On the End-to-End Argument Validation System based on Communicative Discourse Trees

Boris Galitsky[1] and Dmitry Ilvovsky[2]

[1] Oracle Corp. Redwood Shores, CA, USA
boris.galitsky@oracle.com
[2] Higher School of Economics Moscow Russia
dilvovsky@hse.ru

Abstract. We formulate a problem of an assessment of argumentation validity based on rhetorical analysis of text. Argumentation structure can be detected in text in the form of discourse trees extended with edge labels for communicative actions. Extracted argumentation structure is represented as a defeasible logic program and is subject to dialectical analysis to establish the validity of the arguments for the main claim being communicated. We evaluate the accuracy of argument mining and then argument validation as well as an overall performance of an end-to-end argumentation system.

1 Introduction

In this study we focus on validating claims of human agent expressed in text. In non-trivial cases, claim validation relies on an analysis of arguments. When domain knowledge is available and formalized, truthfulness of a claim can be validated directly. However, in most text analysis environments such knowledge is unavailable and other implicit means need to come into play, such as writing style and writing logic, in particular, used argumentation patterns. In this study we employ the discourse analysis in our end-to-end argument validation system for texts and explore which discourse features can be leveraged for argumentation validity analysis.

When an author attempts to provide an argument for something, a number of argumentation patterns can be employed. The basic points of argumentation are reflected in the rhetorical structure of text where an argument is present (Moens et al., 2007). We select the Rhetoric Structure Theory (RST, in Mann and Thompson 1988) as a means to represent discourse features associated with logical argumentation. Nowadays, the performance of both rhetoric parsers and argumentation reasoners has dramatically improved (Feng and Hirst 2014). Taking into account the discourse structure of conflicting dialogs, one can judge on the authenticity and validity of these dialogs in terms of its argumentation. In this work we will evaluate the *combined* argument validity assessment system that includes both the *discourse structure extraction* and *reasoning about it* with the purpose of the validation of an agent's claim. Either approach to argument detection from text or to reasoning about formalized arguments has been undertaken (Galitsky and Pampapathi 2003, Symeonidis et al., 2007), but not the whole argument assessment system.

Most of the modern techniques treat computational argumentation as specific discourse structures and perform detection of arguments of various sorts in text, such as classifying a text paragraph as argumentative or non-argumentative (Moens et al., 2007). A number of systems recognize components and structures of logical arguments (Sardianos et al., 2015). However, these systems do not rely on discourse trees (DTs); they only extract arguments and do not apply logical means to evaluate it. At the same time, a broad corpus of research deals with logical arguments irrespectively of how they may occur in natural language (Bondarenko et al., 1997). A number of studies addressed argument quality in logic and argumentation theory (van Eemeren et al., 1996; Damer, 2009), however the number of systems that assess the validity of arguments in text is very limited (Cabrio and Villata, 2012). Most argument mining systems are either classifiers which recognize certain forms of logical arguments in text, or reasoners over the logical representation of arguments (Amgoud et al., 2015).

To address this shortcoming, in this project, we build an end-to-end argumentation system, augmenting an argument extraction from text with its logical analysis. To represent the linguistic features of text, we use the following sources:

1) *Rhetoric relations* between the parts of the sentences, obtained as a *discourse tree* (DT). Discourse trees encode rhetorical relations such as *Cause, Contrast, Condition, Attribution* which are correlated with argumentation *attack* relation.

2) *Speech acts and communicative actions*, obtained as verbs from the VerbNet resource.

To assess the logical validity of an extracted argument, we apply the Defeasible Logic Program (DeLP; in Garcia and Simari 2004), part of which is built on the fly from facts and clauses extracted from these sources. We integrate argumentation detection and validation components into a decision support system that can be deployed, for example, in the customer relationship management (CRM) domain. To evaluate our approach to extraction and reasoning about argumentation, we chose the dispute resolution / customer complaint validation task because an argumenation analysis plays an essential role in it.

