

Ontological representations of rhetorical figures for argument mining

Jelena Mitrović^{a,*}, Cliff O’Reilly^b, Miljana Mladenović^c and Siegfried Handschuh^a

^a Faculty of Computer Science and Mathematics, University of Passau, Germany

E-mails: Jelena.Mitrovic@uni-passau.de, Siegfried.Handschuh@uni-passau.de

^b Anglia Ruskin University, UK

E-mail: cliff.oreilly@anglia.ac.uk

^c Faculty of Mathematics, University of Belgrade, Serbia

E-mail: ml.miljana@gmail.com

Abstract. This paper surveys ontological modeling of rhetorical concepts, developed for use in argument mining and other applications of computational rhetoric, projecting their future directions. We include ontological models of argument schemes applying Rhetorical Structure Theory (RST); the RhetFig proposal for modeling; the related RetFig Ontology of Rhetorical Figures for Serbian (developed by two of the authors); and the Lassoing Rhetoric project (developed by another of the authors). The Lassoing Rhetoric venture is interesting for its multifaceted approach to linguistic devices, prominently including rhetorical figures, but also RST relations and stylistic models, like the use of historic present. This application takes a natural language text input and uses syntactic parsing tools to produce a knowledge base of linguistic entities using references to an OWL ontological framework, locating these devices using Semantic Web Rule Language (SWRL) logic rules. The paper also reports on a similar approach in research into detecting ironic tweets in a Serbian twitter corpus. The rhetorical schemes used for argument mining are also presented, as well as some suggestions for novel argument schemes based on the ontological approach to rhetorical figuration.

Keywords: Rhetorical figures, ontologies, argument schemes, SWRL, argument mining, computational rhetoric

1. Introduction

Rhetorical figures from *antimetabole* to *zeugma* (A–Z, considering their vast number and ubiquitous presence in text, speech, dialogue) have been rich resources for communication, literature and argument as long as language has existed, and they have been investigated for two and a half millennia. Certain rhetorical figures have their own argumentative functions. In this regard, the most interesting figures are *tropes* and *schemes*. *Tropes* are figures of speech with an unexpected twist in the meaning of words, as opposed to *schemes*, which only deal with patterns of words. That is, *schemes* are figures of speech that deal with word order, syntax, letters, and sounds, rather than the meaning of words. Most of the work done in modeling and detection of rhetorical figures so far has not taken into account their important argumentative roles which we discuss in this paper.

Perelman and Olbrechts-Tyteca looked at rhetorical figures as argumentative, thus making a distinction between “an embellishment, a figure of style and argumentative figures, figures that potentially affect the audience’s rational perception of a standpoint” [53]. Figures are not considered as merely decorative

*Corresponding author. E-mail: Jelena.Mitrovic@uni-passau.de.

1 but as a way to deal with the semantics of argumentative discourse. In their *Treatise on argumentation*, 1
2 the same authors give some hints about how dialog is mirrored in monologue through figures, when 2
3 mentioning contraries or some “quasi-logical” figures, for example, *antimetathesis* (or *antimetabole*), 3
4 *commutation*, *oxymoron*, or *antanaclasis* [54]. 4

5 Tindale wrote that rhetorical figures either constitute or emphasize arguments, arguing that *an-* 5
6 *timetabole*, for instance, involves the audience persuasively because of the rhythm of the discourse and 6
7 because the audience can easily see the reversal [61]. 7

8 Fahnestock tells us figuration is part of the theory of argument. At least some figures epitomize argu- 8
9 ments in their form. Thus, figures are seen as apt or epitomizing iconic forms for the arguments they 9
10 express [19]. Taking into account the argumentative function of rhetorical figures, explored at length 10
11 in the program of figural logic of Fahnestock [18], as well as in [61] and [30], we argue that inclusion 11
12 of formal models of rhetorical figures in argument mining systems would enhance the performance of 12
13 those systems. Different representations of rhetorical figure models might be utilized, some of which we 13
14 present in this paper. 14

15 This paper also explores arguments from the standpoint of argument schemes and different approaches 15
16 to their modeling. We take into account some overlapping terminology that appears both in Classical 16
17 rhetoric tradition and in the Rhetorical Structure Theory practice. While some researchers rightly point 17
18 out RST’s lack of connection to the rhetorical tradition, and the substantial difference between its coher- 18
19 ence relations and rhetorical figures [58], we look at the real similarities between figures and coherence 19
20 relations as fortuitous and discuss the ways in which they can be employed to enhance the existing argu- 20
21 ment mining systems, using our own approaches to knowledge engineering in the domain of rhetorical 21
22 analysis which involve both knowledge modeling and reasoning systems. 22

23 The fields of Computational Linguistics and Natural Language Processing have “exploded in recent 23
24 years” and provide “robust learning and processing systems applied to large corpora” [11]. In this do- 24
25 main, machine learning and stochastic approaches have become the main assets in default toolsets for 25
26 analysing text, with significantly positive results. Ontological approaches we believe, however, can add 26
27 value where statistical options cannot. Knowledge-based systems can be broadly divided into two areas: 27
28 “Knowledge models – formalisms for knowledge representation and reasoning, and Reasoning systems – 28
29 software that generates conclusions from available knowledge using logical techniques such as deduc- 29
30 tion and induction” [3]. Automatic detection of rhetorical figures in text and their annotation based on 30
31 machine learning and ontological approaches has been explored for *metaphors* [29,43,59], *epanaphora* 31
32 [60], *antimetabole* [15] *metonymy* [20,40,50], *irony* [9,47,63], *sarcasm* [13,23] and other figures [22,34], 32
33 but not enough attention has been given to using these approaches for enhancement of argument mining 33
34 systems. 34

35 We formalise the description of particular rhetorical domains which can aid argument mining efforts 35
36 in two styles in this paper. The first is the *RetFig* formal ontology of rhetorical figures approach, and the 36
37 second is the ontology suite within the *Lassoing Rhetoric* project. Both methods also employ reasoning 37
38 approaches using the Semantic Web Rule Language (SWRL) so as to discover new knowledge from 38
39 inference based upon logical rules. Ontological systems, by their nature, also fashion novel information 39
40 from intra-ontology inference, such as via the inheritance of classes and the propagation of types. 40

41 The remainder of the paper is organized as follows. Section 2 describes argument schemes and presents 41
42 their various applications. Section 3 outlines the approaches to rhetorical figuration modeling which can 42
43 be used for enhancement of argument mining systems. In Section 4, a composite figuration approach to 43
44 argument mining is proposed and analyzed. Finally, Section 5 presents our future work and gives many 44
45 concrete suggestions for future developments. 45
46

2. Argument schemes

2.1. Argument representations and rhetorical representations

Many forms of argument were identified and discussed by Aristotle. *Argument schemes* are patterns of human reasoning [64] used for both inventing and evaluating arguments and can be seen as historical descendants of Aristotle's Topics [1]. Chesñevar et al. [10] define argument schemes as argument forms that represent inferential structures of arguments used in everyday discourse, and in special contexts like legal and scientific arguments. According to Reed and Walton [57] argument schemes are forms of inference from premises to a conclusion of the kind used in everyday conversational exchanges in which one party is trying to get another to accept a conclusion at issue. They represent patterns of deductive and inductive reasoning in some instances, but typically they represent defeasible inferences of a kind that are useful heuristics for moving to a plausible hypothesis under conditions of uncertainty and lack of knowledge.

Schemas in *Rhetorical Structure Theory (RST)*, on the other hand, define the structural constituency arrangements of text. As a descriptive framework for text, RST provides features that can be useful in discourse studies. Representation of communicative roles of a text structure via schemas certainly aids in discourse representation for argument mining, although, like rhetorical figures, this was not their primary purpose. RST defines twenty-three relations that can hold between spans of a text. The distinction that RST makes between the part of a text which is the primary goal of the writer (*nucleus*), and the part that provides supplementary material (*satellite*), is crucial for the analysis of argumentative texts. Provided that these abstract patterns consist of a small number of text spans, a specification of relations between them and a specification of how nuclei are related to the entire collection, these characteristics make schemas somewhat similar to grammatical rules.

While RST was developed almost wholly independently from the rhetorical tradition, there are some interesting overlaps. For instance, antithesis is both an RST relation, and a rhetorical figure (a trope), and the two are virtually isomorphic. Antithesis, as an RST relation, is of interest for argument studies which has contrast in positive regard at its core. Mann and Thompson [41] give its functional representation regarding N (nucleus), S (satellite) and their constraints. N and S are in contrast; because of the incompatibility that arises from the contrast, one cannot have positive regard for both of those situations. For example, N can contain an idea that is favored by the author, and S, on the other hand, can contain an idea that is disfavored by the author. Like antithesis the trope, Mann and Thompson's relation is concerned with proximal opposing predictions. Treatments of figure, however, stress a wider range of functions. The RST relation is concerned primarily with quasi-logical and pragmatic relations, and its overall role in achieving text coherence, while the figural tradition stresses the creation of paradoxes, the generation of polarities and binaries in arguments, and the aesthetic function. We feel, however, that these are only differences of theoretical focus and that the relation and the figure are effectively interchangeable.

