Toward Artificial Argumentation

Atkinson, Katie; Baroni, Pietro; Giacomin, Massimiliano; Hunter, Anthony; Prakken, Henry; Reed, Chris

Published in:
AI Magazine

DOI:
10.1609/aimag.v38i3.2704

Publication date:
2017

Document Version
Peer reviewed version

Citation for published version (APA):
Atkinson, K., Baroni, P., Giacomin, M., Hunter, A., Prakken, H., Reed, C., Simari, G., Thimm, M., & Villata, S. (2017). Toward Artificial Argumentation. *AI Magazine*, *38*(3), 25-36. https://doi.org/10.1609/aimag.v38i3.2704

Towards Artificial Argumentation

Katie Atkinson[1], Pietro Baroni[2], Massimiliano Giacomin[2], Anthony Hunter[3], Henry Prakken[4], Chris Reed[5], Guillermo Simari[6], Matthias Thimm[7], and Serena Villata[8]

[1]Department of Computer Science, University of Liverpool, UK
[2]Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Brescia, Italy
[3]Department of Computer Science, University College London, UK
[4]Department of Computer Science, Universiteit Utrecht, The Netherlands
[5]Department of Computer Science, University of Dundee, UK
[6]Department of Computer Science & Engineeering, Universidad Nacional del Sur, Argentina
[7]Institute for Web Science and Technologies, Universität Koblenz-Landau, Germany
[8]Laboratoire I3S, INRIA Sophia Antipolis, France

January 17, 2017

Abstract

The field of computational models of argument is emerging as an important aspect of artificial intelligence research. The reason for this is based on the recognition that if we are to develop robust intelligent systems, then it is imperative that they can handle incomplete and inconsistent information in a way that somehow emulates the way humans tackle such a complex task. And one of the key ways that humans do this is to use argumentation — either internally, by evaluating arguments and counterarguments — or externally, by for instance entering into a discussion or debate where arguments are exchanged. As we report in this review, recent developments in the field are leading to technology for artificial argumentation, in the legal, medical, and e-government domains, and interesting tools for argument mining, for debating technologies, and for argumentation solvers are emerging.

1 Introduction

Humans argue[1]. This distinctive feature is at the same time an important cognitive capacity and a powerful social phenomenon. It has attracted attention and careful analysis since the dawn of civilization, being intimately related to the origin of any form of social organization, from political debates to law, and of structured thinking, from philosophy to science and arts.

As a cognitive capacity, argumentation is important for handling conflicting beliefs, assumptions, viewpoints, opinions, goals, and many other kinds of mental attitudes. When we are faced with a situation where we find that our information is incomplete or inconsistent, we often resort to the use of arguments in favor and against a given position in order to make sense of the situation. When we interact with other people we often exchange arguments in a cooperative or competitive fashion to reach a final agreement and/or to defend and promote an individual position.

Occurring continuously both in our mind and in the social arena, argumentation pervades our intelligent behavior and the challenge of developing artificial argumentation systems appears to be as diverse and exciting as the challenge of artificial intelligence itself.

Indeed, this rich and important phenomenon offers an opportunity to develop models and tools for argumentation and to conceive autonomous artificial agents that can exploit these models and tools in the cognitive tasks they are required to carry out. To this purpose, a number of interesting lines of research are being investigated within artificial intelligence, and several neighbor

[1] "Humans argue" is a truism. Either you already believe it or you would need to argue against it.

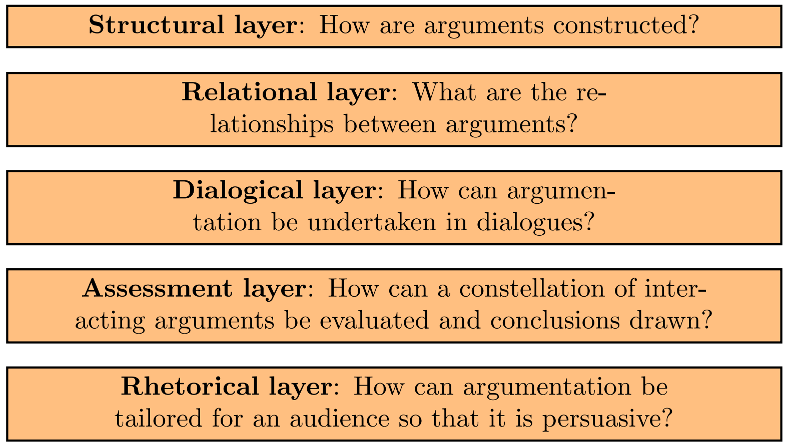

Figure 1: Key aspects of argumentation.

fields, leading to the establishment of computational models of argument as a promising interdisciplinary research area. Progress in this area is expected to contribute to significant advances in the understanding and modelling of various aspects of human intelligence.

In this article, we review formalisms for capturing various aspects of argumentation, and we present advances in their applications, with the aim to communicate how research is making progress towards the goal of making artificial argumentation technologies and systems a mature and widespread reality. In this brief review, we are unable to discuss or cite all the relevant literature, and we suggest that the interested reader seeks more detailed coverage of the foundations from [RS09], of applications from [MTB+13], and of recent developments from the proceedings of the International Conference on Computational Models of Argument series[2], and the Argument and Computation journal[3].

2 Models of argument

Computational models of argument are being developed to reflect aspects of how humans build, exchange, analyse, and use arguments in their daily life to deal with a world where the information may be controversial, incomplete, or inconsistent [BD07, RS09]. The diversity of the manifestations of arguments in real life implies diversity in the relevant models too and the impossibility to reduce the vast available literature to a single reference scheme. It is possible however to identify some *layers* which can be regarded as basic building blocks for the construction of an argumentation model. Specific modelling approaches may differ in the selection of which layers they actually use, in the way the selected layers are combined, and in the formalization adopted within each layer.