2 Rhetorical Representation of Argumentation

We start with a political domain and give an example of conflicting agents providing their interpretation of certain events. These agents provide argumentation for their claims; we will observe how formed rhetoric structures correlate with their argumentation patterns. We focus on the Malaysia Airlines Flight 17 example with the agents exchanging arguments: Dutch investigators, The Investigative Committee of the Russian Federation, and the self-proclaimed Donetsk People's Republic. It is a controversial conflict where each agent attempts to blame its opponent. To sound more convincing, each agent postulates its claim in a way to attack the claims of its opponents, matching their argumentation styles and trying to defeat their claims.

"*Dutch accident investigators say that strong evidence points to pro-Russian rebels as being fully responsible for shooting down plane. The report indicates where the missile was fired from and identifies who was in control of the territory and pins the downing of MH17 on the pro-Russian rebels."* (Fig. 1a).

"*The Investigative Committee of the Russian Federation believes that the plane was hit by a missile, which could not be produced in Russia. The committee cites an investigation that established the type of the missile and disagrees with Dutch accident investigators."(*Fig. 1b)

"*Rebels, the self-proclaimed Donetsk People's Republic, deny that they controlled the territory from which the missile was allegedly fired. They confirm that it became possible only after three months after the tragedy to say if rebels controlled one or another town and the claim of Dutch accident investigators is flawed*"(Fig. 1c).

To show the structure of arguments one needs to merge discourse relations with information from speech acts. We need to know the discourse structure of interactions between agents, and what kinds of interactions they are. For argument identification, we do not need to know the domain of interaction (here, aviation), the subjects of these interaction, what are the entities, but we need to take into account mental, domain-independent relations between them. We accomplish this by introducing the concept of Communicative Discourse Tree (CDT).

CDT is a DT with labels for edges that are the VerbNet expressions for verbs (which are communicative actions, (CA, Galitsky and Kuznetsov 2008)). Arguments of verbs are substituted from text according to VerbNet frames (Kipper et al., 2008). The first and possibly second argument is instantiated by agents. The consecutive arguments are instantiated by noun or verb phrases which are the subjects of CA. For example, the nucleus node for elaboration relation (on the left of Fig. 1a) is labeled with *say(Dutch, evidence)*, and the satellite is labeled with *responsible(rebels, shooting_down)*. These labels are not intended to express that the subjects of Elementary Discourse Units (EDUs) are evidence and *shooting_down* but instead are intended for matching this CDT with others for the purpose of finding similarity between them.

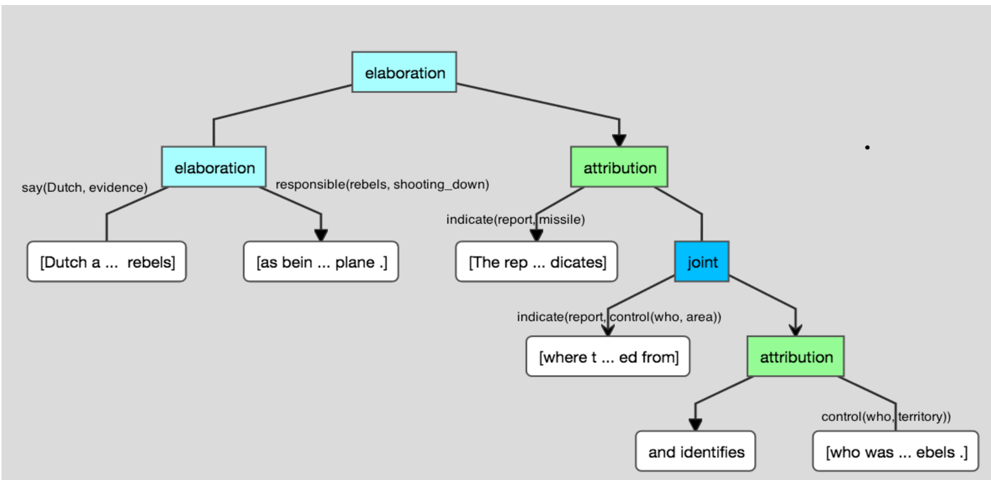

Fig. 1a The claim of the first agent, Dutch accident investigators

Notice that in the CDTs for three paragraphs expressing the views of conflicting parties (Figs 1a, 2b and 2c), communicative actions with their subjects contain the main claims of the respective party, and the DTs without these labels contain information on how these claims are logically packaged. To summarize, a typical CDT for a text with argumentation includes rhetoric relations other than Elaboration and Join, and a substantial number of communicative actions. However, these rules are complex enough so that the structure of CDT matters and tree-specific learning is required (Galitsky et al., 2015).