Schemes also refers to a group of rhetorical figures which represent an artful deviation from the ordinary arrangement of words, as defined in the Classical rhetoric tradition. In this regard, *antimetabole* is a figure belonging to this very important group for argument studies. Due to its distinctive syntactic and semantic profile, it is possible to use parallel phrasing and one set of opposing terms to predispose an audience toward seeing the other pair of terms in the phrases as opposites. A typical and often cited example of an *antimetabole* (along with some other rhetorical devices, all of which will be discussed in more detail later) can be found in President J.F. Kennedy's famous statement: "Ask not what your country can do for you, ask what you can do for your country."

1 To summarize, the task of argument mining can be approached by using three types of schemes, the
2 classical approach of using topoi, reimagined as argument schemes, the RST schemas and our goal
3 here is to show that a third type of schemes, *rhetorical* schemes, can enhance argument mining as well,
4 because they serve argumentative functions.

5 6 2.2. Applications of argument schemes 7

8 Various applications have utilized the capability of RST to describe and help generate argumentative
9 discourse. From a theoretical point of view Azar [2] investigates five RST relations (Evidence, Motiva-
10 tion, Justify, Antithesis, Concession) and their logical/pragmatic equivalents in the realm of arguments
11 (supportive, incentive, justifier, persuader). Strategies for generating evaluative arguments (i.e., argu-
12 ments that attempt to affect attitudes, as opposed to factual and casual arguments, which affect beliefs)
13 are discussed in [7] and [8]. In [67] RST was used to model and analyze a corpus of argumentative
14 discourse structures. Habernal, Eckle-Kohler and Gurevych [28] use a combined approach in which they
15 build upon Walton's argument schemes and combine them with the Claim–Premises scheme [52] and
16 argue that an annotation scheme for argument mining is a function of the task requirements and the
17 corpus properties. Green [24] developed a hybrid argument presentation scheme for biomedical research
18 that enables an analyst to encode argument analysis within the RST framework, which can be used to
19 represent the discourse structure of a text.

20 Many of the most common and important schemes have been identified and analyzed by [33,53,55,
21 66] and [65]. For argument mining purposes, each scheme is associated with a set of identifiers (key
22 words and markers locating premises and conclusions), and when the right grouping of identifiers is
23 located at some place in a text, the argument mining method locates it as an instance of an argument
24 of some particular, identifiable type (from a list of schemes) as shown in [66]. This approach was used
25 for argument mining with argument scheme structures in [39] and it was the basis of the ontological
26 approach to modeling arguments on the Web, as shown in [4].

27 Jeanne Fahnestock's work exploring rhetorical figures in scientific arguments [18] has laid a founda-
28 tion for looking into rhetorical and argumentative structures in highly technical, precise, and rigorous
29 arguments. Similarly, some of the most prominent discourse representation models focusing on argu-
30 ments at the intersection of RST and ontological modeling were developed with the goal of analyzing
31 highly technical, precise, and rigorous publications. Many of these approaches have used Semantic Web
32 standards as the data representation model for arguments. Examples include ScholOnto [6], works pre-
33 sented in [14] and [45], AIF and SALT.

34 ScholOnto (The Scholarly Ontologies project) is based on the decomposition of a scientific publica-
35 tion into elementary discourse knowledge items and their connection via a set of relations. The authors
36 provide a relatively simple set of argument links to make it as easy as possible to add an argument
37 link to a concept or claim. The discourse representation model proposed by de Waard and Tel [14] has
38 an arrangement they call *rhetorical block structure* for scientific publications. The same authors have
39 enriched the initial model by identifying seven basic discourse segment types: Fact, Hypothesis, Goal,
40 Method, Result, Implication, and Problem which correspond to the ones defined by Mizuta et al. [45],
41 who performed an automatic extraction of epistemic items using automated techniques based on argu-
42 mentative zoning, in a corpus of biomedical texts.

43 The AIF (Argument Interchange Format) is a core ontology of argument-related concepts which can
44 be extended to capture a variety of argument formalisms and argument schemes. The argument schemes
45 can be classified further into: rule of inference scheme, conflict scheme, preference scheme, etc. AIF
46

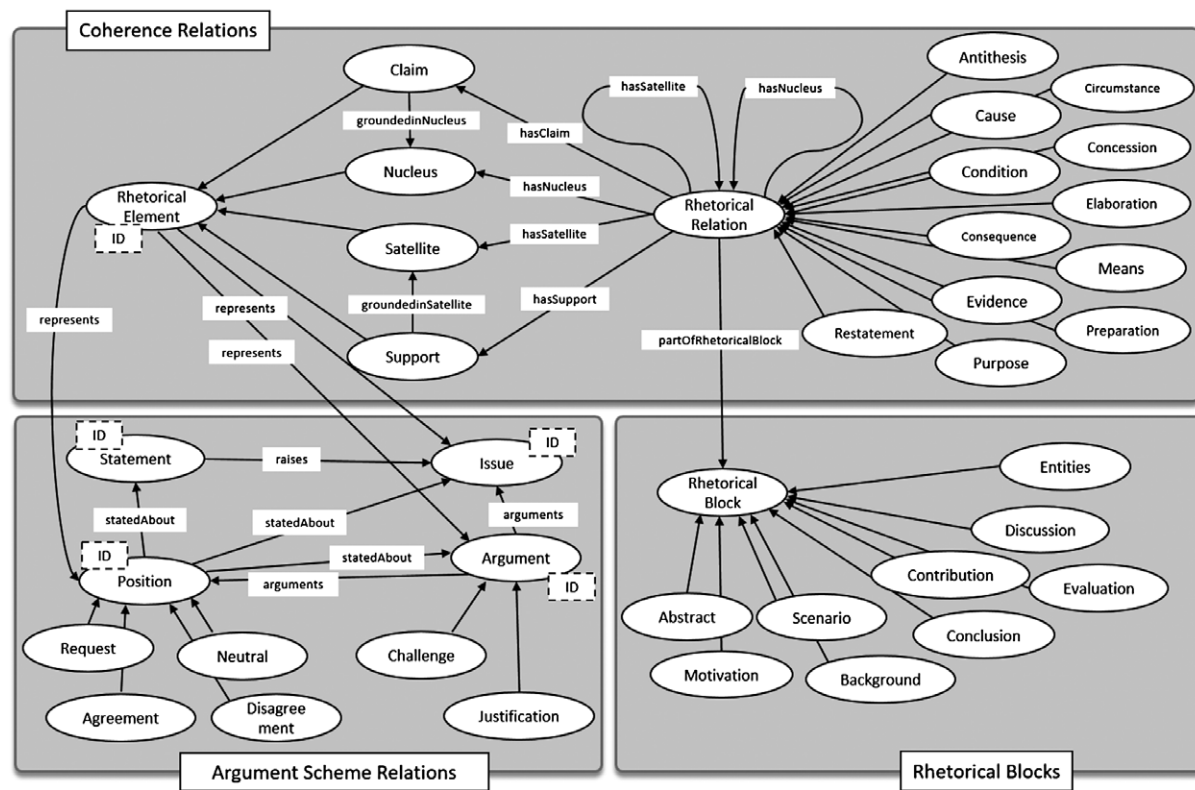


Fig. 1. SALT Rhetorical Ontology Layer (adapted from [25]).

was first introduced in [10]. The original AIF ontology had two disjoint types of nodes: information nodes and scheme nodes. Information nodes are used to represent passive information contained in an argument, such as a claim and premise, while scheme nodes capture the application of schemes (i.e., patterns of reasoning). The AIF description is available as a number of different specifications in different languages (such as OWL-DL, RDF-S and SQL).

The SALT Framework (Semantically Annotated L^AT_EX) [26] is a semantic authoring framework targeting the enrichment of scientific publications with semantic metadata. SALT adopts RST elements and models discourse knowledge items and their intrinsic coherence relations. The framework comprises three ontologies: the Document Ontology, the Rhetorical RST Ontology and the Annotation Ontology that connects the rhetorical structure present within the document's content to the actual content of the document. Applying SALT on a scientific publication leads to a local instance model capturing the inter-connected linear, rhetorical and argument structures within that publication.