We consider the following five main layers: structural, relational, dialogical, assessment and rhetorical. They are described in the following and also summarized in Figure 1. Note that while each layer has its own nature and distinctive traits, the boundaries between layers may not be so neat in some contexts and specific formalisms may inextricably merge together aspects relevant to different layers.

Structural layer. This layer concerns the structure of the arguments and how they are built: essentially it specifies, in a given context, what an argument looks like, in terms of its internal structure, and which are the ingredients for its construction. To exemplify, in contexts where arguments are built from a logical knowledge base the ingredients are the logical formulae included in the knowledge base. Then one way to build arguments is by simply applying the logic of the language in which the knowledge base is stated to derive conclusions. An argument here can

[2] `http://comma.csc.liv.ac.uk/`
[3] `http://www.iospress.nl/journal/argument-computation/`

be seen as a pair $\langle \Phi, \alpha \rangle$ where Φ is a subset of the knowledge base (a set of formulae) that logically entails α (a formula). Here, Φ is called the support, and α is the claim, of the argument. Other approaches consider argument construction from knowledge bases as applying rules to the formulas from the knowledge base, where the rules may be defeasible. In these rule-based approaches an argument is typically seen as a tree whose root is the claim or conclusion, whose leaves are the premises on which the argument is based, and whose structure corresponds to the application of the rules from the premises to the conclusion. Investigations into the structural layer were initiated by Pollock [Pol92]. Prominent examples of rule-based formalisms are ASPIC+, assumption-based argumentation (ABA) and Defeasible Logic Programming (DeLP). For a tutorial introduction to formalisms for structured argumentation, see [BJH$^+$14].

Arguments are not built from knowledge bases only however. For instance, interactive systems acquire arguments from users may adopt the approach of argumentation schemes [WRM08], namely stereotypical reasoning patterns, where in addition to the premises and the claim, a set of critical questions is considered. Critical questions provide a sort of checklist of issues that can be raised to challenge arguments built on the basis of a given scheme. Argumentation schemes have also been used as a source of defeasible inference rules in rule-based approaches to argument construction from knowledge bases. In addition, argumentation schemes are often considered in the context of argument mining (see Section 6) where the goal is to identify and extract the argumentative structures embedded in a natural language source, providing a machine-processable representation of them.

The variety of existing argument models raises the issue of exchanging or sharing arguments among different systems. This problem is addressed by the Argument Interchange Format initiative [CMM$^+$06], aimed at providing an interlingua between various more concrete argumentation languages, on the basis of a generic abstract ontology.

Relational layer. Arguments do not live in isolation and are linked to each other by various types of relations: the relational layer deals with identifying and formally representing them, in view of their use in other layers or even for descriptive and presentation purposes, since they are essential for an understanding of what is actually going on in an argumentation process. Examples of important relationships are:

- the subargument/superargument relationships, indicating how an argument is built incrementally on top of other arguments;
- the attack relationship, indicating that an argument is incompatible with another argument in some sense, e.g. because they have contradictory claims, or one claim contradicts some premise or assumption on which the other is based.
- the support relationship, intuitively meaning that an argument provides some backing to another, and admitting several, even rather dissimilar, interpretations, depending on the actual nature of this backing;
- a preference relationship, ranking arguments according to some criterion, and admitting again a variety of instantiations ranging from strength to credibility to value-based evaluations.

What relationships are significant and how to identify them are highly context-dependent matters. Note in particular that identifying argument relations may be an easy mechanical procedure in settings where arguments are formally built from a knowledge base, while in an argument mining scenario it is a task as challenging as the identification of the arguments themselves.

Dialogical layer. This layer deals with the exchange of arguments among different agents (or even between an agent and itself, in a scenario where argumentative reasoning is conceived as a monological activity) according to formal dialogue rules. Agents may engage in the exchange of arguments for a variety of purposes with several dialogue types having been identified in the literature, like inquiry, negotiation, information-seeking, deliberation, and persuasion. In all cases the exchange can be formalized as a dialogue game, which is normally made up of a set of communicative acts called moves, and a protocol specifying which moves can be made at each step of

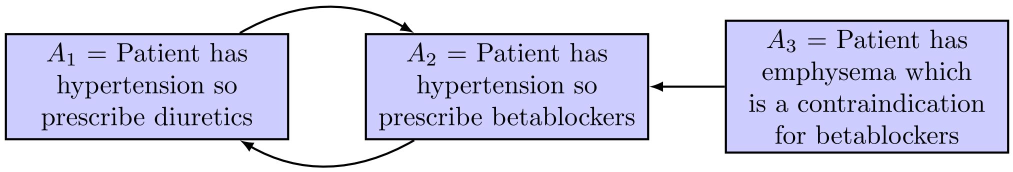

Figure 2: An example of argumentation framework consisting of three arguments in the medical domain and their attacks. Arguments A_1 and A_2 are two alternatives for treating a patient with hypertension, and A_3 provides a reason against one of the options. Here, we assume that A_1 and A_2 attack each other because giving one treatment precludes the other, and we assume that A_3 attacks A_2 because it provides a counterargument to A_2.

the dialogue. It concerns representing and managing the locutions exchanged between the agents involved, as well as specifying the contents of these locutions in terms of entire arguments or components of arguments. Moreover, the dialogue protocol may establish the allowed moves on the basis of argument relationships. For instance, a protocol may specify that a move is legal only if it presents an argument attacking an argument presented in a previous move. For these reasons the dialogical layer requires strict connections with the structural and relational layers. Moreover some dialogue protocols are defined so as to embed an argument assessment method: in these cases the dialogical layer is intertwined with the assessment layer, described next.