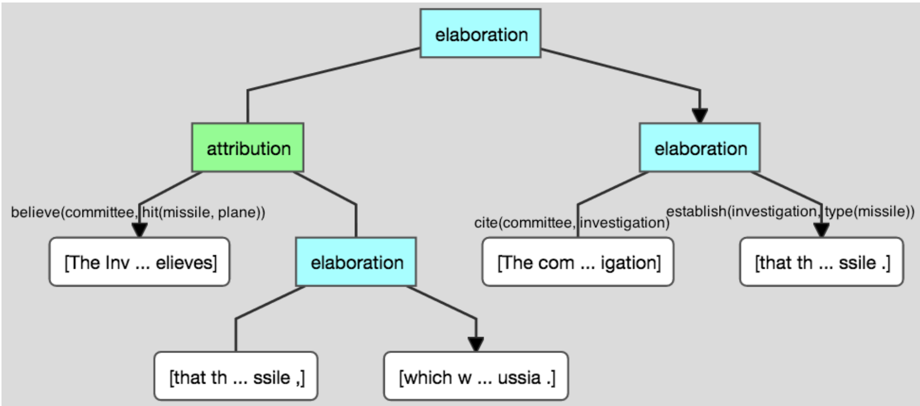

Fig. 1b The claim of the second agent, the Committee

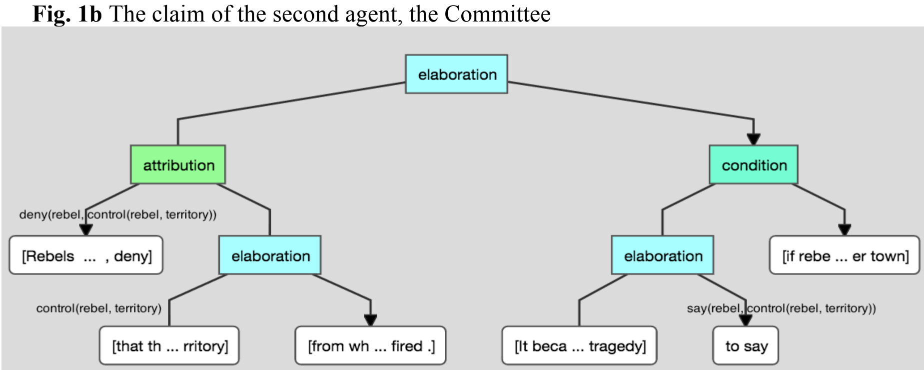

Fig. 1c The claim of the third agent, the rebels

3 Detecting Argumentation in Communicative Discourse Trees

Argumentation analysis needs a systematic approach to learn associated discourse structures. The features of CDTs could be represented in a numerical space so that argumentation detection can be conducted; however, structural information on DTs would not be leveraged. Also, features of argumentation can potentially be measured in terms of maximal common sub-DTs, but such nearest neighbor learning is computationally intensive and too sensitive to errors in DT construction. Therefore, a CDT-kernel learning approach is selected which applies a support vector machine (SVM) learning to the feature space of all sub-CDTs of the CDT for a given text where an argument is being detected.

Tree Kernel (TK) learning for strings, parse trees and parse thickets is a well-established research area nowadays. The CD-TK counts the number of common sub-trees as the discourse similarity measure between two DTs. In this study, we extend the TK definition for the CDT, augmenting DT kernel by the information on CAs. TK-based approaches are not very sensitive to errors in parsing (syntactic and rhetoric) because erroneous sub-trees are mostly random and will unlikely be common among different elements of a training set.