The Rhetorical RST Ontology layer of this framework consists of three major sides, as diagrammed in Fig. 1: (i) coherence relations side that models elements (e.g., claims or supports) and the relations connecting them (e.g., Antithesis, Circumstance, Concession or Purpose); (ii) rhetorical blocks side that provides a coarse-grained structure for modeling the discourse (e.g., abstract, motivation, background or conclusion); (iii) argument side that captures the argument present in the publication via concepts like Issue, Position or Argument. This Rhetorical layer can be enriched with ontologically modeled separate rhetorical figures to enable fine-grained annotations of rhetorical figures which play an important part

in arguments, following approaches to figurational modeling used in [22,32] and [58], as well as our approach to modeling rhetorical figures, as described in Section 3.2 of this paper.

3. Approaches to rhetorical figuration modeling

3.1. Introduction

The discipline of ontological engineering defines many different approaches to conceptualizing a domain. A large body of ontologies are now utilising OWL¹ as an ontology language, for its expressiveness, its broad adoption by the W3C community and the tools and programming environments that are available. Protégé² is one of the major software ontology editors. It is popular, extensive and freely usable. The availability of tools and the need for ontological approaches in academia and commerce has driven a steady increase in the proliferation of ontologies to describe many diverse domains.

OWL is not the only suitable ontological framework for modeling natural language. In our research so far, only occasional workarounds have been needed to manage any inadequacies with practical implementations. However, the investigation of alternatives [36] should be within the scope of all projects that utilize techniques of ontological engineering.

Semantic Web Rule Language (SWRL) rules are Description Logic-based sets of terms that allow logical inference across a knowledge base. SWRL can also be understood as an extension to OWL which provides an efficient way for representing statements that are not expressed by an ontology itself but inferred from them. For example, from the following SWRL rule it can be inferred who is a parent – “a person who has at least one child”.

$$Person(?x), hasChild \min 1 Person(?x) \rightarrow Parent(?x) \quad (1)$$

Ontologies outlined in this paper are designed partly from a *bottom-up*, descriptive perspective, i.e., a word is a textual token. Ontologies are also designed from a *reverse engineering* perspective. Since a SWRL rule must include a set of classes and properties in order to infer something it follows that the ontology on which the rule is based must hold definitions of those same classes and properties. So we might start by designing a rule and then back-fill the entities required to make the rule possible.

The Semantic Web approach to developing a semantic network of readily-comparable and translatable data is still in its relative infancy, but has been growing more quickly in recent years [17]. This movement has opened the door to academic projects that seek to benefit from this approach and interoperability.

Two models of using ontological knowledge, and inferred knowledge based on the rules, in order to detect rhetorical figures in texts in a natural language are presented in the next two subsections. The first of them describes using a formal domain ontology of rhetorical figures for Serbian (The *RetFig* ontology) [48] and Serbian WordNet-based ontology (SWN) [37] in developing tools for automated detection of rhetorical figures in Serbian [46,47]. And the second is the *Lassoing Rhetoric* project that takes a text input and produces a knowledge base output that contains meaning representations of the text including inferred knowledge based on the rules that are applied via the SWRL processing phase [51]. The project was designed as a proof of concept application and therefore the domain has, so far, been investigated only to a limited extent.

¹<https://www.w3.org/OWL/>

²<http://protege.stanford.edu/>

3.2. Ontology-based rhetorical figures detection

A formal domain ontology of rhetorical figures for Serbian (The *RetFig* ontology) presented in [48] is a formal realization of the programs sketched in [32] and [35]. These programs were the basis for the development of methods and tools for automatic detection and annotation of rhetorical figures, as presented in [21,34,60] and [15]. In [60] detecting and classifying *epanaphora*, belonging to the rhetorical group of *schemes*, were treated as a constraint satisfaction problem which makes the greatest likelihood of detecting intentional *epanaphora*. In [21] a model of automatic detection of Schemes of repetition and the figure *oxymoron* belonging to the group of *tropes*, was proposed, and an appropriate tool “Jantor” for their automatic annotation was described. Results have shown that *schemes* can be successfully detected, because it is possible to define syntactic patterns that define them unambiguously. These patterns were based on syntactic and morphological operations within the grammatical rules of a natural language. For automatic detection of rhetorical figures belonging to the group of *schemes*, such as *anadiplosis*, *epanaphora*, *antimetabole* and *epiphora*, proposed in [34], regular expressions were used. In [15] a method for automatic detection of rhetorical figures schemes by introducing four features that indicate the existence of *chiasmus* (a scheme of reversal) was proposed. Tropes are much more problematic to detect, but Gawryjolek and his colleagues showed that some of them are tractable (oxymoron depends on proximal antonyms, which are easier to find than domain-sensitive tropes like metaphor and metonymy).

Taking into account that a rhetorical figure is a fundamentally linguistic phenomenon, in the *RetFig* ontology a top linguistic concept subsumes a rhetorical figure. A rhetorical figure is thus represented by *LinguisticScope* defining a syntax structure (paragraph, strophe, verse, sentence, phrase, word) in which it appears and *LinguisticObject* (verse, sentence, phrase, word) inside *LinguisticScope* whose transformation via *linguisticOperations* leads to the structure that can be detected as a certain rhetorical figure. Transformation can be done either over entire *LinguisticObject* or over its part. Therefore *LinguisticElement* defines what part of a linguistic object is being transformed. For example, if *LinguisticObject* is a *word* and *LinguisticElement* is a *letter* and *linguisticOperation* is *letterOmission*, the rhetorical figure can be: *aphaeresis*, if an omitted letter is the first letter in the word; *apocope*, if an omitted letter is the last letter in the word; *syncope*, if an omitted letter is not in the first or in the last position. Positional relation between *LinguisticElement* and *LinguisticObject* is introduced by the concept *LinguisticPosition*. Object properties *linguisticOperations* connect instances of *RhetoricalFigure* class and instances of *LinguisticElement* class to provide addition, omission, repetition, transposition, joining, separation and symmetrical positioning of a *LinguisticElement*.

A method of detection of rhetorical figures in the Serbian language, specifically *tropes*, was proposed in [46]. The method uses SWRL rules defined in the SWN ontology to identify whether a linguistic structure extracted from text satisfies some of the rules to be labeled as a rhetorical figure. The *RetFig* ontology was used for building the taxonomy of classes concerning *tropes* and SWRL rules defined in the SWN ontology were used for inference about candidate instances of these classes: *irony*, *simile*, *oxymoron* and *periphrasis*. First, a digital corpus of the contemporary Serbian language³ [62] was used for semi-automatic [46] and automatic [49] SWN ontology learning related to different forms of rhetorical figures from examples used in a natural language. Then, the proposed method was used for detection of figures *simile* and *irony*.⁴

³<http://korpus.matf.bg.ac.rs/prezentacija/korpus.html>

⁴While the Waterloo Taxonomy of Rhetorical Figures (utilized by the glossary provided in the introduction) treats irony as chroma [31], *RetFig* follows the more traditional classification of irony as a trope.

Table 1
Datatype properties of the SWN ontology classes

| SWRL rule | Relation chain | Antonymous pair |
|-----------|--|----------------------|
| 1 | happy <i>also_see</i> cheerful <i>near_antonym</i> cheerless | happy – cheerless |
| 1 | happy <i>also_see</i> joyous <i>near_antonym</i> joyless | happy – joyless |
| 1 | happy <i>also_see</i> joyful <i>near_antonym</i> sorrowful | happy – sorrowful |
| 1 | happy <i>also_see</i> glad <i>near_antonym</i> sad | happy – sad |
| 1 | happy <i>also_see</i> euphoric <i>near_antonym</i> dysphoric | happy – dysphoric |
| 1 | happy <i>also_see</i> contented <i>near_antonym</i> discontented | happy – discontented |
| 1 | happy <i>also_see</i> elated <i>near_antonym</i> dejected | happy – dejected |
| 1 | happy <i>also_see</i> felicitous <i>near_antonym</i> infelicitous | happy – infelicitous |
| 2 | happy <i>also_see</i> cheerful <i>near_antonym</i> cheerless <i>synonym</i> depressing | happy – depressing |
| 2 | happy <i>also_see</i> cheerful <i>near_antonym</i> cheerless <i>synonym</i> uncheerful | happy – uncheerful |
| 2 | happy <i>also_see</i> euphoric <i>near_antonym</i> dysphoric <i>synonym</i> distressed | happy – distressed |
| 2 | happy <i>also_see</i> contented <i>near_antonym</i> discontented <i>synonym</i> discontent | happy – discontent |

A tool which uses one SWRL rule to inference if a linguistic structure extracted from a text belonging to the collection of texts in Serbian (consisting of 10 digitized writings of various genres – children’s songs, fairy tales, comedies, novels and essays) represents the figure *simile* was presented in [49]. This ontology-based method for detection of the rhetorical figure *simile* achieved 51.8% accuracy. For automatic detection of *irony*, a set of over four hundred reasoning SWRL rules for extracting indirect antonymous pairs of words were defined [47]. The proposed model was evaluated (achieved accuracy 86.1%) on a collection of 1,732 tweets that had been manually annotated according to *irony*.

