SPP-RATIO (Second Phase) Kick-off Conference 2021

Project Summaries

9th and 10th of June, 2021

Coordinators:

Prof. Dr. Philipp Cimiano
Bielefeld Universität

Prof. Dr. Anette Frank
Universität Heidelberg

Prof. Dr. Benno Stein
Bauhaus-Universität Weimar

Prof. Dr. Michael Kohlhase
FAU Erlangen-Nürnberg

Prof. Dr. Jürgen Ziegler
Universität Duisburg-Essen

Robust Argumentation Machines (RATIO)
Kick-off Conference (2nd Phase)

June 9th-10th, 2021

Dear Attendees,

It is our great pleasure to welcome you to the second phase of the RATIO priority program. This time, the conference will take place in remote mode due to the pandemic situation. We will hold two days conference in which the accepted projects for this phase will be presented, and there will be room for questions in each presentation.

On this occasion, we are honored to have as an invited speaker Prof. Henry Prakken from the University of Groningen, who will talk about *Case-based Argumentation for Explainable AI*.

This booklet comprises the summaries of the projects participating in the second phase of RATIO and the invited talk.

We hope you all enjoy the Kick-off Conference.

Coordination of the RATIO priority program.
Prof. Philipp Cimiano
Dr. Olivia Sanchez-Graillet

Contents

1 Invited Talk

Case-based Argumentation for Explainable AI

Prof. Henry Prakken

Department of Information and Computing Sciences, Utrecht University, and Faculty of Law, University of Groningen, The Netherlands

While machine learning has recently led to spectacular successes in AI, a main issue is the black-box nature of many machine-learning applications. This is especially a problem in applications to decision-making with legal, ethical or social implications. A new subfield of AI has emerged called 'explainable AI'. After a brief overview of this field, I will discuss my recent work with Rosa Ratsma on applying case-based argumentation techniques to explaining the outputs of machine-learning-based decision-making applications. The model draws on AI & law research on argumentation with cases, which models how lawyers draw analogies to past cases and discuss their relevant similarities and differences in terms of relevant factors and dimensions in the problem domain. A case-based approach is natural since the input data of machine-learning applications can be seen as cases. While the approach is motivated by legal decision making, it also applies to other kinds of decision making, such as commercial decisions about loan applications or employee hiring, as long as the outcome is binary and the input conforms to the model's factor- or dimension format.

Reference: H. Prakken & R. Ratsma, A top-level model of case-based argumentation for explanation: formalisation and experiments. Argument and Computation, first online 26 April 2021. DOI: 10.3233/AAC-210009.

2 ACCEPT: Prespectivized Argument Knowledge Graphs for Deliberation Support

Philipp Cimiano (Bielefeld) and Anette Frank (Heidelberg)

Many decisions in daily life and in political or societal contexts are difficult to resolve and require deliberation of arguments in favor or against actions to be taken. In this process, conflicting interests of involved parties need to be balanced to achieve optimal outcomes, general acceptance and fairness. Tools that can support humans in deliberation processes range from debating portals with pre-structured pro and con arguments on widely discussed issues to argument search engines with growing but imperfect capacities for argument structure annotation. Real argumentation, however, goes far beyond coarse-grained structuring of pre-stated pro and con arguments. Deep understanding of debated issues is needed that brings about i) potential implications of decisions; ii) how these affect interested parties, and iii) how to weight potential consequences, in order to find optimal and widely accepted solutions. Importantly, iv) we need to be able to address novel issues not discussed before, and thus, need to build argumentation systems that have access to knowledge resources and offer reasoning capabilities.

Computational argumentation is to date still far from achieving deep understanding of arguments. E.g., tools for classifying support-attack relations perform well in a discourse setting, but our work shows that they crucially rely on discourse markers and perform poorly if these are masked from the input. Similarly, the Argument Reasoning Challenge was designed to enforce deeper reasoning over arguments. But the good performance of BERT-based models turned out to rely on surface cues, and dropped to random behaviour once these cues were eliminated. In prior work we have addressed these weaknesses by integrating background knowledge in an interpretable way, using advanced deep learning methods and uncovering implicit information. But more research is needed to endow systems with deep argument understanding, which includes reasoning about arguments and forseeing consequences of actions, to support deliberation.

