*) which is available online through Wikipedia and related structured knowledge bases such as DBpedia, Wikidata, Yago; (ii) ontological (e.g., *frogs and toads are animal life*) which is available for instance through taxonomies and lexicons such as WordNet, as well as Wikipedia-based knowledge bases; (iii) common sense knowledge (e.g., *dogs usually bark when strangers enter their space*), which is much harder to source and (iv) contextual, such as the purpose of the document, the author, the time, etc. While the first two types of real-world knowledge can be accessed with state-of-the-art entity linking tools, the last two types of knowledge are more challenging, and in general much less researched. Regarding commonsense knowledge, the recent study of Becker et. al (2016) [4] finds that a large majority of commonsense relations captured by implicit unstated statements in arguments can be mapped to ConceptNet [48] relations.

With respect to contextual knowledge, Green (2010) [23] provides evidence that knowledge needed for explicitating enthymemes can often be found in the surrounding context, meta-data about authors and the targeted audience, etc.

3.3 Acceptability-based Explicitation

The previously proposed types of argument explicitation focus solely on the internal structure of the argument. However, everyday arguments rarely occur in isolation or remain unchallenged. A defining property of everyday arguments is precisely their defeasible nature, i.e., their vulnerability to being attacked by other arguments. The ability of arguments to resist such counterarguments has been named *acceptability* [15].

Acceptability-based explicitation aims to expose the relations holding between the targeted argument and other arguments, weaving a macro structure of argumentation. This type of argumentation analysis, whose target are the relations between arguments, has been researched within the context of abstract argumentation frameworks. One of the first and best studied abstract argumentation frameworks was introduced by Dung (1995) [15]. It defines only one type of relation between arguments, that of *attack* or *defeat*. Dung [15] defines a set of arguments as *acceptable* (by a rational agent), if it can defend itself against all attacks on it. More recent lines of work on argumentation frameworks extend Dung's framework by defining two types of relations between arguments, *attack* and *support* [12, 13]. Drawing inspiration from these frameworks, much of the recent computational linguistic analysis of arguments has focused on automated support/attack relation classification between pairs of arguments [6, 11, 20].

Much of the research on argument analysis considers attack and support relations to exist within a single argumentative text [36, 40, 41, 50, 51]. This is often the case in everyday argumentation, in a rhetorical technique for displaying the argument's ability to defend itself against predictable counter arguments. In order to disentangle the argumentative text in such a way as to explicitate the acceptability of its arguments, one challenge is to identify and extract the *atomic arguments*: (i) *the main argument* - the one whose conclusion is the main conclusion of the text, (ii) *the supporting arguments* - sub-arguments whose conclusions act as premises to the main argument and (iii) (anticipated) *counterarguments* - arguments that attack the main argument. Our intuition is that counterarguments are indicated by what seems like attack relations between premises of the same argument.

Fig. 6 illustrates an explicitation of an argumentative text adapted from the Microtexts of Peldszus and Stede(2015) [40], by isolating two atomic arguments – the main argument and the anticipated counterargument. As shown in Fig. 6, a counterargument can be anticipated and defeated, hence increasing the acceptability of the main argument. In our example, the premise of the main argument attacks the ability of the counterargument's premise to entail the implicit conclusion (since reported relief of complaints is not a scientific proof).

We envisage two levels of acceptability-based explicitation: (i) a shallow explicitation in which an attack or support relation is indicated between pairs

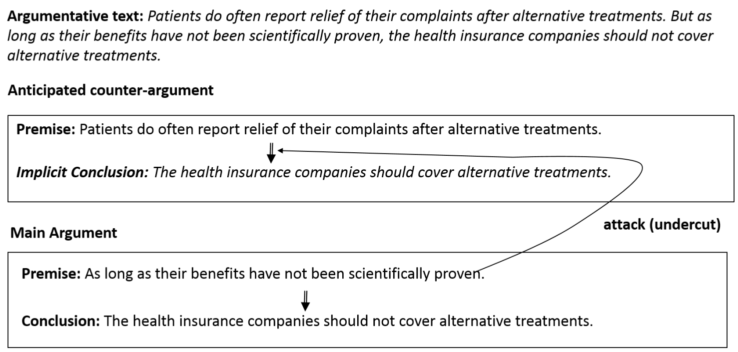

Fig. 6. Example of argumentative text containing attacking statements that are shown to belong to two different arguments.