With the help of antonymous pairs, the real meaning of an ironic tweet can be comprehended, as in the following example: “O, baš sam srećan što je Dr Igi najavio svoj povratak!” (O, how happy I am to see Dr Iggy’s comeback!) where in the SWN ontology there are: a direct antonymous pair (happy – unhappy), eight indirect antonymous pairs inferred by SWRL rule

$$swn : also_see(?a, ?b) \wedge swn : near_antonym(?b, ?c) \rightarrow swn : irony(?a, ?c) \quad (2)$$

and four indirect antonymous pairs inferred by

$$swn : also_see(?a, ?b) \wedge swn : near_antonym(?b, ?c) \wedge swn : synonym(?c, ?d) \rightarrow swn : irony(?a, ?d) \quad (3)$$

The tweet can therefore be interpreted with an opposite of the intended meaning: “O, how cheerless I am to see Dr Iggy’s comeback!” or any one of the adjectives opposite to the adjective “happy” from the third column of the Table 1 can be used. As it was mentioned above, inferred examples of *irony* are figure candidates. Research presented in [34,47,60] and [15] pointed out the importance of literary devices near to the candidate figure (called either features, constraints or markers) which indicate the presence of the figure. For example, in [47] stylistic irony markers like punctuation marks, exclamation points, and italicized letters were used to indicate the existence of *irony*. In the previous tweet example, stylistic markers are the presence of emphasis “baš” (O, how) and the exclamation point.

The rhetorical figure *oxymoron* represents, by definition, merging of terms with opposite meanings into a new concept. Examples of *oxymoron* candidates found in SWN ontology are: “virtualna stvarnost” (virtual reality), “glasna tišina” (loud silence), “luda pamet” (crazy clever), “živi fosil” (living fossil),

“vatreni led” (fire-ice), etc. The rules for generating candidates for the figure *oxymoron* (4), (5) take into account adjectives that are mutual antonyms as candidates for instantiating of this figure when one of the adjectives is related by relations *derived_pos* or *be_in_state* with the corresponding noun. In SWN, relations *derived_pos* and *be_in_state* are lexical relations between synsets belonging to different parts-of-speech (cross-part of speech relations – XPoS). SWRL rules in the SWN ontology can be expressed by the following rules:

$$\text{near_antonym}(?p1, ?p2) \wedge \text{derived_pos}(?p2, ?n) \rightarrow \text{oxymoron}(?p1, ?n) \quad (4)$$

$$\text{near_antonym}(?p1, ?p2) \wedge \text{be_in_state}(?p2, ?n) \rightarrow \text{oxymoron}(?p1, ?n) \quad (5)$$

The *synonymy* relation can be used for obtaining candidates for instances of the *antonomasia* (*periphrasis*) class (6), which is a well-known descriptive phrase that replaces named entities. For example, we found that there are two phrases related to the city of Paris: “the capital of France” and “the City of Light”.

$$\begin{aligned} &\text{synonym}(?n1, ?n2) \wedge \text{hasDomain}(?n2, ?s) \wedge \text{swrlb} : \text{equal}(?s, \text{'administration'}) \\ &\rightarrow \text{antonomasia}(?n1, ?n2) \end{aligned} \quad (6)$$

3.3. A second model for inferential rhetorical figure discovery

An approach that attempts to infer the location and structure of rhetorical devices was developed for the *Lassoing Rhetoric* project in 2010. A computer application was developed that takes text inputs and parses and augments them to form a knowledge base of meaning representations. These formal structures that capture information about linguistic expressions have the potential to “bridge the gap from linguistic inputs to the non-linguistic knowledge of the world needed to perform tasks involving the meaning of linguistic inputs” [42]. Using the Semantic Web,⁵ more structure is added to the knowledge base, thereby increasing the complexity and coverage of meaning representations that describe the text. The project’s scope was as a proof of concept, developed primarily in order to show the possibilities of a semantic and pragmatic linguistic analysis using Linked Data⁶ as a source of knowledge, and drawing on the rhetorical tradition. Resources that conform and relate to the Semantic Web movement are used dynamically within the process of generating meaning representations, e.g., using DBPedia⁷ to determine the date of death of an individual. Any ontology or knowledge base that observes and co-ordinates with Semantic Web principles can be utilised in this way to provide contextual or world knowledge to analyses. A notable advance in this domain is the Linguistic Linked Open Data movement⁸ which groups and relates resources for Linguistic analysis that are openly available in order to create an online “collection of interrelated datasets [that are] available in a standard format, reachable and manageable by Semantic Web tools” [12].

To be able to structure the required elements in the knowledge base, a suite of ontologies was created. Separating OWL ontology URIs in this way allows for greater flexibility and more specific re-use across

⁵<https://www.w3.org/standards/semanticweb/>

⁶<https://www.w3.org/standards/semanticweb/data>

⁷<http://wiki.dbpedia.org/>

⁸<http://linguistic-lod.org/llod-cloud>

the Semantic Web. This design choice benefits users of the online resources and is an important paradigm of the Semantic Web project.

Pre-processing of the text data consists of a standard parsing routine (using GATE⁹ and the Stanford Parser¹⁰). After pre-processing, a knowledge base is constructed that instantiates various classes and properties contained in the suite of ontologies. The Protégé API¹¹ was used to programmatically generate the knowledge base. The ontologies were developed and encoded so as to provide functionally-distinct, but linked, resources of OWL classes, properties and instances. These are used, at runtime, by the Java program to generate an in-memory representation of the supplied text. This is referred to as the knowledge base. This representation is, in effect, a set of RDF triples that represents the instantiated classes and properties, located during the pre-processing phase of execution. Each ontology from the suite is listed, below.

1. The GATE processing output, including the Stanford Parser – *gate.owl*¹²
2. Linguistic entities and relationships – *langtag.owl*¹³
3. Document structure, language/text entities and relationships – *DocStruct.owl*¹⁴
4. External entities and relationships and associated intra-ontology relationships – *VerbNoun-Combo.owl*
5. Rhetorical devices, figures of speech and RST coherence relations – *RhetoricalDevices.owl*
6. The project ontology – *LassoingRhetoric.owl*

The knowledge base contains the various instantiated classes that relate to, firstly, linguistic entities: words, punctuation, sentence, etc. (*gate.owl*); grammatical structures: determiners, verbs etc. (*langtag.owl*) and document structure: properties such as *hasNextSentence* and *hasLastWord* etc. (*DocStruct.owl*).

Many rhetorical figures are structured around surface-level cues such as repeating phrases or token positioning. This is intuitively processed by a human reader without regard, but must be specified if a logic rule is to be used to infer anything on this basis. The Document Structure ontology (*DocStruct.owl*) is the specification for how entities in text relate to one another from a structural perspective. For example, each word in a document can have another that follows, and also one that precedes. This is encoded in the ontology by the OWL Object Properties of *hasNextWord* and *hasPreviousWord* (the Domain and Range of the property are the *gate:word* abstract class). Similarly, the properties of *hasNextSentence* and *hasPreviousSentence* help to describe how sentence-level structure is organised. A representation of the Document Structure ontology is shown in Fig. 2.

Two other ontologies in the suite are used to specify more complex structures and rhetorical patterns by name. The *VerbNounCombo.owl* ontology describes odd pairings of verbs and nouns. This is used primarily in the rule that attempts to discover the figure of *catachresis*, which is a “farfetched metaphor” [38] (the farfetched metaphor of the phrase *Lassoing Rhetoric* is inspired by this trope). A further example from Shakespeare’s *Timon of Athens* is “Tis deepest winter in Lord Timon’s purse”. This pattern describes an unusual combination of words that are used for effect. It is a subjective definition, but by