In this project we aim to move beyond our achievements in knowledge-based argument analysis: i) We aim to extend our neural-symbolic methods to achieve **deeper understanding of arguments** supported by background knowledge. This includes the induction of **deeper relations between argument components** leading to **enhanced argument explicitation**. ii) We represent, contextualize and enrich arguments in a **multi-factorial argument knowledge graph** that goes beyond current work by including stakeholder perspectives and their interests, values and goals. iii) Based on the argument graph we develop methods that perform **reasoning to derive novel conclusions from many perspectives**, will de- and recompose arguments to perform **perspectivized argument graph completion**, and analyze debated issues from multiple perspectives and learn to reason towards alternative premises, to offer **interpretable deliberation support**.

3 ACQuA-2.0: Answering Comparative Questions with Arguments

Chris Biemann (Hamburg) and Matthias Hagen (Halle)

In the second phase of the ACQuA project in the DFG-SPP RATIO we continue our work on arguments in comparison scenarios. Comparisons are an interesting and

valuable subject of research as people often need to compare different options and need argumentative explanations of pros and cons to come up with an informed own opinion.

Confronted with a question like 'Is X better than Y with respect to Z?', even the big commercial web search engines today often only show facts from a knowledge base that match 'X', 'Y', or 'Z', or they show some related questions mined from question answering websites as kind of a direct answer above the classic ten blue links. One of the main issues—besides the usual dilemma of short direct answers not being able to cover all angles of more complex information needs—is that mining question answering websites and extracting facts from knowledge bases can only help to answer a minority of comparative questions. By analyzing large Yandex query logs in our first project phase, we have shown that most real comparative questions need argumentative explanations that can be found on the web, but not necessarily on question-answering websites. We thus developed the comparative argumentation machine CAM that retrieves sentences from the web related to a comparative information need. We could show that CAM users are faster and more correct than users of a traditional keyword-based search. To identify the most argumentative sentences (and text passages), we developed the neural argument tagger TARGER, and we started to work on distinguishing comparative questions that can be answered by facts from questions that need argumentative answers. Some pilot studies in the first ACQuA phase showed promising preliminary results for extracting facts from a knowledge base to answer factual comparisons. Further, we have co-organized the Touché lab at CLEF 2020: the first shared task on argument retrieval; we will continue to co-organize it in the coming years.

In the second project phase, we plan to (1) strengthen the connection of knowledge bases and textual evidence as sources for argumentative answers on comparative questions, to (2) diversify arguments via axiomatically re-ranking search results from web-scale multilingual and multicultural sources for a broad coverage of different viewpoints (e.g., for "controversial" topics like programming languages or cuisine), and to (3) develop prototypes to embed summarized argumentative answers on natural language comparative questions in applications with interactive conversational interfaces, especially paying attention to argument provenance to ensure a high transparency for the user.

4 BEA: Bulding Engaging Argumentation

Elisabeth André (Ausburg) and Wolfgang Minker (Ulm)

For humans, a natural way of resolving different points of view or forming an opinion is through conversation, i.e., through the exchange of arguments and knowledge. On the other hand, numerous opinions and information on almost every topic are nowadays also directly available online. However, the screening and evaluation of these multiple sources can be a time-consuming and demanding task. Filter algorithms aim to reduce this effort by taking previous user requests into account to serve the users' interest. As a result, people tend to focus on a biased subset of sources that repeat or strengthen an already established or convenient opinion (so-called filter bubbles). In order to avoid this (often unconscious) process of intellectual isolation, we herein propose an approach to explore a vast amount of different information in a natural and intuitive way – through a conversation. To this end, we propose a virtual agent that engages in a deliberative dialogue with a human user in order to support a fair and unbiased opinion building process. In contrast to persuasive systems, the envisioned agent has no topic-specific agenda, but instead aims for providing a diverse and representative overview - embedded

in a conversation with the user.

The proposed system will extend technologies investigated and developed in the scope of the EVA project during Phase I of the SPP RATIO. Whereas in the EVA project users assumed the (passive) role of the audience, the system envisioned for Phase II will allow the user to actively participate in the argumentation process through a natural language interface. Therefore, the system will be able to react to user utterances, such as opinions, requests and questions and provide an individualized conversation over the discussed topic. This will be supported by a transparent presentation of pros and cons towards the discussed topic through the agent's access to recently developed argument search engines and thus contribute to explainability and deep understanding.