of arguments and (ii) a deep explicitation in which the particular components (statements) participating in the relation are highlighted. Pollock [42] identifies two common types of attack relations: *rebuttals*, which directly attack the conclusion of an argument, and *undercuts*, which attack the logical entailment of the conclusion given the premise. From this perspective, in Fig. 6, the attack relation between the premise of the main argument and the anticipated counter-argument is an undercut. Acceptability-based explicitation is complementary to the previously defined types of explicitation: the identified individual arguments can be further explicitated with other types of explicitation.

3.4 Knowledge Enhancement-based Explicitation

The last type of explanation that we propose is knowledge enhancement-based explicitation, which provides additional background information about the entities and concepts mentioned in the argument's text, as well as the relations between them. The idea is to activate knowledge which is needed to understand the content of the argument components and how they are linked semantically. Consider the following argumentative text example: *Acetylsalicylic acid helps in case of a myocardial infarct as it reduces the platelet adhesion.*

A potential explicitation of this example for the lay person would be to add background knowledge in the form of additional statements such as *Acetylsalicylic acid is the active ingredient in Aspirin.*, or *Myocardial infarct is another term for Heart Attack.*, or *Reducing the platelet adhesion prevents blood clotting.* A medical doctor would most likely not benefit from this type of explicitation. Instead, they may be interested to know why the prior doctor has preferred *Acetylsalicylic acid* over alternative treatments, etc. Therefore, the challenge for this type of explicitation is to determine what information should be added. This type of explicitation therefore lends itself most naturally to personalization.

Knowledge enhancement-based explicitation bares some similarities to enthymeme reconstruction, but differs from it in that the provided knowledge statements do not need to be premises. Thus, this type of explicitation does not require any argumentation knowledge. Nonetheless, we expect the extracted knowledge to oftentimes contain the premises required for enthymeme reconstruction and hence provide satisfactory explanations for the end-user. Still, we want to underline the less constrained nature of the knowledge presented in knowledge enhancement-based explicitation, and that while this step might help the user make sense of the argument, it does not reveal how the reasoning behind the argument works.

4 A Framework for Argument Explicitation

In this section, we propose a framework for argument explicitation that considers the presented explicitation facets, as well as how they relate to each other. The framework is illustrated in Figure 7. Given an argumentative text, the first steps towards its explicitation are (i) to enhance it with background knowledge (step K), by retrieving entities and relations that are relevant to the argument from external knowledge bases, and (ii) the identification of the atomic arguments and counterarguments (step A). The extracted background knowledge can assist the acceptability-based explicitation of the argument. For instance, recent work in Kobbe et al. [27] uses DBpedia and ConceptNet in order to classify support/attack relations between argumentative statements.

Once the atomic arguments are identified, the argument explicitation system can proceed to explicitate the argument based on the model(s) of choice. The first and minimal step in this direction is to detect the argumentative units and classify them as premise or conclusion. A more elaborate explicitation is to identify the Toulmin model elements in each argument, as well as their Walton scheme. These two tasks can support each other since in some Walton schemes, the premises can be mapped to either data or warrant elements in the Toulmin model. Furthermore, as discussed earlier, the relevant background knowledge can provide valuable insights for the classification of Walton schemes or Toulmin model elements. Lastly, after each identified argument has been explicitated based on the chosen model(s), the explicitation machine can proceed with enthymeme reconstruction (step E). This step brings further detail into the model-based explicitations by filling in the blank slots of the identified models, and can further explicitate the acceptability of the main argument.

5 Discussion and Implications

In this paper, we introduce the notion of *argument explicitation* as an overarching task that makes the reasoning mechanisms required for understanding natural language arguments explicit to the end-user. The perspective we take in this work is to analyze the very diverse research directions in argumentation from the same viewpoint: that of *explaining arguments*, and to integrate these different research contributions in a common Framework of Argumentation Explicitation.