⁹<https://gate.ac.uk/>

¹⁰<http://nlp.stanford.edu/software/lex-parser.shtml>

¹¹<http://protege.stanford.edu/>

¹²<http://www.repositori.com/sw/onto/gate.owl>

¹³<http://www.repositori.com/sw/onto/langtag.owl>

¹⁴<http://www.repositori.com/sw/onto/DocStruct.owl>

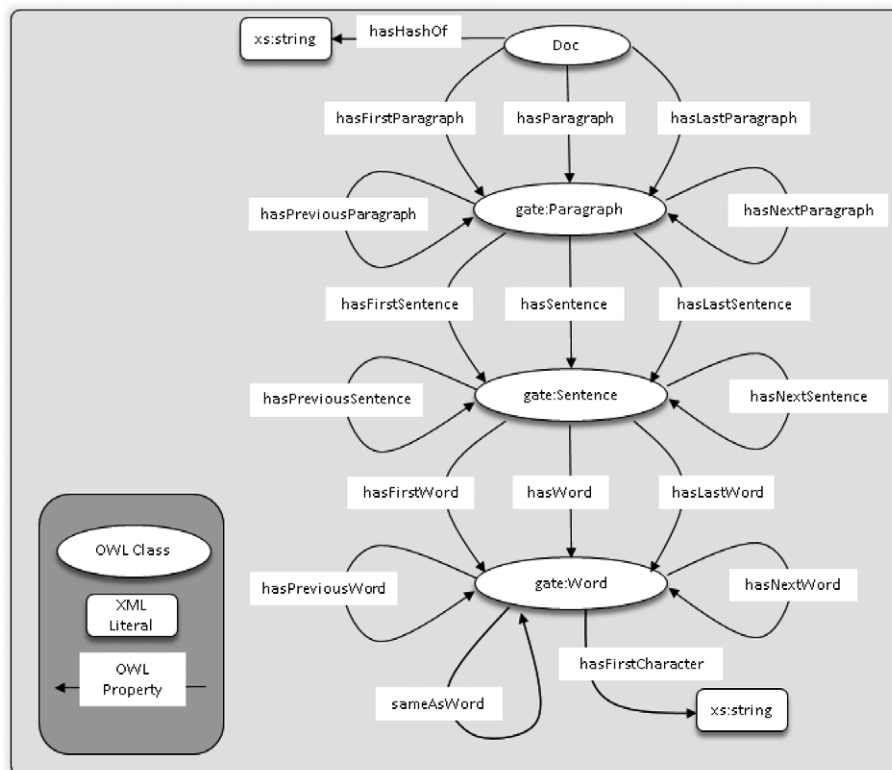


Fig. 2. Document Structure ontology class diagram.

highlighting unusual pairings it is possible to capture possible examples in text. This approach is similar to the idea of antonymous pairs, from [46].

The Rhetorical Devices ontology¹⁵ contains various classes for describing complex language structures as well as the `RhetoricalDevice` abstract class. The owl file also lists a number of instances of the `RhetoricalDevice` class, e.g., *epanaphora*, *anadiplosis*, *chiasmus* etc., as well as properties that relate them to a document, e.g., `hasRhetoricalDevice` (with Domain as `DocStruct:Doc` class and Range as the `RhetoricalDevice` class).

Lastly, each of the ontologies in the suite are imported, via OWL syntax, into a project-level ontology.¹⁶ This over-arching ontology also contains the SWRL rules stored in XML. Separating out the different elements of the work into owl files that are functionally specific allows other projects in the future to use elements that relate to their work without including entities that do not.

An example of a rule designed to discover a figure defined perhaps more readily by surface cues, is the rule for *epanaphora*. This SWRL rule seeks to discover and locate examples of repeating phrases, perhaps best exemplified by famous speeches. Winston Churchill's speech to Parliament on the 4th June 1940, in the midst of war, describes, by a repeated phrase, the emphatic stance to be conveyed... "We shall go on to the end. We shall fight in France, we shall fight on the seas and oceans...". In this notable example, the surface cue is a repeated phrase at the start of successive clauses, however the more general

¹⁵<http://www.repositori.com/sw/onto/RhetoricalDevices.owl>

¹⁶<http://www.repositori.com/sw/onto/LassoingRhetoric.owl>

figure of *epanaphora* can be more complex in practice and range over paragraphs, across clauses and concepts that are beyond the scope of the rule designed for this project. The rule defined here is limited to inferring the existence of the Rhetorical Device of *epanaphora* where two successive sentences have the first two words the same. A horn clause that represents the rule is shown, below.

$$\begin{aligned}
& \text{Doc}(\?h) \wedge \text{hasParagraph}(\?h, \?i) \wedge \text{hasSentence}(\?i, \?z) \wedge \text{word}(\?x) \\
& \wedge \text{hasNextWord}(\?x, \?y) \wedge \text{Sentence}(\?z) \wedge \text{hasFirstWord}(\?z, \?x) \wedge \text{word}(\?a) \\
& \wedge \text{hasNextWord}(\?a, \?b) \wedge \text{Sentence}(\?c) \wedge \text{hasFirstWord}(\?c, \?a) \\
& \wedge \text{hasNextSentence}(\?z, \?c) \wedge \text{hasString}(\?x, \?d) \wedge \text{hasString}(\?y, \?e) \\
& \wedge \text{hasString}(\?a, \?f) \wedge \text{hasString}(\?b, \?g) \wedge \text{swrlb: equal}(\?d, \?f) \\
& \wedge \text{swrlb: equal}(\?e, \?g) \wedge \text{hasStartNode}(\?x, \?j) \wedge \text{hasEndNode}(\?b, \?k) \\
& \Rightarrow \text{hasRhetoricalDevice}(\?h, \text{epanaphora}) \\
& \wedge \text{hasStartNode}(\text{epanaphora}, \?j) \wedge \text{hasEndNode}(\text{epanaphora}, \?k)
\end{aligned}$$

This is a simplified version of a rule that could be very much more complex if we wanted to capture further examples of the figure. In order to capture more examples of *epanaphora* the rule would need to be amended significantly, for example to capture phrases as well as single words that repeat over sentences, clauses, paragraphs etc.

Strictly not a rhetorical figure and perhaps best considered a *stylistic method* or a *move*, but as a further example we look at *historic present*, which describes a reference to a person or thing in the present tense who is no longer alive or has some element of anachronism. By referencing an instantiated class of **DeadPerson** and a verb in the present tense, we can determine whether these two events co-occur using the grammatical dependency chain that was produced from the parsing phase. The SWRL rule for discovering *historic present* is shown in horn clause form, below:

$$\begin{aligned}
& \text{Doc}(\?x) \wedge \text{paragraph}(\?y) \wedge \text{hasParagraph}(\?x, \?y) \wedge \text{Sentence}(\?z) \\
& \wedge \text{hasSentence}(\?y, \?z) \wedge \text{word}(\?a) \wedge \text{hasWord}(\?z, \?a) \wedge \text{word}(\?b) \\
& \wedge \text{hasWord}(\?z, \?b) \wedge \text{DeadPerson}(\?a) \wedge \text{hasDependency}(\?b, \?a) \\
& \wedge \text{NonThirdPersonSingularPresentVerb}(\?b) \wedge \text{hasStartNode}(\?a, \?c) \\
& \wedge \text{hasEndNode}(\?a, \?d) \\
& \Rightarrow \text{hasRhetoricalDevice}(\?x, \text{Historic_Present}) \\
& \wedge \text{hasStartNode}(\text{Historic_Present}, \?c) \\
& \wedge \text{hasEndNode}(\text{Historic_Present}, \?d)
\end{aligned}$$

Even though *historic present* doesn't fit neatly into a category of rhetorical figure, we include this example because it shows the benefits of drawing contextual information into our model. Many figures rely on background knowledge to form elements of their influence on a reader and therefore the technique of including this kind of knowledge in an automated and structured way can be quite powerful.

1 This SWRL set of terms is acting as a production rule, which indicates that the output is the generation
2 of a new set of RDF triples into the same knowledge base that describes the document. In this example,
3 if all the various terms in the body of the rule are held true, then it is logically asserted that the head is
4 held true also and therefore that the text document has an example of *historic present* and that it starts
5 and ends at the node values indicated. This is an important function of the model – that it generates new
6 knowledge about the text that was input and also records it in the output.

7 8 3.4. Inferential rhetorical figure discovery 9

10 The initial *Lassoing Rhetoric* project, again, was a proof of concept. Since that time, it has been ex-
11 tended only partially and is still very much a template for a fully-developed application to automatically
12 discover rhetorical forms and to record meaning representations of text documents. In effect, the set
13 of public domain documents processed and uploaded as well as the list of rules encoded in SWRL is
14 limited. The rhetorical figures with rules in place currently number thirteen and includes *adjunctio*, *al-*
15 *literation*, *antonomasia*, *apocope*, *asyndeton*, *catachresis*, *dirimens copulatio*, *epanaphora*, *epistrophe*,
16 *epizeuxis*, *historic present*, *parelcon* and *syncope*. There is future work in exploring other structures and
17 encoding rules that would infer their existence in text.

18 An important part of the output of the design is the HTML viewer. This approach to visualisation
19 is similar to others in the field, e.g., Michael Ullyot in his blog,¹⁷ and Gawryjolek's work [22], which
20 was similar in nature to the project outlined here. An HTML file and a Cascading Style Sheet file are
21 output along with the knowledge base, showing the original text marked up with the locations of each
22 discovered rhetorical form in a different colour. When Javascript code is added, this web page becomes
23 interactive so that the user can select the figures in which they are interested and display only those at
24 any one time. Figure 3 shows an example of the HTML output for a specifically contrived test example.