A CDT can be represented by a vector V of integer counts of each sub-tree type (without taking into account its ancestors):

$V(\boldsymbol{T}) = (\#\ of\ subtrees\ of\ type\ 1, \ldots, \#\ of\ subtrees\ of\ type\ \boldsymbol{I}, \ldots, \#\ of\ subtrees\ of\ type\ n)$. Given two tree segments CDT_1 and CDT_2 , the tree kernel function is defined: $K(CDT_1, CDT_2) = <V(CDT_1), V(CDT_2)> = \Sigma i\ V(CDT_1)[i], V(CDT_1)[i] =$

$\Sigma n_1 \Sigma n_2\ \Sigma_i\ I_i(n_1) * I_i(n_2)$, where $n_1 \in N_1$, $n_2 \in N_2$ and N_1 and N_2 are the sets of all nodes in CDT_1 and CDT_2 , respectively; $\boldsymbol{I}_i(n)$ is the indicator function:

$\boldsymbol{I}_i(n) = \{1$ iff a subtree of type i occurs with a root at a node; 0 otherwise$\}$. Further details for using TK for paragraph-level and discourse analysis are available in (Galitsky 2017).

Only the arcs of the same type of rhetoric relations (presentation relation, such as antithesis, subject matter relation, such as condition, and multinuclear relation, such as List) can be matched when computing common sub-trees. We use N for a nucleus or situations presented by this nucleus, and S for a satellite or situations presented by this satellite. Situations are propositions, completed actions or actions in progress, and communicative actions and states (including beliefs, desires, approve, explain, reconcile and others). Hence we have the following expression for RST-based generalization '^' for two texts *text1* and *text2* :

text1 ^ text2 = ∪i,j (*rstRelation1i*, (…,…) ^ *rstRelation2j* (…,…)), where I ∈ (RST relations in *text1*), j ∈ (RST relations in *text2*). Further, for a pair of RST relations their generalization looks as follows: *rstRelation1(N1, S1) ^ rstRelation2 (N2, S2) = (rstRelation1^ rstRelation2)(N1^N2, S1^S2).*

We define CA as a function of the form verb (agent, subject, cause), where verb characterizes some type of interaction between involved agents (e.g., explain, confirm, remind, disagree, deny, etc.), subject refers to the information transmitted or object described, and cause refers to the motivation or explanation for the subject. To handle meaning of words expressing the subjects of CAs, we apply word2vec models (Mikolov et al., 2015).

We combined Stanford NLP parsing, coreferences, entity extraction, DT construction (discourse parser, Surdeanu et al., 2016 and Joty et al., 2013), VerbNet and Tree Kernel builder into one system available at

https://github.com/bgalitsky/relevance-based-on-parse-trees.

4 Claim Validation via Dialectical Analysis

To convince an addressee, a message needs to include an argument and its structure needs to be valid. Once an argumentation structure extracted from text is represented via CDT, we need to verify that the main point (target claim) communicated by the author is not logically attacked by her other claims. To assess the validity of the argumentation, a Defeasible Logic Programming (DeLP) approach is selected. It is an argumentative framework based on logic programming (García and Simari, 2004; Alsinet et al., 2008).

A DeLP is a set of facts, strict rules Π of the form (A:-B), and a set of defeasible rules Δ of the form A-<B, whose intended meaning is "if B is the case, then usually A is also the case". Let P=(Π, Δ) be a DeLP program and L a ground literal. Let us now build an example of a DeLP for legal reasoning about facts extracted from text (Fig. 2). A judge hears an eviction case and wants to make a judgment on whether rent was provably paid (deposited) or not (denoted as *rent_receipt)*. An input is a text where a defendant is expressing his point. Underlined words form the clause in DeLP, and the other expressions formed the facts.