5 CAML-2: Causality, Argumentation and Machine Learning

Kristian Kersting (Darmstadt) and Matthias Thimm (Koblenz)

While traditional machine learning methods have led to great advances in artificial intelligence, there are a number of obstacles complicating their wider adoption. Pearl argues that these obstacles arise from the fact that current machine learning methods are associational learning systems which do not attempt to understand cause and effect relationships (Pearl, 2018). Thus, in order to overcome these obstacles, machines must be equipped with the ability to reason about cause and effect. The overall objective of the project CAML2 is to use formal argumentation techniques for causal machine learning. While large amounts of observed data are increasingly easy to obtain, it is well known that causal relationships cannot, in general, be discovered on the basis of observed data alone. This is a process that depends on additional knowledge, such as background knowledge provided by domain experts or results of controlled experiments. A machine that reasons about cause and effect must therefore interact with users. Given a query formulated by the user, it must be able to explain an answer, show which additional knowledge is necessary to arrive at an answer, and be able to integrate knowledge that is provided by the user. Formal argumentation provides tools that facilitate this kind of interaction. In formal argumentation, answers to queries are accompanied by a dialectical analysis showing why arguments in support of a conclusion are preferred to counterarguments. Such an analysis acts as an explanation for an answer, and can furthermore be made interactive, by allowing users to pinpoint errors or missing information and to formulate new arguments that can be integrated into the analysis. More concretely, we will develop an approach that allows for learning causal models from observed data and, in particular, to reason about conflicting causal models using arguments and counterarguments for and against certain models, respectively. The approach allows for the injection of background knowledge and expert information can be elicited through an interactive dialogue system. In order to address problems pertaining to uncertainty and noisy information, we will not only consider qualitative approaches to formal argumentation but also quantitative ones that are able to argue about probabilities. An important notion in causal discovery (and machine learning in general) is that of an intervention, i.e., a "manual" setting of a variable to a certain value. This allows to guide the learning algorithm in the appropriate direction. We will model interventions through the use of counterfactual reasoning methods of formal argumentation, yielding a comprehensive framework for interactive argumentative causal discovery. We will focus on the setting of supervised

learning, but also consider our approach in an unsupervised setting, in particular, for clustering and density estimation.

6 CUEPAQ: Visual Analytics and Linguistics for Capturing, Understanding, and Explaining Personalized Argument Quality

Miriam Butt (Konstanz) and Daniel Keim (Konstanz)

In our project, Visual Analytics and Linguistics for Capturing, Understanding, and Explaining Personalized Argument Quality (CUEPAQ), we combine methods from the fields of visual analytics and computational linguistic to generate new approaches for the analysis of argument quality in terms of various metrics on different levels of linguistic analysis. Based on this analysis, we provide so-called preference profiles, enabling users to gain insights into their personal argumentation behavior and compare it to other users' behavior.

This project's main goal is to capture, understand, and explain the perceived quality of arguments. To that end, we collect various stylistic, content, and semantic features that influence how arguments are framed and perceived. This project's central question is how these elements interact to produce arguments that are perceived as high-quality.

In answering this question, we contribute to the research on argument quality a visual analytics framework for the rating and ranking of arguments. The system enables rapid analysis of interactions between argument quality and the linguistic expression of an argument. Our framework extracts preference profiles, which capture users' annotation behavior by indicating the content and stylistic features that mainly affect their argument rating. These preference profiles may vary from user to user or across different user groups. To account for this, we include both expert knowledge on the annotation of argument quality and results of non-expert user ratings in our analysis of argument quality.

Based on relative preference comparisons between arguments, the system extracts patterns of linguistic features, both stylistic and interpretational. These features are expected to capture the users' preferences and would thus be reflected in their rating behavior. This externalized knowledge is visualized based on specific guidance strategies and allows both the user and the system to learn from each other. Since the system can keep track of the annotation behavior of different users, this co-adaptive process does not only allow a user to understand their own argumentation preferences but also to compare them with other users of the system, as well as the expert opinions on high-quality argumentation.

In this project, we use computational linguistic methods to explore the relationship between linguistic choices and the ranking of arguments by users and systems based on expert-opinion. Concretely, we contribute to the uniform annotation of linguistic features of arguments that are relevant to the judgment of argument quality.