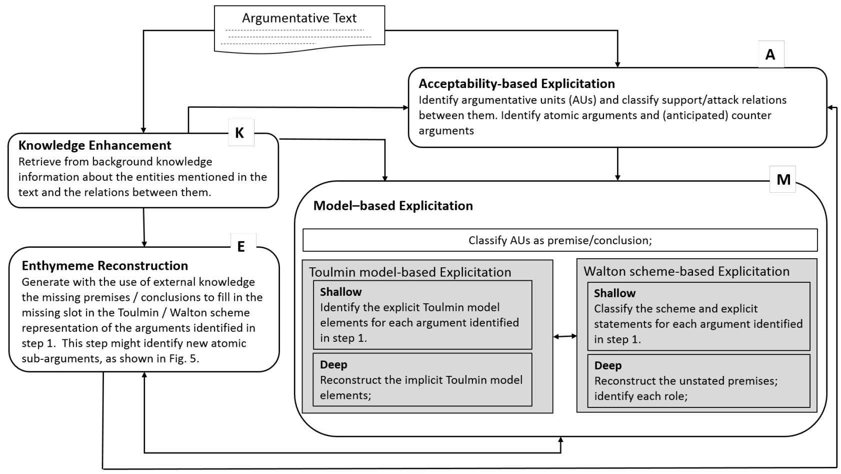

Fig. 7. Proposed Framework for Argument Explicitation

In doing so, we are able to identify the research challenges and opportunities that lie ahead. We are summarizing the most important implications of our analysis:

(i.) For uncovering the reasoning behind arguments, it is of great importance to apply and improve formal argument structure analysis, following *detailed, content-driven argument schemes* such as Toulmin's [54] or Walton's [59] schemes.

(ii.) Throughout the paper we stress and exemplify the importance of extending argument analysis with *enthymeme reconstruction*, by completing arguments with implicit argument components. This requires access to different types of knowledge that may support and validate a given argument in terms of linguistic, encyclopedic or commonsense knowledge. Clearly, this is a challenging aim. Strong NLP and AI capabilities are required in order to fully assess the explicit meaning of a given argument. Strong *reasoning capacities* are needed to be able to *select appropriate knowledge* and to verify the enriched argument to ensure its *validity and soundness* – or else to uncover inconsistencies that are revealed by assuming further information.

(iii.) Besides *appropriate repositories of background or domain knowledge*, alternative ways of identifying relevant knowledge need to be considered, such as *link prediction methods* and *on-the-fly knowledge retrieval* from textual sources, to make implicit assumptions in the NL argument (structure)s explicit.

(iv.) To support this process, *machine reasoning techniques* should be used to enforce high-level constraints over argumentation models, as well as for detecting inconsistencies in content or argument structures.

(v.) Real life arguments are rarely isolated, as they are most often part of debates. In this context, arguments should be treated as belonging to sets of arguments (following Dung). The retrieval of an *assembly of further supporting or defeating arguments* from additional sources should be considered, to facilitate the judgement of the validity or generality of an argument from a more global perspective.

(vi.) Since the reconstruction of argument components can be highly subjective, the *explicitation of reconstructed knowledge* can be realized e.g. by way of *natural language generation techniques*, to allow end users identify what additional assumptions have been made to support the conclusion. This is especially relevant for argumentation machines, but may also serve humans to fully understand the logics and possible background assumptions of an argument.

While most of the above considerations have been discussed in the theoretical literature, they constitute true challenges to computational treatments of argumentation and need to be addressed in a step-wise fashion.

Acknowledgments

This work has been funded by the Deutsche Forschungsgemeinschaft (DFG) within the project ExpLAIN, Grant Number STU 266/14-1 and FR 1707/-4-1, as part of the Priority Program "Robust Argumentation Machines (RATIO)" (SPP-1999) as well as by the Leibniz Science Campus Empirical Linguistics & Computational Language Modeling, supported by Leibniz Association grant no. SAS2015-IDS-LWC and by the Ministry of Science, Research, and Art of Baden-Wurttemberg.