The complaint is as follows: *The landlord contacted me, the tenant, and the rent was requested. However, I refused the rent since I demanded repair to be done. I reminded the landlord about necessary repairs, but the landlord issued the three-day notice confirming that the rent was overdue. Regretfully, the property still stayed unrepaired.*

Defeasible Rules Prepared In Advance
rent_receipt -< rent_deposit_transaction.
rent_deposit_transaction -< contact_tenant.
¬ *rent_deposit_transaction -<contact_tenant, three_days_notice_is_issued.*
¬ *rent_deposit_transaction -< rent_is_overdue.*
¬ *repair_is_done -< rent_refused, repair_is_done.*
repair_is_done -< rent_is_requested.
¬ *rent_deposit_transaction -< tenant_short_on_money, repair_is_done.*
¬ *repair_is_done -< repair_is_requested.*
¬ *repair_is_done -<rent_is_requested.*
¬ *repair_is_requested -< stay_unrepaired.* ¬ *repair_is_done -< stay_unrepaired.*
Target Claim to be Assessed
? - rent_receipt
Clauses Extracted from text
repair_is_done -< rent_refused.
Facts from text
contact_tenant. rent_is_requested. rent_refused. remind_about_repair. three_days_notice_is_issued.
rent_ is_overdue. stay_unrepaired.

Fig. 2 An example of a Defeasible Logic Program for modeling category mapping

We outline the algorithm for validation of a domain-specific claim for arguments extracted from text:

1. Build a DT from input text;
2. Attach communicative actions to its edges to form CDT;
3. Detect argumentation from this CDT using SVM learning; Stop if not detected.
4. Extract subjects of communicative actions attached to CDT and add to 'Facts' section (Fig. 3 on the left);
5. Extract the arguments for rhetoric relation *contrast* and communicative actions of the class *disagree* and add to 'Clauses Extracted FromText' section of Fig. 2;
6. Add a domain-specific section to DeLP;
7. Having the DeLP formed, build a dialectical tree and assess the claim (Fig. 3 on the right).

We use the Tweety (2017) system for DeLP implementation (Thimm 2014).

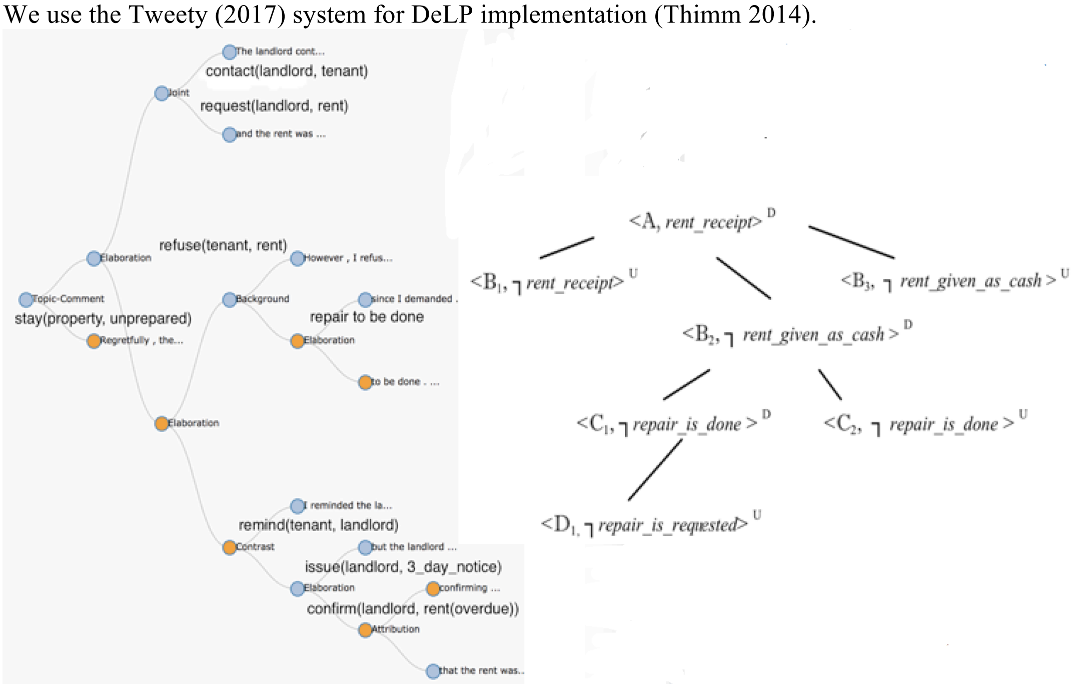

Fig. 3 The CDT for the complaint (on the left, (Joty et al. 2013) visualization) and the dialectical tree for target claim *rent_receipt* (on the right)