7 FAME: A Framework for Argument Mining and Evaluation

Gerhard Heyer (Leipzig), Ringo Baumann (Leipzig) and Gregor Wiedemann (Hamburg)

Project period: May 2019 – April 2022

Two different perspectives on argumentation have been pursued in computer science research, namely approaches of argument mining from natural language texts on the one hand and formal argument evaluation on the other. So far these research areas are largely independent and unrelated. The overall goal of the FAME project is to link these two perspectives and their respective research agendas. We develop a framework which integrates argument mining and formal argument evaluation. Evaluation results will allow for new types of queries to be answered by argumentation retrieval systems and large-scale content analysis of empirical argument use. Moreover, feeding back evaluation results in the mining process can be utilized to improve the obtained results.

The challenge here is to bridge the gap between processing texts in natural language and evaluating arguments expressed in some formal language with a corresponding formal semantics. The hypothesis underlying this proposal is that controlled natural language (CNL) can provide the necessary link for bridging this gap. CNL can serve as an intermediate representation of argumentative text. They have the look and feel of natural language and are thus close enough to natural language based argument processing. On the other hand, the CNL we will be using is one which also possesses a well-defined formal semantics and is thus amenable to formal evaluation methods. CNLs of the kind mentioned already exist. A prominent example is Attemp to Controlled English (ACE). ACE is easily understood by humans. At the same time, its semantics is formally defined based on a translation into discourse representation structures, which can be further translated into predicate logic. Since ACE was not primarily developed with argumentation in mind, we will thoroughly investigate whether the language constructs provided are sufficient for our purposes, or whether we need to extend the language. Within the project, we investigate argumentation in public, political discourse represented in newspaper text and related user comments with respect to five selected, controversial issues (e.g. free trade agreements). By modeling specific, empiric issues as CNL statements in a knowledge base, we extend existing approaches of argument mining from identification of generic, functional types of argument structures (e.g. premise, claim) towards semantic argument constituents. At the same time, representing such constituents as formally evaluable statements in a knowledge base allows for combination with approaches of abstract argumentation evaluation. We expect this to be a suitable framework to handle specifics argument structures as they generally occur in empiric communication (e.g. incomplete or inconsistent arguments).

8 INAS: Interactive Argumentation Support for the Scientific Domain of Invasion Biology

Tina Heger (München), Birgitta König-Ries (Jena) and Sina Zarrieß (Jena)

Developing a precise argument is not an easy task. In real-world argumentation scenarios, arguments presented in texts (e.g. scientific publications) often constitute the end result of a long and tedious process. A lot of work on computational argumentation has focused on analyzing and aggregating these products of argumentation processes, i.e. argumentative texts. In this project, we adopt a complementary perspective: we aim to develop an argumentation machine that supports users during the argumentation process in a scientific context, enabling them to follow ongoing argumentation in a scientific

community and to develop their own arguments. To achieve this ambitious goal, we will focus on a particular phase of the scientific argumentation process, namely the initial phase of claim or hypothesis development. According to Toulmin (1986), the starting point of an argument is a claim, and also data that serves as a basis for the claim. In scientific argumentation, a carefully developed and thought-through hypothesis (which we see as Toulmin's "claim" in a scientific context) is often crucial for researchers to be able to conduct a successful study and, in the end, present a new, high-quality finding or argument. Thus, an initial hypothesis needs to be specific enough that a researcher can test it based on data, but, at the same time, it should also relate to previous general claims made in the community. We investigate how argumentation machines can (i) represent concrete and more abstract knowledge on hypotheses and their underlying concepts, (ii) model the process of hypothesis refinement, including data as a basis of refinement, and (iii) interactively support a user in developing her own hypothesis based on these resources. This project will combine methods from different disciplines: natural language processing, knowledge representation and semantic web, philosophy of science and – as an example for a scientific domain – invasion biology. Our starting point is an existing resource in invasion biology that organizes and relates core hypotheses in the field and associates them to meta-data for more than 1000 scientific publications, which was developed over the course of several years based on manual analysis. This network, however, is currently static (i.e. needs substantial manual curation to be extended to incorporate new claims) and, moreover, is not easily accessible for users who miss specific background and domain knowledge in invasion biology. Our goal is to develop (i) a semantic model for representing knowledge on concepts and hypotheses, such that also non-expert users can use the network; (ii) a tool that automatically computes links from publication abstracts (and data) to these hypotheses; and (iii) an interactive system that supports users in refining their initial, potentially underdeveloped hypothesis.