Assessment layer. This layer concerns the assessment of a set of arguments and of their conclusions in order to establish their justification status. The need for this layer arises from the presence of attacks among arguments, preventing them so as to be accepted altogether and calling for a formal method to solve the conflict. This problem is addressed in a principled and highly stylized form in the context of the theory of abstract argumentation frameworks [Dun95], where arguments are treated as abstract entities, deprived of any structural property and of all their relations but attack. We give an example of an argumentation framework, based on textual arguments, in Figure 2. Given its abstract nature, an argumentation framework is often referred to as *argument graph*, and this term is also used to refer to similar representations where additional relations, like support, are considered.

An *abstract argumentation semantics* is a formal criterion to determine which sets of arguments, called *extensions*, are able to survive the conflict together and can be regarded as collectively acceptable. Abstract argumentation theory is probably the subfield of computational models of argument that has attracted most research attention in the last two decades, due to its generality and theoretical cleanness. In particular Dung [Dun95] has shown the ability of the formalism to capture as instances several other approaches, especially in the area of non-monotonic reasoning. In general, Dung's abstract approach has been combined with models at the structural layer of argument to define attack in terms of preferences or argument strength while taking the structure of arguments into account, though this remains implicit in the abstract representation. It is, however, worth remarking that the basic assumption that the assessment of argument acceptability depends only on the attack relation in abstract terms, while reasonable in many contexts, may not be universally applicable. For this reason, other approaches, e.g. the relation of support too or other argument properties like strength or some reference values (see next layer), are available in the literature. Moreover, while most approaches consider a qualitative notion of acceptance, quantitative assessments methods are being investigated too.

Further, it must be noted that the evaluation of argument acceptability is only a part, actually the most basic one, of the assessment tasks required in an argumentative process. In particular the final goal of an agent is usually the assessment of the justification status of the statements supported by arguments, which, in the end, amounts to determine what to believe or what to do. Since many arguments may have the same conclusion, assessing the status of a statement involves a synthesis of the statuses of the arguments supporting it. As in real life, the task of deciding what to believe may be carried out adopting different attitudes, ranging from extremely skeptical to extremely credulous, corresponding to different formal methods for statement justification synthesis.

Rhetorical layer. Normally argumentation is undertaken in some wider context of goals for the agents involved, and so individual arguments are presented with some wider aim and according to some strategical considerations. For instance, if an agent is trying to persuade another agent to do something, then it is likely that some rhetorical device is harnessed and this will reflect the nature of the arguments used (e.g. a politician may refer to investing in the future of the nation's children as a way of persuading colleagues to vote for an increase in taxation). With the roots of the study of rhetoric going back to Aristotle[4], recent studies into aspects of the rhetorical level include believability and impact of arguments from the perspective of the audience, use of threats and rewards, appropriateness of advocates, and values of the audience. The rhetorical layer may be absent in some contexts, e.g. when neutrally building arguments from a knowledge base, but can permeate all the other layers in other contexts since goal-oriented considerations may drive the decisions of which arguments to build, taking into account their relations with other arguments, of whether, how, and when to use the arguments in a dialogue, and of which assessment method (e.g. a more skeptically or more credulously oriented) to apply.

The following sections review several prominent domains which exploit computational models of argument for the development of actual applications and, at the same time, stimulate the relevant theoretical development by providing case studies and important modelling challenges.

3 Legal argumentation

The law is an obvious application domain for argumentation research, since legal reasoning is essentially argumentative and to a large extent recorded in documents. This has led to highly stylized forms of argumentation, which makes it easier to formulate and validate formal and computational models of argument than in many other domains. In this section, we briefly discuss work and research themes in this area. A more detailed survey can be found in [PS15].

In legal cases, first the facts have to be determined. Because of the diverse nature of the evidence in most cases and the need for explanation to statistical laypeople, legal evidential reasoning is an excellent testbed for combined qualitative and quantitative models of defeasible reasoning. At the practical side, so-called 'sense-making systems' have been proposed, with which crime investigators or triers of fact can structure their arguments and scenarios to make sense of a large body of evidence.

After the facts of a case have been established, they must be classified under the conditions of legal rules, which involves interpreting these rules. Two influential AI & Law models of this are the HYPO system by Kevin Ashley and its successor CATO by Vincent Aleven, which model how lawyers in common-law jurisdictions make use of past decisions when arguing a case. Their underlying argumentation model is for 'factor'- or 'dimension'-based reasoning, where cases are collections of abstract fact patterns that favour or oppose a conclusion, either in an all-or nothing fashion (factors) or to varying degrees (dimensions). This work inspired subsequent formal work using the tools of formal argumentation, resulting in formalized versions of traditional legal argument forms such as appeal to precedent and policy and the balancing of goals, values and interests [HBC12].

Finally, when the facts have been classified, the legal rules must be applied to them. Legal rules can have exceptions or conflict on other grounds. Rule-based argumentation logics with preferences have proved useful here.

Legal reasoning usually takes place in the context of a dispute between adversaries, within a prescribed legal procedure. This makes the setting inherently dynamic and multi-party, and raises issues of strategy and choice. For example, there is work on optimal strategies for adversaries in debates with an adjudicator, given their preferences over the possible outcomes of a debate and their estimates of what the adjudicator will likely accept.