5 Evaluation and Conclusions

The objective of argument detection task is to identify all kinds of arguments, not only the ones associated with customer complaints. We formed the positive dataset from textual customer complaints dataset (Galitsky et al., 2009, Github 2018) scraped from consumer advocacy site PlanetFeedback.com. The domain of residential real estate complaints was selected and a DeLP ontology was built for this domain. Automated complaint processing system can be essential, for example, for property management companies in their decision support procedures (Constantinos et al., 2003).

This dataset is used for both argument detection (first step) and argument validity (second step) tasks. For argument detection, we attempt to identify if a given paragraph of text has contains an argument, in a domain-independent manner. For argument validation, in the second step, if we detected an argument in the first step, we try to validate it having the domain-ontology built in a given vertical domain such as landlord-tenant dispute. If an argument has not been detected in the first step, we have nothing to validate.

Table 1 Evaluation results for argument detection

Method / sources	P	R	F1
Bag-of-words	57.2	53.1	55.07
WEKA-Naïve Bayes	59.4	55.0	57.12
SVM TK for RST and CA (full parse trees)	77.2	74.4	75.77
SVM TK for DT	63.6	62.8	63.20
SVM TK for CDT	82.4	77.0	79.61

For the *negative* dataset, only for the argument detection task, we used *Wikipedia*, factual news sources, and also the component of (Lee, 2001) dataset that includes such sections of the corpus as: instructions for how to use software, hardware, presentations of a news article in an objective, independent manner, and others. Further details on the data set are available in (Galitsky et al 2015).

Each row indicates a method used to detect a presence of argumentation in a paragraph. We start with baseline methods, based on keywords and their frequencies (second and third row on the top, Table 1). Second column shows precision (P), third – recall and the fourth – F1 measure. Frequently, a coordinated pair of communicative actions (so that at least one has a negative sentiment polarity related to an opponent) is a hint that logical argumentation is present. This naïve approach is outperformed by the top performing TK learning CDT approach by 29%. SVM TK of CDT outperforms SVM TK for RST+CA and RST + full parse trees (Galitsky et al., 2018) by about 5% due to noisy syntactic data which is frequently redundant for argumentation detection.

In our validity assessment, we focus on target features (claims) related to how a given complaint needs to be handled, such as *compensation_required, proceed_with_eviction, rent_receipt* and others. System decision is determined by whether claim is validated or not: if it is validated, then the decision support system demands compensation, and if not validated, decides that compensation should not be demanded (for the *compensation_required* claim).

Table 2 Evaluation results for argument validation

Types of complaints	P	R	F1 of validation	F1 of total
Single rhetoric relation of type *contrast*	87.3	15.6	26.5	18.7
Single communicative action of type *disagree*	85.2	18.4	30.3	24.8
Two or three specific relations or communicative actions	80.2	20.6	32.8	25.4
Four and above specific relations or communicative actions	86.3	16.5	27.7	21.7

Validity assessment results are shown in Table 2. These results are computed together for detection and validation steps. In the first and second rows, we show the results of the simplest complaint with a single rhetoric relation such as *contrast* with a single CA indicating an extracted argumentation attack relation respectively. In the third row we assess complaints of average complexity, and in the bottom row, the most complex, longer complaints in terms of their CDTs. The third column shows detection accuracy for invalid argumentation in complaints in a stand-alone argument validation system. Finally, the fourth column shows the accuracy of the integrated argumentation extraction and validation system.

In these results recall is low because in the majority of cases the invalidity of claims is due to factors other than being self-defeated. Precision is relatively high since if a logical flaw in an argument is established, most likely the whole claim is invalid because other factors besides argumentation (such as false facts) contribute as well. As complexity of a complaint and its discourse tree grows, F1 first improves since more logical terms are available and then goes back down as there is a higher chance of a reasoning error due to a noisier input.