9 LARGA: Learning Argumentation Axioms from Monological and Dialogical Texts

Manfred Stede (Potsdam) and Benno Stein (Weimar)

In many real-life situations, from political debates to paper writing, the effectiveness of argumentation depends not only on picking the best arguments, but also on following the best strategy to deliver them. We hence ask: In what order should arguments be presented in a text? What rules or conventions guide these ordering decisions? In what way does a specific linearisation improve or diminish the acceptability of the author's standpoint? Do principles exist that apply across text genres?

An example of such organisation knowledge is "Arguments with units of anecdote type should precede those with units of statistical type.", which is a pattern often found in editorials. With our research, we want to systematise the identification and the analysis of such patterns (or axioms, as they are called here). Most existing work on computational argumentation concentrates on argument mining and argumentation assessment, while the empirical knowledge about what arrangement strategies are effective for which text genre or mode as well as how to identify such kind of strategy knowledge is missing so far.

We want to address this gap by introducing an axiomatic approach for modeling argument arrangement preferences on the basis of "topic-agnostic" attributes. We assign these attributes, which we have been compiling in the course of our recent research, to three

abstraction levels: the argument unit level, the argument level, and the discourse level. Based on adequate datasets, which have to be properly annotated at these three levels, we seek to uncover and analyse interpretable axiomatic knowledge from argumentative texts of different genres, modes, and levels of writing expertise.

Our project provides a concrete plan to acquire the necessary datasets, to induce axioms of the described kind both in monological and dialogical settings, and to analyse interesting relations among these axioms. For instance, how do expert argumentative texts differ from non-expert argumentative texts in terms of these axioms? In addition, we plan to investigate the consequences that our axiom-based approach has on current theories of text structure, and we have devised experiments to demonstrate the benefit of operationalising axiomatic argumentation strategy knowledge in two downstream applications: augmented writing and dialog assistance. The resources developed as part of this project, including the annotations, code, axiomatic knowledge, and prototypical tools, will be made freely available, contributing to the RATIO priority programme and research on argumentation in general.

10 MARDY-2: Modeling Argumentation Dynamics in Political Discourse

Sebastian Haunss (Bremen), Jonas Kuhn and Sebastian Padó (Stuttgart)

In the first phase of SPP 1999, the project MARDY (Modeling Argumentation Dynamics in Political Discourse) brought together machine learning methods from Natural Language Processing and theoretically grounded analysis frameworks from Political Science in order to better understand how actors in political debates (politicians, parties, demonstrators) articulate their positions by making certain claims (e.g., "migrants should be integrated into the labor market"), how they form 'discourse coalitions' to achieve shared goals, and how debates evolve over time. To make this vision feasible, MARDY1 limited its focus to the study of a single topically confined debate at a time, namely the German national debates on migration and on pensions, using data from one newspaper source, and restricting network modeling to actors and claims. The interdisciplinary collaboration turned out to be highly successful, showing that the analysis of long-running political debates can indeed be sped up considerably while maintaining quality, and yielding new types of insights. Yet, the necessary limitations restrict the general applicability of the project's results, and form the starting point for the second phase of MARDY.

The goal of MARDY2 is to generalize in two crucial dimensions. First, to substantially generalize both the methodology and statistical models of MARDY to enable not only the analysis of new debates on arbitrary new topics, but also cross-national (and therefore, cross-lingual) comparisons. This capability will be leveraged to analyse the ongoing debate on COVID-19 and its policy implications in Germany, France, the UK, and the US. Second, to systematically include frames into the models. Frames are argumentative patterns used to substantiate a claim and add an extra level of structure. For example, the claim that migrants should be integrated into the labor market can justified by left-wing actors by referring to concerns for general societal integration, or by center/right wing actors by referring to economic gains. Modeling the dynamics of actors, claims, and frames concurrently creates additional leverage for machine learning based detection and classification of actors' statements. To take full advantage of the benefits of frames, we extend our research approach to include longer argumentative texts.

Not only can the broadened analytical scope lead to enriched corpus analyses in (computational) social science, it also helps to push the methodological development to a more systematic level, working towards generic workflows and tools supporting the analysis of argumentation in any given complex discourse setting.