[4] `http://plato.stanford.edu/entries/logic-informal/`

While the theoretical advances on models of legal argument have been impressive and a number of valuable prototype systems have been developed, no systems have been deployed in everyday practice yet. One reason is the conservative attitude to technology in the legal world and its 'billing by the hour' culture, which does not stimulate innovation. Another reason is the fact that building realistic systems of legal argument requires vast amounts of commonsense knowledge. However, recently things have changed. First, clients of law firms increasingly demand the use of modern technology. Moreover, the recent advances in natural-language processing, machine learning and data science combined with the massive digital availability of legal data and information have created the prospects for combining AI models of legal argument with argument mining techniques (see Section 6). In fact, two of the first argument mining projects were on legal argument [PM11]. If this technology is combined with AI & Law's computational models of argument, then practical applications of these models could be well within reach.

4 Medical argumentation

Healthcare is a potentially important domain for developing and applying computational models of argument. It is common for healthcare information to be complex, heterogeneous, incomplete and inconsistent, and therefore argumentation is appealing for those involved as it allows for important conflicts to be highlighted and analyzed and unimportant conflicts to be suppressed.

One of the pioneers of argumentation technology, John Fox, developed a number of prototype systems for medical decision-support such as the Capsule system [WGY^{+}97]. Capsule supports a family practitioner when s/he is about to prescribe a specific drug for a patient. The system uses a standard database of "equivalent treatments" that is routinely used by clinicians, and the patient records, to provide arguments pro and con each of the alternatives. The arguments are based on whether the patient has previously expressed a preference for/against the alternative, whether the patient has previously exhibited a negative reaction to the alternative, whether there is possible negative interaction with other drugs being taken by the patient, and the relative cost of the alternative. In a formal trial of the Capsule system, with 42 clinicians using 36 simulated cases, the system was shown to help clinicians improve the quality of their prescribing and to improve their compliance with medical guidelines.

Over recent years, there has been substantial shift in healthcare to evidence-based practice. This means that healthcare professionals need to use the best available evidence to inform their decision making. For deciding on interventions, this normally calls for evidence from randomized clinical trials. The problem with this is that there are many such trials published each year, and it is difficult for clinicians to keep abreast of this literature. To help them, there are medical guidelines and systematic reviews that aggregate this evidence by providing recommendations. Unfortunately, these recommendations can rapidly become out of date, they do not take local circumstances into account, and they normally do not consider patients with comorbidities. To address these problems, an argument-based approach to aggregating clinical evidence has been proposed by Hunter and Williams [HW12]. The framework is a formal approach to synthesizing knowledge from clinical trials involving multiple outcome indicators (where an outcome indicator is either positive such as the number of patients who survive the disease after 1 year, or 2 years, etc, or negative such as the proportion of those treated who have a particular side-effect). Based on the available evidence, evidence-based arguments are generated for claiming that one treatment is superior to another for a given patient. Preference criteria over evidence-based arguments are specified in terms of the outcome indicators, and the magnitude of those outcome indicators, in the evidence. Various kinds of counter-arguments attack the evidence-based arguments depending on the quality of evidence used (e.g. evidence could be attacked because a trial was not conducted correctly). The arguments and counter-arguments constitute an argument graph, and using abstract argumentation semantics, the winning arguments are identified, and thereby argument-based recommendations for which treatments are superior can be identified. The approach has been evaluated by comparison with recommendations made in published healthcare guidelines [HW12] and it has been used to publish, in the medical literature on lung cancer, a more refined systematic review of the evidence. An ongoing study is using this technique in a systematic review on brain cancer for publication by Cochrane.

The above examples are just two of a number of applications of argumentation being developed for supporting healthcare professionals and patients. Further applications include dealing with the conflicts that can occur when using multiple clinical guidelines, supporting multi-disciplinary teams of healthcare professionals when dealing with difficult clinical cases, and supporting medical image interpretation.

5 eGovernment

An important feature of democracies is that citizens can engage their governments in dialogues about policies. Traditionally this was done by writing letters and holding town hall debates, but over the past two decades new methods of interaction have been developed to exploit the benefits of current technology. Citizens may wish to respond in several ways to policy proposals made by their governments. They may simply seek a justification of the proposed policy; they may wish to object to the proposed policy; or they may want to propose policies of their own. Such dialogues can be facilitated through tools to support e-democracy, and computational models of argument can be put to effective use in such tools.

For example, consider a local government authority that is deciding what to build on a plot of wasteland in a community. One proposal by the local authority may be to permit the building of a new supermarket on the grounds that this will provide jobs for the local community and shopping facilities for local residents. Citizens may be consulted on this proposal and critique this policy as well put forward their own proposals. For example, the local authority's proposal could be critiqued by stating that the action of building a new supermarket will not have the intended effect of creating jobs as there will be job losses from local shop owners being put out of business by the supermarket. An alternative proposal could be that instead of building a new supermarket, the site should be used to build a new play centre for local residents' children. Such opinions can be formed into arguments by distinguishing the premises (for example, there is little unemployment in the community and play centres promote social interaction) and conclusion (we should build a new children's play centre). Argumentation-based tools can then be put to use to facilitate such debates.

Simple tools like e-petitions[5] can transform traditional paper-based communication into digital communication, but recent advances have been made in the development of tools that exploit the web, such as the on-line argument mapping tools Debategraph[6] and Debatabase[7], which enable users to model debates by considering issues, and their pros and cons, relevant to a debate. With such tools users can freely insert and modify contributions, but the arguments entered are not required to conform to any particular semantics that would support coherence and argument evaluation. A comprehensive survey of the state-of-the-art in web-based argumentation tools can be found in [SGP13].

Early collaborative decision support systems such as Zeno [GK97] and HERMES [KP01] contained more structure by making use of the IBIS (Issue Based Information Systems) model of argument. Use of this model enabled a particular problem or issue to be decomposed into a number of different positions for which arguments can then be created to attack or defend the positions until the issue is settled.