For decision support systems, it is important to maintain a low false positive rate. It is acceptable to miss invalid complaints, but for a detected invalid complaint, confidence should be rather high. If a human agent is recommended to look at a given complaint as invalid, her expectations should be met most of the time. Although F1-measure of the overall argument detection and validation system is low in comparison with modern recognition systems, it is still believed to be usable as a component of a CRM decision-support system.

We observed that by relying on discourse tree data, one can reliably detect patterns of logical argumentation. Communicative discourse trees become a source of information to form a defeasible logic program to validate an argumentation structure. Although the performance of the former being about 80% is significantly above that of the latter (29%), the overall pipeline can be useful for detecting cases of invalid argumentation, which is important in decision support for CRM.

To the best of our knowledge, this is the first study building a whole argument validity pipeline in the industrial setting. Hence although the overall detection rate for invalid argument is fairly low, there is no existing system to compare this performance against. All detected cases with invalid claims are very valuable for a business or a legal case. In this paper we attempted to combine the best of both worlds, argumentation mining from text and reasoning about the extracted argument. Whereas applications of either technology are limited, the whole argumentation validation system is expected to find a broad range of applications. In this work, we focused on a very specific legal area such as customer complaints, but it is easy to see a decision support system employing the proposed argumentation pipeline in other domains of CRM.

Acknowledgements

The article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

References

1. Alsinet, T., Carlos Iván Chesñevar, Lluis Godo, Guillermo Ricardo Simari. 2008. A logic programming framework for possibilistic argumentation: Formalization and logical properties. Fuzzy Sets and Systems 159(10): 1208-1228.
2. Amgoud, L., Besnard, P. and Hunter, A. 2015. Representing and Reasoning About Arguments Mined from Texts and Dialogues. ECSQARU, pp 60-71.
3. Bondarenko, A., Dung, P., Kowalski, R., Toni, F. 1997. An abstract, argumentation-theoretic approach to default reasoning. Artificial Intelligence 93, pp 63–101.
4. Cabrio, E. and Villata, S. 2012. Combining textual entailment and argumentation theory for supporting online debates interactions. In ACL.
5. Constantinos JS, Sarmaniotis, C., Stafyla, A. CRM and customer-centric knowledge management: an empirical research. 2003. Business Process Management Journal, Vol. 9, Issue: 5, 617-634.
6. Damer, T.E. 2009. Attacking Faulty Reasoning: A Practical Guide to Fallacy-Free Reasoning. Wadsworth Cengage Learning.
7. Feng, WV and Hirst, G. 2014. A linear- time bottom-up discourse parser with constraints and post-editing. In ACL.
8. Galitsky, B. and Pampapathi, R. 2003. Deductive and inductive reasoning for processing the claims of unsatisfied customers. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems. 21-30.
9. Galitsky, B. and Kuznetsov S.O. 2008. Learning communicative actions of conflicting human agents. Journal of Experimental & Theoretical Artificial Intelligence 20 (4), 277-317
10. Galitsky, B., González, MP, CI Chesñevar. 2009. A novel approach for classifying customer complaints through graphs similarities in argumentative dialogue. Decision Support Systems, 46-3, 717-729.
11. Galitsky, B., de la Rosa, J.L. 2011. Concept-based learning of human behavior for customer relationship management. Information Sciences 181 (10), 2016-2035.
12. Galitsky, B., Detecting Rumor and Disinformation by Web Mining. 2015. AAAI Spring Symposium.