25 As well as the HTML output, a textual report of the elements calculated via the various processes is
26 also generated. This is important to allow external scrutiny of the process involved in calculating the
27 various elements in the text and the logical processes undertaken to reach the output. And further, a per-
28 centage figure is calculated to indicate how "rhetorical" the text is. This is obviously a very subjective
29 score and only simplistically calculated as a ratio of the phrases containing rhetorical forms discovered
30 by the rule system to those phrases not containing figures. Similar work in this area has been conducted
31 on the Faciloscope Project.¹⁸

32 Summarisation and quantification will, we believe, become an important feature of the field of com-
33 putational rhetoric in the future. The desire to understand the overall form of an author's persuasive
34 intent is palpable in both rhetorician and non-rhetorician alike. The level at which quantification can be
35 applied varies across a document and even to individual figures or types of figures. Chiastic figures (that
36 is, figures of the Chiastic Suite, as defined in the glossary appended to the introduction of this issue)
37 are those which have a criss cross nature at, for example, lexical level. Dubremetz and Nivre [15,16]
38 analysed chiastic figures in terms of the extent to which they vary across this general prototypical form.
39 By ranking the inverse lexical repetition configurations they were able to show quantified variation of
40 *chiasmus* across various corpora.

41 This type-ranking process of quantification naturally leads to further understanding of the prototypes
42 and components of a figure such as *chiasmus*. For example, two figures: *parison*, which is a parallel
43 syntactic structure, and *antithesis*, which is characterised by proximal opposing predications, can be

44 ¹⁷<http://ullyot.ucalgaryblogs.ca/2015/05/31/augmented-criticism-and-rhetorical-figures/>

45 ¹⁸<http://faciloscope.cal.msu.edu/facilitation/>



Fig. 3. Lassoing Rhetoric – HTML output example.

additionally quantified by the *chiasmus* calculations when considering other elements such as parts-of-speech and lexical clues.

The *Lassoing Rhetoric* project's aim is to take a text input, parse and augment the contents and upload the resulting representations to the Internet so that knowledge can be shared over the Semantic Web. In addition, by reference to a suite of ontologies, the existence and location of rhetorical forms can be inferred from a set of rules that apply across each knowledge base. This method has been shown to work, within a limited corpus and for a restricted domain.

Since the structure of the project allows for complete contrivance towards the goal, i.e., a rule can be defined in advance and subsequently the ontological structure put in place along with the computational discovery logic to populate the knowledge base, it is likely that more complex rhetorical forms can be processed and located in text. The possibilities are huge compared with other techniques where complex logical structures have to be interpreted or learned, for example, in machine learning algorithms. This is an area for future development.

The potential benefits of this system to general knowledge cannot be understated. It is possible to create semantically-tagged representations of any textual work (where language parsers exist) that relate to a growing, coherent set of concepts and world knowledge data. The uses for this approach, from corpus analysis to rhetorical figuration scholarship, are not yet fully investigated.

It is recognised that the Semantic Web movement hasn't yet fully realised the potential that was originally promised. Many of the reasons for this were predicted, however the progress is still moving forward and it is hoped that a semantically rich network of data will soon start to make more sense as a real prospect. In the meantime, the proof of concept *Lassoing Rhetoric* project stands as a starting point for further work.

4. Argument mining improvements based on ontological rhetorical models

Argument models differ in granularity, expressive power, and other properties. Because of the lack of a single general-purpose argument model there is a need for using different schemes for different scenarios. [56] has shown that many schemes are structurally related in an interesting way – many of the most common schemes have an interlocking relationship with other schemes so that one scheme can be classified as a subspecies of another, but only in a complex manner. Building a robust model for argument mining based on combining different approaches to acquiring computational understanding of the way arguments work is a goal which can, in part, be achieved by combining different methods for designing the building blocks of the arguments themselves. Adding rhetorical figures to these models would increase granularity and leverage two and half millennia of formal study, because the arguments drawn from the figures according to [61], engage the audience at a deep, often emotional level, before reason moves in as an organizing force. Thus, formal models of rhetorical figures used for enhancing argument mining systems are expected to bring better results in performance of these systems.

Composite effects of rhetorical figures have always been known in terms of an increased aesthetic appeal in some configurations. New areas of rhetorical research are being discovered now, almost three millennia into the scholarship on figures. Composite figuration makes detection more reliable and constrains functions and is therefore a boon to argument mining and other computational-textual activities [31,58].

“And so, my fellow Americans: ask not what your country can do for you – ask what you can do for your country.”

This famous statement uttered during the inaugural speech address of the 35th president of the United States of America, John F. Kennedy, can be seen as an argument scheme of its own, in the ontological rhetoric modeling sense. It is considered as a highly persuasive statement. It contains three distinct rhetorical figures: the whole statement is an *antimetabole*, yet, it also contains a *mesodiplosis* (“can do for”) and an *antithesis*, the contrary propositions “Ask not ...” and “Ask ...”. The three figures carry the functional freight, but other figures can also be spotted here. Hyperbaton (evoking archaic or biblical phrasing) seen in “Ask not” instead of using the contemporary “Do not ask. . .”. The loose epanaphora of “Ask not what” and “ask what”, as well as the parison of “what your country can do for you” and “what you can do for your country” [31]. Kennedy’s statement can also be looked at in terms of classical argument scheme modeling, such as the one proposed by Walton [64], but it is not quite clear at this point how that scheme would be defined. One solution would be a scheme consisting of a *claim* and *rebuttal* part, provided that rebuttal requires a prior argument that it is directed against and the rebuttal itself is an argument that is directed against this prior argument. These conditions are met in president Kennedy’s statement, but considering its semantic complexity, a special kind of scheme could be invented, after careful deliberation.

Still, detecting composite rhetorical devices such as these is indispensable and can lead to a better understanding of the way arguments are formed, which can secure a more straightforward machine analysis and understanding of those arguments. Ruan et al. [58] analyze this famous sentence and argue that it is so widely known and frequently invoked because of the “schematic congruence with which the form matches the Reject–Replace function its arrangement serves” and because of the cognitive affinities humans have for the structural properties of this kind of figural stacking – opposition, repetition, and symmetry. The same authors also point out that this kind of statement is especially interesting because of its form-functional correlation, stemming from the program of figural logic of Jeanne Fahnestock.

1 Thus, the form of this famous Kennedy's statement (written by his speech writer, Ted Sorensen) makes 1
2 it tractable for automated detection, while the function accounts for its rhetorical purpose. Detection of 2
3 composite rhetorical figures which appear in the statement itself can be achieved following the ontolog- 3
4 ical modeling approach presented in Sections 3.2 and 3.3. 4

5 Imagining further the nature of mixtures of rhetorical figures (which is surely what most documents 5
6 are), we foresee a fruitful effort in combining analytical and statistical ranking methods [15] with a 6
7 larger-document approach [51] and an approach to annotating stacks of figures proposed in [58]. Un- 7
8 der such scrutiny, we predict that the models we have for figures will uncover a greater complexity 8
9 of foundational constituents (such as parallelism and symmetry) and complementary elements (such as 9
10 semantic and pragmatic augmentation) that brings insights into the cognitive basis for understanding. 10
11 Another interesting approach by Mehlenbacher [44] proposes that *prolepsis*, the figure of anticipation, 11
12 can also be seen as an argument strategy and an argument scheme in and of itself. She proposes the 12
13 proleptic suite as a preliminary map for special argument schemes which would also be beneficial for a 13
14 combined approach to argument mining using rhetorical figure modeling. 14
15

16 5. Future work 16

17 We summarise four areas of future development: 17
18

19 5.1. Broadening of logic rules 19

20 In this paper we showed that semantic knowledge of domain ontologies like the *RetFig* ontology of 20
21 rhetorical figures, and ontologies such as the WordNet-based ontology, can be used for detecting *tropes* 21
22 by defining inference rules. We showed that tropes whose effect is based on antonymy and synonymy 22
23 of words can be detected in this way. In future work, we plan on broadening the set of antonymous 23
24 pairs using ontological rules in which different lexical relations, existing in WordNet-based ontology 24
25 participate. We also plan to expand the set of ontological rules to include procedures of generalization 25
26 and specifications of concepts that are already included by the rules laid down by this system. This 26
27 system will also be used for detection of *litotes*, *antithesis*, as well as for modeling an argument scheme 27
28 model consisting of rhetorical figures which appear in unison in some of the strongest arguments given 28
29 in political speeches, as presented in Section 4 of this paper. Therefore, the discovery and description of 29
30 composite rhetorical figure schemes is an area of obvious future investigation. As remarked previously, 30
31 a complex rhetorical interplay can sometimes be found in political statements. A set of logic rules, 31
32 perhaps run in sequence and consecutively in order to build on the inferences of previous analyses, 32
33 would be an interesting next step in bringing together work across different domains of Computational 33
34 Rhetoric. An obvious choice for a data set on which all of these improvements could be tested is a Corpus 34
35 of Political Speeches such as the one presented in [27]. We will also look into corpora of web arguments, 35
36 such as the ComArg corpus of online discussions [5]. Arguments obtained from user-generated Web 36
37 content often lack proper substantiation and reasoning and are usually ambiguous, implicit and poorly 37
38 worded, but the Web provides us with a steady flow of subjective information and remains the largest 38
39 source of opinions. 39
40