In recent work to consolidate different tasks relevant for e-democracy tools, on a recent European project called IMPACT[8], an "argumentation toolbox" was created that consists of four interconnected modules: an argument reconstruction tool; a structured consultation tool; an argument visualisation tool; and, a policy modelling tool. This is a decision-support tool that makes use of structured theories of argument representation and evaluation to enable public opinion gathering

[5] `https://petition.parliament.uk/`
[6] `http://debategraph.org/`
[7] `http://idebate.org/debatabase`
[8] `http://www.policy-impact.eu/`

on political issues from which conclusions can be drawn concerning how government policies are presented, justified and viewed by the users of the system. The tool uses argumentation schemes [WRM08] to structure the information presented to the users, but within the front-end interface, this structure is implicit in order to facilitate ease of learning and use. Once opinions have been gathered, the arguments generated are evaluated through the use of value-based argumentation frameworks (see chapter in [RS09]) to provide users with support for which actions are justified, according to the facts and social interests promoted by the different arguments. Debates that have been modelled, using in particular the Parmenides system (see [RS09]), cover local issues, such as whether to introduce more speed cameras on dangerous stretches of road, and wider national issues such as the UK debate about whether to ban fox hunting. This strand of work continues to be expanded within the Structured Consultation Tool developed as part of the IMPACT project mentioned above and the Carneades tools[9].

The richness of policy debates clearly makes the domain of e-government an ideal one for the study and application of tools that use computational models of argument on a large scale.

6 Argument mining

In recent years, the growth of the web, and the rapidly increasing amount of diverse textual data published there, have highlighted the need for methods to identify, structure and summarize this huge resource. Online newspapers, blogs, online debate platforms and social networks, as well as normative and technical documents, provide a heterogeneous flow of information where natural language arguments can be identified, and analyzed. The availability of such data, together with the advances in natural language processing and machine learning, have supported the rise of a new research area called *argument mining*. The main goal of argument mining is the automated extraction of natural language arguments and their relations from generic textual corpora, with the final goal to provide machine-readable structured data for computational models of argument and reasoning engines.

Two main stages have to be considered in the typical argument mining pipeline, from the unstructured natural language documents towards structured (possibly machine-readable) data:

Argument extraction The first stage of the pipeline is to detect arguments within the input natural language texts. The retrieved arguments will thus represent the nodes in an argument graph returned by the system. This step may be further split into two different stages such as the extraction of arguments and the further detection of their boundaries. Many approaches have recently been applied to tackle this challenge adopting different methodologies like for instance support vector machines, naïve Bayes classifiers, logistic regression.

Relation extraction The second stage of the pipeline consists in constructing the argument graph to be returned as output of the system. The goal is to identify what are the relations holding between the arguments identified in the first stage. This is an extremely complex task, as it involves high-level knowledge representation and reasoning issues. The relations between the arguments may be of a heterogeneous nature, like attack, support or entailment. This stage is also responsible for identifying, in structured argumentation, the internal relations of the components of an argument, such as the connection between the premises and the claim. Being an extremely challenging task, existing approaches assume simplifying hypotheses, like the fact that evidence is always associated with a claim.

To illustrate, we consider the following example adapted from an online debate about *random sobriety tests for drivers and their consistency with human rights*[10]. We start with the unstructured natural language text from which we first aim to extract the arguments, and then to identify their relations:

[9] `http://carneades.github.io/`

[10] `http://debatepedia.idebate.org/en/index.php/Debate:_Random_alcohol_breath_tests_for_drivers`

Random breath tests to public vehicle drivers can hardly be called an invasion of privacy or an investigation without due cause, because public safety is at stake. Random tests are routinely carried out by many train and bus companies and are being introduced on airlines as well. The same applies for other drivers, who are a major liability to the safety and lives of other drivers.

Randomly testing employees cannot be considered an invasion of privacy. People who have to take random breath tests to drive trucks or fly planes as part of their jobs are taking the test as part of their job. They are being paid and must do what their employer wants them to do in order to keep their job. Searching random people outside of the context of employment with no suspicion of a crime is very different as it erodes civil liberties and sets a dangerous precedent.

The first goal of the argument mining pipeline consists in extracting the arguments from this text. In the example above, we highlight the four arguments that can be identified:

Random breath tests to public vehicle drivers can hardly be called an invasion of privacy or an investigation without due cause, because public safety is at stake. *[A_1] Random tests are routinely carried out by many train and bus companies and are being introduced on airlines as well.* ***The same applies for other drivers, who are a major liability to the safety and lives of other drivers.*** *[A_2]*

Randomly testing employees cannot be considered an invasion of privacy. *[A_3] People who have to take random breath tests to drive trucks or fly planes as part of their jobs are taking the test as part of their job. They are being paid and must do what their employer wants them to do in order to keep their job.* ***Searching random people outside of the context of employment with no suspicion of a crime is very different as it erodes civil liberties and sets a dangerous precedent.*** *[A_4]*

Given these four arguments (i.e., A_1, A_2, A_3, and A_4), the relations among them have to be identified. Let us consider for the explanatory purpose of this example that the two relations we aim at identifying are the *attack* and the *support* relations only. In this case, we have that, taking into account the temporal dimension of the debate to decide the direction of the relations, argument A_3 supports argument A_1, and argument A_4 attacks argument A_2.[11]

It is worth noticing that the identification of the arguments and their relations is much more subtle and ambiguous than what emerges from this explanatory example, and may often be a matter of interpretation that current state of the art argument mining systems cannot tackle yet. For instance, argument A_1 can be considered as a sub-argument of argument A_2 as "The same applies ..." refers to A_1, and argument A_3 can be interpreted as a kind of *persuasive statement* meant to strengthen argument A_4.