References

1. Alvarado, S., Dyer, M.G., Flowers, M.: Editorial comprehension in oped through argument units. UCLA Computer Science Department (1986)
2. Amgoud, L., Prade, H.: Using arguments for making and explaining decisions. Artificial Intelligence **173**(3-4), 413–436 (2009)
3. Becker, M., Palmer, A., Frank, A.: Clause types and argumentative texts. In: ArgMining Workshop (2016)
4. Becker, M., Staniek, M., Nastase, V., Frank, A.: Enriching Argumentative Texts with Implicit Knowledge. In: NLDB, LNCS. pp. 21–30 (2016)
5. Birnbaum, L., Flowers, M., McGuire, R.: Towards an ai model of argumentation. In: AAAI. pp. 313–315. AAAI Press (1980)
6. Boltuzic, F., Snajder, J.: Back up your stance: Recognizing arguments in online discussions. In: ArgMining@ACL. pp. 49–58 (2014)
7. Boltuzic, F., Šnajder, J.: Fill the gap! analyzing implicit premises between claims from online debates. In: ArgMining Workshop. pp. 124–133 (2016)
8. Botschen, T., Sorokin, D., Gurevych, I.: Frame-and entity-based knowledge for common-sense argumentative reasoning. In: ArgMining Workshop. pp. 90–96 (2018)
9. Brem, S.K., Rips, L.J.: Explanation and evidence in informal argument. Cognitive science **24**(4), 573–604 (2000)
10. Cabrio, E., Tonelli, S., Villata, S.: From discourse analysis to argumentation schemes and back: Relations and differences. In: International Workshop on Computational Logic in Multi-Agent Systems. pp. 1–17. Springer (2013)

11. Cabrio, E., Villata, S.: Combining textual entailment and argumentation theory for supporting online debates interactions. In: ACL. pp. 208–212 (2012)
12. Cayrol, C., Lagasquie-Schiex, M.C.: On the acceptability of arguments in bipolar argumentation frameworks. In: ECSQARU, Barcelona. pp. 378–389 (2005)
13. Cayrol, C., Lagasquie-Schiex, M.C.: Bipolarity in argumentation graphs: Towards a better understanding. International Journal of Approximate Reasoning **54**(7), pp–876 (2013)
14. Chesñevar, C.I., Maguitman, A.G., Loui, R.P.: Logical models of argument. ACM Computing Surveys (CSUR) **32**(4), 337–383 (2000)
15. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial intelligence **77**(2), 321–357 (1995)
16. Ennis, R.H.: Identifying implicit assumptions. Synthese **51**(1), 61–86 (1982)
17. Feng, V.W., Hirst, G.: Classifying arguments by scheme. In: ACL. pp. 987–996 (2011)
18. Florou, E., Konstantopoulos, S., Kukurikos, A., Karampiperis, P.: Argument extraction for supporting public policy formulation. In: LaTeCH (2013)
19. Freeman, J.B.: Argument strength, the toulmin model, and ampliative probability. Informal Logic **26**(1), 25–40 (2008)
20. Ghosh, D., Muresan, S., Wacholder, N., Aakhus, M., Mitsui, M.: Analyzing argumentative discourse units in online interactions. In: ArgMining Workshop. pp. 39–48 (2014)
21. Gough, J., Tindale, C.: 'missing' premises. Informal Logic **7**(2), 99 (1985)
22. Govier, T.: Analogies and missing premises. Informal Logic **11**(3) (1989)
23. Green, N.L.: Representation of argumentation in text with rhetorical structure theory. Argumentation **24**(2), 181–196 (2010)
24. Habernal, I., Gurevych, I.: What makes a convincing argument? empirical analysis and detecting attributes of convincingness in web argumentation. In: EMNLP. pp. 1214–1223 (2016)
25. Hitchcock, D.: Enthymematic arguments. In: Argumentation: Across the lines of discipline. Proceedings of the Conference on Argumentation. pp. 289–298 (1987)
26. Hitchcock, D., Verheij, B.: Arguing on the Toulmin model. Springer (2006)
27. Kobbe, J., Opitz, J., Becker, M., Hulpus, I., Stuckenschmidt, H., Frank, A.: Exploiting Background Knowledge for Argumentative Relation Classification. In: LDK. vol. 70, pp. 8:1–8:14. Dagstuhl, Germany (2019)
28. Kuhn, D., Pearsall, S.: Developmental origins of scientific thinking. Journal of cognition and Development **1**(1), 113–129 (2000)
29. Lawrence, J., Reed, C.: Combining argument mining techniques. In: ArgMining Workshop. pp. 127–136 (2015)
30. Lawrence, J., Reed, C.: Argument mining using argumentation scheme structures. In: COMMA. pp. 379–390 (2016)
31. Levi, D.S.: The case of the missing premise. Informal Logic **17**(1) (1995)
32. Levy, R., Bilu, Y., Hershcovich, D., Aharoni, E., Slonim, N.: Context dependent claim detection. In: COLING. pp. 1489–1500 (2014)
33. Lugini, L., Litman, D.: Argument component classification for classroom discussions. In: Workshop on Argument Mining. pp. 57–67 (2018)
34. Mayes, G.R.: Argument explanation complementarity and the structure of informal reasoning. Informal Logic **30**(1), 92–111 (2010)
35. Moens, M.F., Boiy, E., Palau, R.M., Reed, C.: Automatic detection of arguments in legal texts. In: Proceedings of ICAIL 2007 (2007)