41 5.2. Publishing online and defining standards 41

42 Several projects were reported in this paper, each taking a slightly different approach. These applica- 42
43 tions discover rhetorical forms from natural language and by using OWL ontologies and SWRL rules, 43
44

1 are able to extend the work by publishing the results to the internet and, in tandem, the ontological
2 framework and SWRL rules. Ultimately this allows for unprecedented sharing and collaboration. Other
3 projects in this field we recognise are Ruan et al. [58] and ROAP.¹⁹

4 A specific problem to be addressed in the near future is agreeing to a standard for ontological ap-
5 proaches in this domain. There are a number of groups working on ontologies of rhetorical figures that
6 have many shared concepts [58]. These concepts have been modeled in different ways, depending on
7 the application or language within which the work is undertaken, but a translation mechanism and set
8 of agreements would be beneficial. Defining, publishing and sharing these common elements is essen-
9 tial while these projects remain relatively young. Once knowledge bases of meaning representations for
10 specific texts are published online, there is also a need to define these in a form that is re-usable by
11 others in this or associated fields. For example, marked-up corpora that describe a text would be more
12 useful if they were in a standard and recognised format. There is a wealth of public domain and research-
13 friendly text resources available. The project aimed at parsing these documents and processing them into
14 re-usable knowledge bases would be useful for many others.

15 16 5.3. *Improvements in effectiveness*

17
18 A number of challenges are becoming more clear as work in this field develops, especially around the
19 writing of logic rules and ontology design. Utilising statistical techniques in tandem with ontological
20 and inferential ones, we would expect improvements in effectiveness. A future area for investigation
21 would therefore be machine learning augmentation of ontological approaches. This is a huge area for
22 investigation, but some specific examples are:

- 23 ● executing a supervised learning algorithm on a corpus of known figures to produce probabilistic
24 models of key terms that can be used in logic rules
- 25 ● developing inter-inference filters that operate alternatively and in tandem with an inferential ap-
26 proach, e.g., between iterative applications of rule-based logical inferences each rule set is aug-
27 mented by the results of probabilistic models
- 28 ● executing a clustering algorithm over a text to highlight rudimentary terms to be collated into rule
29 sets

30
31 An even more challenging area of investigation is pragmatic augmentation. Many facets of under-
32 standing rely on contextual knowledge. Therefore increasing the quantity and scope of world knowledge
33 injection into rules and knowledge bases is an area that would also be fruitful. The Linked Data move-
34 ment has grown considerably in the last five years and it is growing ever plausible to create useful links
35 across data sets.

36 37 5.4. *Potential applications*

38
39 As discussed earlier in this paper, there are numerous and potentially fruitful areas to investigate
40 regarding quantification of rhetorical figures at various levels. The hypothesis that an author has a distinct
41 profile of quantified figures is one that could also be very useful, especially for the following ideas for
42 applications:

- 43 ● author recognition and discovery

44
45 ¹⁹<http://bakulf.github.io/raop/>

- plagiarism detection
- comparing authors and speech-writers
- political discourse analysis
- argument/intent analysis, e.g., hate speech detection

The rudimentary score developed in the *Lassoing Rhetoric* project should be extended into a more appropriate set of metrics by which to measure documents. It is imagined that this set of measurements would be a complex, but rewarding project to realise.

As we mentioned previously in this paper, we hope that further analysis of rhetorical figures will reveal a deeper understanding of the cognitive aspects that underlie the impact of persuasive communication. A conjunction of domains such as psychology, cognitive science, philosophy, computer science and rhetoric provide a rewarding furrow to plough.

Acknowledgement

A series of workshops on Computational Rhetoric held at the University of Waterloo strive to bring together researchers from the fields of Rhetoric, Computer Science, Linguistics, Cybernetics, Information Science, Literary studies, Psychology and Philology. This paper was inspired by the presentations and rich discussions experienced at the last Computational Rhetoric workshop of the University of Waterloo, held in August 2016.

Acknowledgements

References

- [1] Aristotle, *Topics*, Oxford University Press, 1958.
- [2] M. Azar, Argumentative text as rhetorical structure: An application of rhetorical structure theory, *Argumentation* **13**(1) (1999), 97–114. doi:10.1023/A:1007794409860.
- [3] M. Bergman, KBAI, 2014, <http://www.mkbergman.com/1816/knowledge-based-artificial-intelligence/>, accessed 2017-01-10.
- [4] T. Blount, D.E. Millard and M.J. Weal, An ontology for argumentation on the social web: Rhetorical extensions to the AIF, in: *Computational Models of Argument – Proceedings of COMMA 2016*, Potsdam, Germany, 12–16 September, 2016, 2016, pp. 119–126. doi:10.3233/978-1-61499-686-6-119.
- [5] F. Boltužić and J. Šnajder, Back up your stance: Recognizing arguments in online discussions, in: *Proceedings of the First Workshop on Argumentation Mining*, Association for Computational Linguistics, Baltimore, Maryland, 2014, pp. 49–58. doi:10.3115/v1/W14-2107.
- [6] S.J. Buckingham Shum, V. Uren, G. Li, B. Sereno and C. Mancini, Modeling naturalistic argumentation in research literatures: Representation and interaction design issues, *International Journal of Intelligent Systems*. **22**(1) (2007), 17–47. doi:10.1002/int.20188.
- [7] G. Carenini and J.D. Moore, A strategy for generating evaluative arguments, in: *Proceedings of the First International Conference on Natural Language Generation – Volume 14. INLG'00*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2000, pp. 47–54. doi:10.3115/1118253.1118261.
- [8] G. Carenini and J.D. Moore, Generating and evaluating evaluative arguments, *Artificial Intelligence*. **170**(11) (2006), 925–952. doi:10.1016/j.artint.2006.05.003.
- [9] P. Carvalho, L. Sarmento, J. Teixeira and M.J. Silva, Liars and saviors in a sentiment annotated corpus of comments to political debates, in: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers – Volume 2. HLT'11*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2011, pp. 564–568, available from: <http://dl.acm.org/citation.cfm?id=2002736.2002847>.
- [10] C. Chesñevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari et al., Towards an argument interchange format, *Knowl Eng Rev*. **21**(4) (2006), 293–316.