To address this kind of issues and build more capable applications, it is necessary to enhance the existing tools used to analyze, aggregate, synthesize, structure, summarize, and reasoning about arguments in texts, with more sophisticated natural language processing (NLP) methods. However, and considering the complexity of the task, to do so it is still necessary to reach a deeper level of understanding of the inner workings of natural language, and evolve new methods expanding the ones currently found in NLP.

Moreover, to tackle these challenging tasks, high-quality annotated corpora are needed, as those proposed in [RR04, PM11, APL+14, SG14], to be used as a training set for any kind of aforementioned identification. These corpora are mainly composed by three different elements: an *annotated* dataset which represents the gold standard whose annotation has been checked and validated by expert annotators and is used to train the system for the required task (i.e., arguments

[11] Argument A_2 has to be read as *[Random breath tests to] other drivers [can hardly be called an invasion of privacy or an investigation without due cause as they are] a major liability to the safety and lives of other drivers.*

or relations extraction), a set of guidelines to explain in a detailed way how the data has been annotated, and finally, the *unlabelled* raw corpus that can be used to test the system after the training phase. The reliability of a corpus is ensured by the calculation of the inter-annotator agreement that measures the degree of agreement in performing the annotation task among the involved annotators.[12] Current prototypes of argument mining systems require to be trained against the data the task is addressed to, and the construction of such annotated corpora remains among the most time consuming activities in this pipeline.

For an exhaustive state of the art review on argument mining techniques and applications, we refer the reader to [LT16].

7 Debating technologies

There is a long tradition of computer-aided debate systems with roots in e-democracy, decision support and so on. These systems have two things in common: first, they implement idiosyncratic and new dialogical structures, or games, with little re-use or incremental development. The second is that little or no contribution to the debate itself is offered by the machine. The system role has been one of support and facilitation only. With a variety of techniques for automatically mining argument structures from both monological and dialogical resources, an exciting new possibility is opened up for not just supporting and enhancing new human-human arguments but also developing new human-machine arguments: this is the space of debate technology, a specific subfield of argument technology in general.

Several systems have demonstrated stand-alone applications of debate technology focusing on domains of use such as pedagogy [PM12], in which both responsibility for the structuring of a debate and its automatic furthering are taken on by the machine. The key bottleneck in such systems, however, is the availability of data.

Many of the resources available in the largest openly accessible datasets provided by the Argument Web (searchable at aifdb.org, see [BLSR13]), also indicate argument provenance and authorship. By associating arguments with arguers, a ready-built mechanism becomes available for populating agent knowledge bases. This forms the foundation of the Arvina system, shown in Figure 2, which makes use of a general purpose platform for executing dialogue games or protocols. There are two approaches to such generalized dialogue execution which allow systems to deliver *mixed initiative argumentation* whereby humans and software agents can play one of a number of debating dialogue games on a level playing field (one in which, indeed, infrastructure may have no way of distinguishing human from software players). The first, *lightweight*, approach extends existing programming languages with a small set of communicative coordination constructs. The advantage of this approach is that systems for debating technology can be rapidly prototyped with few new concepts required. The problem is that for engineering practical systems, it leaves a very large amount of dialogical componentry to be defined by researchers and developers. The solution to this problem is offered in *rich* dialogue execution, by which all the usual components of dialogue games (participants, commitments, turn-taking, backtracking, etc.) are baked in to a rich domain-specific language, which, though less elegant than a lightweight approach, gives the developer a much more extensive language for engineering. This is the approach taken in the DGDL language used by Arvina. Though such an approach lacks the general purpose flexibility that might be hoped for from some future system capable of full NLP understanding, it provides a flexible, intuitive and naturalistic mechanism for navigating complex information spaces (such as climate change, abortion, civil liberties, and so on).

These complex information spaces are often sufficiently intricate and detailed that users – both those who have engaged in the debate and also those who are reviewing it post hoc – need additional mechanisms to make sense of the otherwise potentially overwhelming deluge of data. This has led to two distinct approaches. The first involves a range of *debate analytics* which aggregate, calculate and interrogate the argument structures created in a debate allowing insight into, for example,

[12] The number of involved annotators should be > 1 in order to allow for the calculation of this measure and, as a consequence, produce a reliable resource.

strong and weak arguments, stimulating and boring participants, central and peripheral issues and so on. The second approach focuses on *augmented debate*, adding rich streams of additional information concerning the debate presented visually and often tied to a video recording. One of the most sophisticated and rich systems of augmented debate with analytics is the EDV project, which was trialled with the 2014 UK election debates [Pluss-etal2016]. In both cases, situated, linguistic and dialogical metrics (such as dominance and relationships between speakers) are used in combination with metrics based on structured argumentation (which yield insights into the inferential structures created by participants) and those based on abstract argumentation (which can contribute to assessing debate-wide features such as which arguments are winning).

As debate technology starts to mature, we are thus seeing not only increasingly sophisticated systems for supporting and contributing to debates, but also complementary systems for making sense of the large datasets that result.