36. Nguyen, H.N., Litman, D.J.: Context-aware argumentative relation mining. In: ACL. pp. 1127–1137 (2016)
37. Opitz, J., Frank, A.: Dissecting content and context in argumentative relation analysis. arXiv preprint arXiv:1906.03338 (2019)
38. Osborne, J.F., Patterson, A.: Scientific argument and explanation: A necessary distinction? Science Education **95**(4), 627–638 (2011)
39. Peldszus, A., Stede, M.: From argument diagrams to argumentation mining in texts: A survey. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI) **7**, 1–31 (2013)
40. Peldszus, A., Stede, M.: Joint prediction in mst-style discourse parsing for argumentation mining. In: EMNLP. pp. 938–948 (2015)
41. Persing, I., Ng, V.: End-to-end argumentation mining in student essays. In: HLT-NAACL. pp. 1384–1394 (2016)
42. Pollock, J.L.: Defeasible reasoning. Cognitive science **11**(4), 481–518 (1987)
43. Rago, A., Cocarascu, O., Toni, F.: Argumentation-based recommendations: Fantastic explanations and how to find them. In: IJCAI. pp. 1949–1955 (2018)
44. Rinott, R., Dankin, L., Perez, C.A., Khapra, M.M., Aharoni, E., Slonim, N.: Show me your evidence - an automatic method for context dependent evidence detection. In: EMNLP. pp. 440–450 (2015)
45. Schank, R.C., Abelson, R.P.: Scripts, plans, and knowledge. In: IJCAI. pp. 151–157 (1975)
46. Schank, R.: Explanation patterns: Understanding mechanically and creatively. Psychology Press (2013)
47. Scriven, M.: Reasoning (1976)
48. Speer, R., Chin, J., Havasi, C.: ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In: AAAI (2017)
49. Stab, C., Gurevych, I.: Annotating argument components and relations in persuasive essays. In: COLING. pp. 1501–1510 (2014)
50. Stab, C., Gurevych, I.: Identifying argumentative discourse structures in persuasive essays. In: EMNLP. pp. 46–56 (2014)
51. Stab, C., Gurevych, I.: Parsing argumentation structures in persuasive essays. Computational Linguistics **43**, 619–659 (2017)
52. Teufel, S.: Argumentative zoning : Information extraction from scientific text (1999), doctoral dissertation, University of Edinburgh
53. Thagard, P.: Explanatory coherence. Behavioral and brain sciences **12**(3), 435–467 (1989)
54. Toulmin, S.E.: The uses of argument. Cambridge university press (2003)
55. Vreeswijk, G.: Reasoning with defeasible arguments: Examples and applications. In: JELIA. pp. 189–211. Springer (1992)
56. Wachsmuth, H., Stede, M., El Baff, R., Al Khatib, K., Skeppstedt, M., Stein, B.: Argumentation synthesis following rhetorical strategies. In: COLING. pp. 3753–3765 (2018)
57. Wachsmuth, H., Stein, B., Hirst, G., Prabhakaran, V., Bilu, Y., Hou, Y., Naderi, N., Thijm, T.A.: Computational argumentation quality assessment in natural language. In: EACL. pp. 176–187 (2017)
58. Walton, D.: Argument mining by applying argumentation schemes (2012)
59. Walton, D., Reed, C.A.: Argumentation schemes and enthymemes. Synthese **145**(3), 339–370 (2005)
60. Walton, D.N.: Enthymemes, common knowledge, and plausible inference. Philosophy & rhetoric **34**(2), 93–112 (2001)