11 OASiS: Objective Argument Summarization in Search

Martin Potthast (Leipzig) and Henning Wachsmuth (Paderborn)

Conceptually, an argument logically combines a claim with a set of reasons. In real-world text, however, arguments may be spread over several sentences, often intertwine multiple claims and reasons along with context information and rhetorical devices, and are inherently subjective. This project aims to study how to computationally obtain an objective summary of the gist of an argumentative text. In particular, we aim to establish foundations of natural language processing methods that (1) analyze the gist of an argument's reasoning, (2) generate a text snippet that summarizes the gist concisely, and (3) neutralize potential subjective bias in the summary as far as possible.

The rationale of the planned project is that argumentation machines, as envisioned by the RATIO priority program, are meant to present the different positions people may have towards controversial issues, such as abortion or social distancing. One prototypical machine is our argument search engine, *args.me*, which opposes pro and con arguments from the web in response to user queries, in order to support self-determined opinion formation. A key aspect of *args.me* and comparable machines is to generate argument snippets, which give the user an efficient overview of the usually manifold arguments. Standard snippet generation has turned out to be insufficient for this purpose. We hypothesize that the best argument snippet summarizes the argument's gist objectively.

Building upon existing argument mining techniques, the project centers around text summarization and style transfer. Both tasks are hard in general, but extensive research in the last years has shown potential in focused domains. For arguments, short content summaries have been generated successfully, yet without attention to the argument's reasoning. Style transfer even remains fully unstudied so far in computational argumentation. We aim to fill these gaps through the project's objectives.

In particular, we plan to (1) create a first corpus of thousands of human-written argument summaries of different kinds. On this basis, we develop and evaluate genuine computational methods that (2) generate a summary of an argument's reasoning and (3) neutralize the style of the summaries to reduce subjective bias. In (4) empirical studies, we explore what kinds of summaries are seen as best and why. We expect to obtain new knowledge about what is important in summarizing subjective language such as argumentation. The corpus enables more systematic research on argument generation, and the methods will be useful for various types of argumentation machines and other applications of computational argumentation. Making all outcomes publicly available, the project substantially contributes to the goals of RATIO in particular, and to natural language processing research in general.

12 Open Argument Mining

Iryna Gurevych (Darmstadt) and Steffen Staab (Stuttgart)

Project period: May 2019 – April 2022

Open debates include so many arguments that sound decision making exceeds cognitive capabilities of the interested public or responsible experts. Until now, argument mining approaches typically map from a closed set of given texts into a formal argumentation model. However, this does not fully cater to the nature of open, ongoing debates because of the following challenges:

(C1) Following a Continuous Debate: Participants in open, mass debates continuously introduce new arguments with novel aspects relevant to the debate topic. This leads to brittleness of state-of-the-art argument extractors, which are trained once and for all time.

(C2) Dealing with Incomplete Arguments: Textual arguments are often incomplete because the participants in the debate can understand them based on common background or shared knowledge. Hence argument structures which are identified by current argument mining methods are often incomplete.

(C3) Establishing Open Knowledge for Argumentation: Interpreting and understanding textual arguments requires additional facts and common knowledge which are often absent in existing knowledge graphs.

Our project "Open Argument Mining" investigates computational methods that i) continuously improve their capability to recognize arguments in ongoing debates, ii) align incomplete arguments with previous arguments and enrich them with automatically acquired background knowledge, and iii) continuously extend semantic knowledge bases with information required to understand arguments. We intend to achieve this by combining the two research fields argument mining and knowledge graph construction.

13 RAND: Reconstructing Arguments from Newsworthy Debates

Stefan Evert (Erlangen) and Lutz Schröder (Erlangen)

Large portions of ongoing political debates are available in machine-readable form nowadays, ranging from the formal public sphere of parliamentary proceedings to the semi-public sphere of social media. This offers new opportunities for gaining a comprehensive overview of the arguments exchanged, using automated techniques to analyse text sources. The goal of the RANT/RAND project series within the priority programme RATIO (Robust Argumentation Machines) is to contribute to the automated extraction of arguments and argument structures from machine-readable texts via an approach that combines logical and corpus-linguistic methods and favours precision over recall, on the assumption that the sheer volume of available data will allow us to pinpoint prevalent arguments even under moderate recall. Specifically, we identify logical patterns corresponding to individual argument schemes taken from standard classifications, such as argument from expert opinion; essentially, these logical patterns are formulae with placeholders in dedicated modal logics. To each logical pattern we associate several linguistic patterns corresponding to different realisations of the formula in natural language; these patterns are developed and refined through corpus-linguistic studies and formalised in terms of corpus queries. Our approach thus integrates the development of automated argument extraction methods with work towards a better understanding of the linguistic aspects of everyday political argumentation.