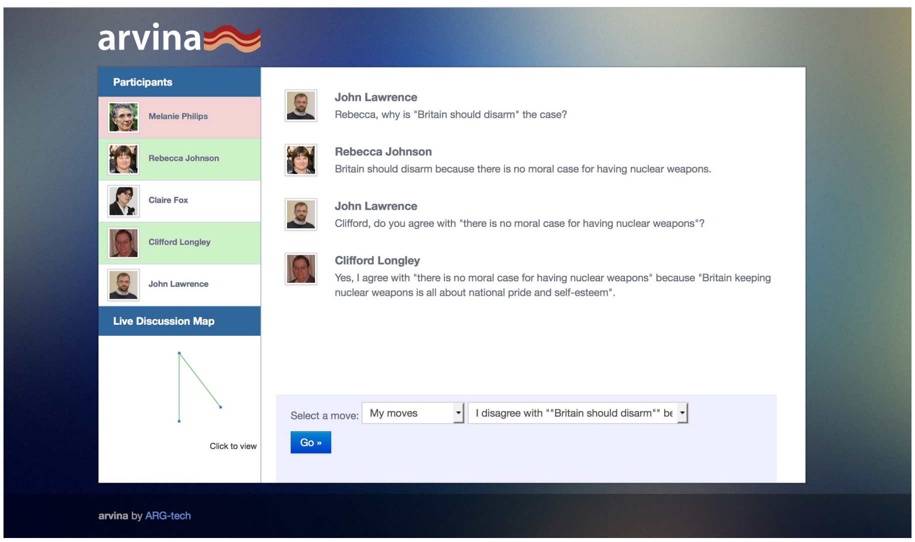

Figure 3: Arvina

8 Argumentation solvers

The scenarios outlined above require the use of effective systems for solving various problems related to computational argumentation. Recall the approach of abstract argumentation discussed in the assessment layer in Section 2 (see also Figure 2). This approach models argumentation scenarios as graphs and a central question is how to determine the *extensions* of such a graph, i. e., sets of arguments that can collectively be accepted, given the attack relation represented by the arcs. The literature offers various ways to formally define this concept of *acceptability*, but the computational problem of extracting any such a set from a given graph is usually hard to solve and can exhibit complexity beyond P and NP. For example, deciding for a given graph and a given argument, whether the argument is contained in all extensions under the so-called *preferred semantics* [Dun95] is Π_2^P-complete, which is regarded as highly infeasible. Roughly speaking, while (deterministic) algorithms to NP-hard problems usually require exponential worst-case runtime, algorithms for Π_2^P-complete problems may even exhibit super-exponential worst-case runtime. However, solutions to these kind of problems are required by applications utilising argumentative decision procedures and, recently, the community started to address these challenges by developing specific *argumentation solvers* for both abstract and structured argumentation settings.

Solvers for abstract argumentation are usually general-purpose tools similar to SAT-solvers (for a review of SAT-solvers, see [GKSS08]) and solve the computational problem of determining acceptable arguments from a given graph. Solvers for this setting usually fall into one of two categories, the *reduction-based approach* and the *direct approach* [CDG+15]. In the reduction-based approach, the problem at hand is translated into another formalism (such as SAT) and specialized solvers for that formalism are used to solve the original problem. In the direct approach, the peculiarities of abstract argumentation frameworks are exploited to directly solve the problem at hand without the use of another formalism.

Systems addressing the structured argumentation setting (see the "Structural layer" in Section 2) are additionally concerned with problems related to argument construction and defeat discovery. In many application scenarios, knowledge is represented as facts and rules or, more generally, as formulas in some logic. In order to apply argumentation technology, arguments have to be constructed by combining formulas and conflicts between different arguments have to be detected.[13,14,15,16] Many systems for structured argumentation generate argumentation graphs such as the one shown in Figure 2 as output and use abstract argumentation solvers for the actual determination of acceptable arguments. However, as actual application contexts may require the user to specify facts and rules rather than the induced arguments, systems for structured argumentation are a key element for the adoption of argumentation technology in actual applications.

In order to evaluate the state-of-the-art of argumentation solvers, the International Competition on Computational Models of Argumentation (ICCMA)[17] has been initiated in 2014 and organised its first contest in 2015 [TVC+16]. For the first contest, the focus was on problems related to abstract argumentation and solvers were evaluated based on their runtime performance for computing extensions or deciding on acceptance of arguments with respect to complete, preferred, stable, and grounded semantics of abstract argumentation, cf. [Dun95]. There were 18 participating systems and the best performing ones achieved significant improvements with respect to the state of the art. Based on these encouraging results, a second contest will be held in 2017.

9 Conclusions

Developing artificial tools capturing the human ability to argue is an ambitious research goal, and it may ultimately prove to be as difficult as developing AI in general.

As described in this paper, current research addresses a range of applications like law, medicine, eGovernment, debating, where argument-based approaches have shown to be beneficial for intelligent activities like sense-making and decision-making. These areas witness an increasing integration of proactive support and automated reasoning capabilities, complementing the functionalities offered by other useful but more passive tools like argument visualization systems.

Even more sophisticated roles for artificial argumentation tools are sought in the medium term and are the subject of recent research initiatives. For instance, there is a growing interest in developing *computational persuasion* systems able to assist people in making the right choice in their daily activities. Consider scenarios such as a doctor persuading a patient to drink less alcohol, a road safety expert persuading drivers to not text while driving, or an online safety expert persuading users of social media sites to not reveal too much personal information online. These all involve the persuader finding the right arguments to use with respect to the persuadee's knowledge, priorities, and biases. Using artificial argumentation to build automated persuaders provides several interesting research challenges the community is starting to tackle. For example, the Framework for Computational Persuasion project[18] is developing a computational model of argument for behaviour change in healthcare. This kind of application calls for the development of rhetorical and dialogical layers in Figure 1.