13. Galitsky, B, Ilvovsky, D. and Kuznetsov SO. 2015. Rhetoric Map of an Answer to Compound Queries. ACL-2, 681–686.
14. Galitsky, B. 2017. Using Extended Tree Kernel to Recognize Metalanguage in Text. In Uncertainty Modeling, Volume 683 of the series Studies in Computational Intelligence, pp.71-96, Springer.
15. Galitsky, B., Ilvovsky, D., Kuznetsov, SO. 2018. Detecting logical argumentation in text via communicative discourse tree. Journal of Experimental & Theoretical Artificial Intelligence, 1-27.
16. Galitsky, B. and Taylor, J. 2018. Discovering and accessing heated arguments at the discourse level. Dialogue 2018, Moscow, Russia.
17. Garcia, A. and Simari GR. 2004. Defeasible Logic Programming: An Argumentative Approach. Theory and Practice of Logic Programming 4(1-2):95--138.
18. Debanjan, G and Smaranda Muresan, Nina Wacholder, Mark Aakhus, and Matthew Mitsui. 2014. Analyzing argumentative discourse units in online interactions. In Proceedings of the First Workshop on Argumentation Mining, pages 39–48, Baltimore, Maryland, June. ACL.
19. Github. 2018. Intense Argumentation dataset. https://github.com/bgalitsky/relevance-based-on-parse-trees/blob/ master/src/test/resources/opinionsFinanceTagged.xls.zip.
20. Joty, Shafiq R, Giuseppe Carenini, Raymond T Ng, and Yashar Mehdad. 2013. Combining intra-and multi- sentential rhetorical parsing for document-level dis- course analysis. In ACL (1), pages 486–496.
21. Kipper, K. Korhonen, A., Ryant, N. and Palmer, M. 2008. A large-scale classification of English verbs. Language Resources and Evaluation Journal, 42, pp. 21-40.
22. Lawrence, J. and Reed, C. 2015. Combining Argument Mining Techniques. ArgMining@ HLT-NAACL, 127-136.
23. Mann, William and Sandra Thompson. 1988. Rhetorical structure theory: Towards a functional theory of text organization. Text-Interdisciplinary Journal for the Study of Discourse, 8(3):243–281.
24. Mikolov, Tomas, Chen, Kai, Corrado; G.S., Dean; Jeffrey (2015). Computing numeric representations of words in a high-dimensional space. US Patent 9,037,464, Google, Inc.
25. Moens, Marie-Francine, Erik Boiy, Raquel Mochales Palau, and Chris Reed. 2007. Automatic detection of arguments in legal texts. In Proceedings of the 11th International Conference on Artificial Intelligence and Law, ICAIL ’07, pages 225–230, Stanford, CA.
26. Pendyala, V.S., Figueira, S. 2015. Towards a truthful world wide web from a humanitarian perspective. Global Humanitarian Technology Conference,2015 8-11.
27. Sardianos, C. Katakis, IM, Petasis, G. and Karkaletsis, V. 2015. Argument extraction from news. In Proceedings of the 2nd Workshop on Argumentation Mining, pages 56–66, Denver, CO, USA.
28. Sibley, C., Coyne, J., and Sherwood, S. 2017. Research Considerations and Tools for Evaluating Human-Automation Interaction with Future Unmanned Systems. In Autonomy and Artificial Intelligence: A Threat or Savior? Springer, Cham, Switzerland.
29. Surdeanu, Mihai, Thomas Hicks, and Marco A. Valenzuela-Escarcega. Two Practical Rhetorical Structure Theory Parsers. Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies: Software Demonstrations (NAACL HLT), 2016.
30. Symeonidis, A.L., Chatzidimitriou, K.C., Athanasiadis, I.N.and Mitkas, P.A.. 2007. Data mining for agent reasoning: A synergy for training intelligent agents. Eng. Appl. Artif. Intell. 20, 8 (December 2007), 1097-1111.

31. Thimm, M. 2014. Tweety - A Comprehensive Collection of Java Libraries for Logical Aspects of Artificial Intelligence and Knowledge Representation. In Proceedings of the 14th International Conference on Principles of Knowledge Representation and Reasoning.
32. Tweety 2016. https://javalibs.com/artifact/net.sf.tweety.arg/delp. Last downloaded Dec 12.
33. van Eemeren, Frans H., Rob Grootendorst, and Francisca Snoeck Henkemans. 1996. Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments. Routledge, Taylor & Francis Group.
34. Walton, D. N., Reed, C., & Macagno, F. (2008) Argumentation Schemes. Cambridge: Cambridge University Press.