Research in the ongoing first project phase is focused on designing and evaluating patterns and queries for individual arguments, with a large corpus of English Twitter messages used as a running case study. In the second project phase, we plan to test

the robustness of our approach by branching out into additional text types, in particular longer coherent texts such as newspaper articles and parliamentary debates, as well as by moving to German texts, which present additional challenges for the design of linguistic patterns (i.a. due to long-distance dependencies and limited availability of high-quality NLP tools).

Crucially, we will also introduce similarity-based methods to enable complex reasoning on extracted arguments, representing the fillers in extracted formulae by specially tailored neural phrase embeddings. Moreover, we will extend the overall approach to allow for the high-precision extraction of argument structure, including explicit and implicit references to other arguments. We will combine these efforts with more specific investigations into the logical structure of arguments on how to achieve certain goals and into the interconnection between argumentation and interpersonal relationships, e.g. in ad-hominem arguments.

14 ReCAP-II: Information Retrieval and Case-Based Reasoning for Robust Deliberation and Synthesis of Arguments – Architecture and Applications

Ralph Bergmann (Trier) and Ralf Schenkel (Trier)

The ReCAP–II proposal continues the research performed in the ReCAP project, currently funded within the RATIO DFG Priority Programme. ReCAP–II aims at significantly contributing to the foundations for building argumentation machines according to the overall goal of the Priority Programme. The project focus remains on argumentation to support researchers, journalistic writers, as well as human decision makers to obtain a comprehensive overview of current arguments and opinions related to a certain topic. Such argumentation machines automatically explore and process available information sources on the Web, particularly argumentative texts and factual content relevant for the specific topic under discussion. Unlike existing search engines, which primarily operate on the textual level, such argumentation machines reason on the knowledge level formed by arguments and argument structures. For a given particular context, such reasoning will actively support the deliberation of arguments and counter-arguments for the issue under consideration. In addition, it will support the synthesis of new arguments, based on analogical transfer from similar related contexts and topics. For this purpose, the project aims at novel contributions to and confluence of methods from information retrieval (IR) and knowledge representation and reasoning, in particular case-based reasoning (CBR), for building argumentation machines. The aim is to develop methods that are able to capture arguments in a robust and scalable manner, in particular representing, contextualising, and aggregating arguments and making them available to a user.

In particular, RECAP-II focusses on two major goals. The first goal is to further improve the methods already developed in RECAP with respect to specific challenges. In particular, we will address interactive, explainable CBR for argument synthesis as well as argument validation and evaluation. The second goal is to achieve, the integration of the developed methods in order to approach the overall goal of RECAP and the RATIO Priority Program, namely to build an argumentation machine. For this purpose, we will additionally investigate state-of-the-art methods from argument mining for extracting argument graphs from German texts as well as methods for user interaction, visualization, and context representation. As a major result, we aim at developing a technical architecture of argumentation machines with clearly specified services and interfaces. Based on

this, several use cases will be implemented as demonstrator application, which will then allow to perform an end-to-end evaluation using real users from journalism and political research.

15 The Bayesian Approach to Robust Argumentation Machines

Stephan Hartmann (München)

It is well-known that the Bayesian approach to argumentation (i) has a solid normative foundation and (ii) connects well with empirical data from experiments in the psychology of reasoning and argumentation. The main objective of this research proposal is to demonstrate that it also has the computational resources to allow for large-scale applications in the context of robust argumentation machines. We will adapt some of the available computational tools and methods to the study of argumentation and develop new tools and methods if needed. More specifically, our project has the following four objectives: (1) To use machine learning tools to learn Bayesian networks from large data sets. (2) To develop adequate argument generation and evaluation algorithms from these Bayesian networks. (3) To set up tools for testing perceived argument quality of generated arguments. (4) To use these tools to test the arguments we generated.