In the longer term, there are exciting possibilities for developing artificial agents able to use argumentation as a general pattern of interaction with other agents, exactly like humans are able to argue with other humans to achieve collectively useful behaviors. Consider a situation where

[13] http://lidia.cs.uns.edu.ar/delp_client/
[14] http://toast.arg-tech.org
[15] http://tweetyproject.org
[16] http://robertcraven.org/proarg
[17] http://argumentationcompetition.org
[18] http://www.computationalpersuasion.com

heterogeneous robots need to work together to survey a situation such as a large building on fire. Exactly like in a team of firefighters, each robot will have direct perception of some local situation and will need to exchange information and coordinate actions with other robots in a dynamic environment where, altogether, information will always be incomplete and inconsistent and, consequently, goals and action plans might need to be revised at any moment. Different capabilities of the team members will have to be taken into account too. These features call for high level arguing capabilities, applicable in a variety of contexts among heterogeneous agents whose unique common property might be the capability to argue itself. In this sense artificial argumentation promises, in the long term, to provide a sort of universal social glue for linking together, in a "plug and play" and cooperative manner, robots and any other kind of intelligent agents.

Acknowledgements

This work has been partially supported by EPSRC grants EP/N008294/1 and EP/N014871/1, by EU H2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 690974 for the project MIREL: MIning and REasoning with Legal texts, and by funds provided by the Institute for Computer Science and Engineering, Universidad Nacional del Sur, Argentina.

References

[APL+14] E. Aharoni, A. Polnarov, T. Lavee, D. Hershcovich, R. Levy, R. Rinott, D. Gutfreund, and N. Slonim. A benchmark dataset for automatic detection of claims and evidence in the context of controversial topics. In *Proceedings of the First Workshop on Argumentation Mining*, pages 64–68. Association for Computational Linguistics, 2014.

[BD07] T.J.M. Bench-Capon and P. Dunne. Argumentation in artificial intelligence. *Artificial Intelligence*, 171(10-15):619–641, 2007.

[BJH+14] Ph Besnard, A. Javier Garca, A. Hunter, S. Modgil, H. Prakken, G. Simari, and F. Toni. Introduction to structured argumentation. argument & computation. *Argument and Computation*, 5(1):1–4, 2014.

[BLSR13] F. Bex, J. Lawrence, M. Snaith, and C. Reed. Implementing the argument web. *Communications of the ACM*, 56(10):951–989, 2013.

[CDG+15] G. Charwat, W. Dvorak, S. Gaggl, J. Wallner, and S. Woltran. Methods for Solving Reasoning Problems in Abstract Argumentation - A Survey. *Artificial Intelligence*, 220:28–63, 2015.

[CMM+06] C. Chesñevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk, and S. Willmott. Towards an argument interchange format. *Knowledge Engineering Review*, 21(4):293–316, 2006.

[Dun95] P. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.

[GK97] T. F. Gordon and N. I. Karacapilidis. The Zeno argumentation framework. In *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL '97)*, pages 10–18. ACM Press, 1997.

[GKSS08] C. Gomes, H. Kautz, A. Sabharwal, and B. Selman. Satisfiability solvers. In F. Van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation*, page 89134. Elsevier, 2008.

[HBC12] J. Horty and T.J.M. Bench-Capon. A factor-based definition of precedential constraint. *Artificial Intelligence and Law*, 20:181–214, 2012.

[HW12] A. Hunter and M. Williams. Aggregating evidence about the positive and negative effects of treatments. *Artificial Intelligence in Medicine*, 56:173–190, 2012.

[KP01] N. I. Karacapilidis and D. Papadias. Computer supported argumentation and collaborative decision making: the HERMES system. *Information Systems*, 26(4):259–277, 2001.

[LT16] M. Lippi and P. Torroni. Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology*, 2016.

[MTB+13] S. Modgil, F. Toni, F. Bex, I. Bratko, C. Chesñevar, W. Dvořák, M. Falappa, X. Fan, S. Gaggl, A. García, M. González, T. Gordon, J. Leite, M. Možina, C. Reed, G. Simari, S. Szeider, P. Torroni, and S. Woltran. The added value of argumentation. In Sascha Ossowski, editor, *Agreement Technologies*, pages 357–403. Springer, 2013.

[PM11] R. Mochales Palau and M-F Moens. Argumentation mining. *Artificial Intelligence in Law*, 19(1):1–22, 2011.

[PM12] N. Pinkwart and B. M. McLaren, editors. *Educational Technologies for Teaching Argumentation Skills*. Bentham Science Publishers, Sharjah, UAE, 2012.

[Pol92] J. Pollock. How to reason defeasibly. *Artificial Intelligence*, 57(1):1 – 42, 1992.

[PS15] H. Prakken and G. Sartor. Law and logic: a review from an argumentation perspective. *Artificial Intelligence*, 227:214–245, 2015.

[RR04] C. Reed and G. Rowe. Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, 13(4):983, 2004.

[RS09] I. Rahwan and G. Simari, editors. *Argumentation in Artificial Intelligence*. Springer, 2009.

[SG14] C. Stab and I. Gurevych. Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing, (EMNLP'14)*, pages 46–56, 2014.

[SGP13] J. Schneider, T. Groza, and A. Passant. A review of argumentation for the social semantic web. *Semantic Web*, 4(2):159–218, 2013.

[TVC+16] M. Thimm, S. Villata, F. Cerutti, N. Oren, H. Strass, and M. Vallati. Summary report of the first international competition on computational models of argumentation. *AI Magazine*, 37(1):102–104, April 2016.

[WGY+97] R. Walton, C. Gierl, P. Yudkin, H. Mistry, M. Vessey, and J. Fox. Evaluation of computer support for prescribing CAPSULE using simulated cases. *British Medical Journal*, 315(7111):791–795, 1997.

[WRM08] D. Walton, C. Reed, and F. Macagno. *Argumentation Schemes*. Cambridge University Press, 2008.