- [11] A. Clark, C. Fox and S. Lappin, *The Handbook of Computational Linguistics and Natural Language Processing*, John Wiley & Sons, 2013.
- [12] Consortium WWW, Data – W3C, <https://www.w3.org/standards/semanticweb/data>, accessed 2016-11-29.
- [13] D. Davidov, O. Tsur and A. Rappoport, Semi-supervised recognition of sarcastic sentences in Twitter and Amazon, in: *Proceedings of the Fourteenth Conference on Computational Natural Language Learning, CoNLL'10*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2010, pp. 107–116, available from: <http://dl.acm.org/citation.cfm?id=1870568.1870582>.
- [14] A. de Waard and G. Tel, The ABCDE format – enabling semantic conference proceedings, in: *Proceedings of 1st Workshop: SemWiki2006 – from Wiki to Semantics*, Budva, Montenegro, 2006.
- [15] M. Dubremetz and J. Nivre, Rhetorical figure detection: The case of chiasmus, in: *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature*, Denver, CO, 2015.
- [16] M. Dubremetz and J. Nivre, Syntax matters for rhetorical structure: The case of chiasmus, in: *Proceedings of the Fifth Workshop on Computational Linguistics for Literature*, Association for Computational Linguistics, San Diego, California, USA, 2016, pp. 47–53. doi:10.18653/v1/W16-0206.
- [17] I. Ermilov, J. Lehmann, M. Martin and S. Auer, LODStats: The data web census dataset, in: *International Semantic Web Conference*, Springer, 2016, pp. 38–46.
- [18] J. Fahnestock, *Rhetorical Figures in Science*, Oxford University Press, 1999.
- [19] J. Fahnestock, Figures of argument, *Informal Logic*. **24**(2) (2004), 115–135. doi:10.22329/il.v24i2.2139.
- [20] R. Farkas, E. Simon, G. Szarvas and D. Varga, Maxent metonymy resolution, in: *Proceedings of the 4th International Workshop on Semantic Evaluations, SemEval'07*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 161–164, available from: <http://dl.acm.org/citation.cfm?id=1621474.1621507>. doi:10.3115/1621474.1621507.
- [21] J. Gawryjolek, C. Di Marco and R.A. Harris, An annotation tool for automatically detecting rhetorical figures – system demonstration, in: *Proceedings of the IJCAI-09 Workshop on Computational Models of Natural Argument (CMNA IX)*, Pasadena, CA, 2009.
- [22] J.J. Gawryjolek, *Automated Annotation and Visualization of Rhetorical Figures*, University of Waterloo, Canada, 2009.
- [23] R. González-Ibáñez, S. Muresan and N. Wacholder, Identifying sarcasm in Twitter: A closer look, in: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers – Volume 2, HLT'11*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2011, pp. 581–586, available from: <http://dl.acm.org/citation.cfm?id=2002736.2002850>.
- [24] N.L. Green, Representation of argumentation in text with rhetorical structure theory, *Argumentation* **24**(2) (2010), 181–196. doi:10.1007/s10503-009-9169-4.
- [25] T. Groza, *Advances in Semantic Authoring and Publishing*, Vol. 13, IOS Press, 2012.
- [26] T. Groza, H.L. Kim and S. Handschuh, SALT: Enriching LATEX with semantic annotations, in: *Proceedings of the 5th International Semantic Web Conference (ISWC 2006)*, Athens, GA, 2006.
- [27] M. Guerini, C. Strapparava and O. Stock, CORPS: A corpus of tagged political speeches for persuasive communication processing, *Journal of Information Technology & Politics*. **5**(1) (2008), 19–32. doi:10.1080/19331680802149616.
- [28] I. Habernal, J. Eckle-Kohler and I. Gurevych, Argumentation mining on the web from information seeking perspective, in: *Proceedings of the Workshop on Frontiers and Connections Between Argumentation Theory and Natural Language Processing*, Forlì-Cesena, Italy, 2014.
- [29] A. Hardie, V. Koller, P. Rayson and E. Semino, Exploiting a semantic annotation tool for metaphor analysis, in: *Proceedings of the Corpus Linguistics 2007 Conference*, M. Davies, P. Rayson, S. Hunston and P. Danielsson, eds, 2007.
- [30] R. Harris, Figural logic in Gregor Mendel's experiments on plant hybrids, *Philosophy and Rhetoric*. **46**(4) (2013), 570–602. doi:10.5325/philrhet.46.4.0570.
- [31] R.A. Harris, The fourth master trope, antithesis, in: *Meeting of the Rhetoric*, Society of America, 2016.
- [32] R.A. Harris and C. Di Marco, Constructing a rhetorical figuration ontology, in: *Symposium on Persuasive Technology and Digital Behavior Intervention, Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)*, Edinburgh, 2009, pp. 47–52.
- [33] A. Hastings, A reformulation of the modes of reasoning in argumentation, PhD dissertation, University of Evanston, Illinois, 1963.
- [34] D.D. Hromada, Initial experiments with multilingual extraction of rhetoric figures by means of PERL-compatible regular expressions, in: *Proceedings of the Student Research Workshop Associated with the 8th International Conference on Recent Advances in Natural Language Processing, RANLP 2011*, Hissar, Bulgaria, 13 September, 2011, 2011, pp. 85–90, available from: <http://www.aclweb.org/anthology/R11-2013>.
- [35] A.R. Kelly, N.A. Abbott, R.A. Harris and C. Di Marco, Toward an ontology of rhetorical figures, in: *Proceedings of the 28th ACM International Conference on Design of Communication, SIGDOC'10*, ACM, New York, NY, USA, 2010, pp. 123–130. doi:10.1145/1878450.1878471.
- [36] M. Krötzsch and V. Thost, Ontologies for knowledge graphs: Breaking the rules, in: *International Semantic Web Conference*, Springer, 2016, pp. 376–392.

- [37] C. Krstev, *Processing of Serbian: Automata, Texts and Electronic Dictionaries*, University of Belgrade, Faculty of Philology, 2008.
- [38] R. Lanham, *A Handlist of Rhetorical Terms*, 2nd edn, University of California Press, 1991.
- [39] J. Lawrence and C. Reed, Argument mining using argumentation scheme structures, in: *Computational Models of Argument – Proceedings of COMMA 2016*, Potsdam, Germany, 12–16 September, 2016, 2016, pp. 379–391. doi:10.3233/978-1-61499-686-6-119.
- [40] J. Leveling, FUH (FernUniversität in Hagen): Metonymy recognition using different kinds of context for a memory-based learner, in: *Proceedings of the 4th International Workshop on Semantic Evaluations. SemEval'07*, Association for Computational Linguistics, Stroudsburg, PA, USA, 2007, pp. 153–156, available from: <http://dl.acm.org/citation.cfm?id=1621474.1621505>. doi:10.3115/1621474.1621505.
- [41] W.C. Mann and S.A. Thompson, *Rhetorical Structure Theory: A Theory of Text Organization*, USC Information Sciences Institute, 1987.
- [42] J.H. Martin and D. Jurafsky, *Speech and Language Processing*, International Edition, 2000, p. 710.
- [43] Z.J. Mason, CorMet: A computational, corpus-based conventional metaphor extraction system, *Comput Linguist.* **30**(1) (2004), 23–44. doi:10.1162/089120104773633376.
- [44] A.R. Mehlenbacher, Rhetorical figures as argument schemes – the proleptic suite, *Argument & Computation*, This issue.
- [45] Y. Mizuta, A. Korhonen, T. Mullen and N. Collier, Zone analysis in biology articles as a basis for information extraction, *International Journal of Medical Informatics* **75**(6) (2006), 468–487. doi:10.1016/j.ijmedinf.2005.06.013.
- [46] M. Mladenović, Ontology-based recognition of rhetorical figures, *Infotheca, Journal for Digital Humanities.* **16**(1–2) (2016), 24–47. doi:10.18485/infotheca.2016.16.1_2.2.
- [47] M. Mladenović, C. Krstev, J. Mitrović and R. Stanković, Using lexical resources for irony and sarcasm classification, in: *8th Balkan Conference in Informatics*, Skopje, Macedonia, 2017.
- [48] M. Mladenović and J. Mitrović, Ontology of rhetorical figures for Serbian, in: *Proceedings of Text, Speech, and Dialogue – 16th International Conference, TSD2013*, 2013, pp. 386–393. doi:10.1007/978-3-642-40585-3_49.
- [49] M. Mladenović, J. Mitrović and C. Krstev, A language-independent model for introducing a new semantic relation between adjectives and nouns in a WordNet, in: *Proceedings of Eight Global WordNet Conference, GWC2016*, 2016, pp. 218–225.
- [50] C. Nicolae, G. Nicolae and S.M. Harabagiu, UTD-HLT-CG: Architecture for metonymy resolution and classification of nominal relations, in: *Proceedings of the 4th International Workshop on Semantic Evaluations, SemEval@ACL 2007*, Prague, Czech Republic, June 23–24, 2007, 2007, pp. 454–459, available from: <http://aclweb.org/anthology/S/S07/S07-1101.pdf>. doi:10.3115/1621474.1621575.
- [51] C. O'Reilly and S. Paurobally, Lassoing Rhetoric with OWL and SWRL, 2010, available http://www.academia.edu/2095469/Lassoing_Rhetoric_with_OWL_and_SWRL.
- [52] A. Passant, P. Ciccarese, J.G. Breslin and T. Clark, SWAN/SIOC: Aligning scientific discourse representation and social semantics, in: *Proceedings of the Workshop on Semantic Web Applications in Scientific Discourse (SWASD 2009) at (ISWC-2009)*, Vol. 523, IOS Press, 2009.
- [53] C. Perelman and L. Olbrechts-Tyteca, *The New Rhetoric: A Treatise on Argumentation*, University of Notre Dame Press, 1969.
- [54] C. Plantin, A place for figures of speech in argumentation theory, *Argumentation* **23**(3) (2009), 325–337. doi:10.1007/s10503-009-9152-0.
- [55] J.L. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA, USA, 1995.
- [56] H. Prakken, On the nature of argument schemes, in: *Dialectics, Dialogue and Argumentation an Examination of Douglas Walton's Theories of Reasoning and Argument*, 2010, pp. 167–185.
- [57] C. Reed and D. Walton, Argumentation schemes in dialogue, in: *Dissensus and the Search for Common Ground (Proceedings of OSSA 2007)*, H.V. Hansen, C.W. Tindale, R.H. Johnson and J.A. Blair, eds, 2007.
- [58] S. Ruan, C. Di Marco and R.A. Harris, Rhetorical figure annotation with XML, in: *Computational Models of Natural Argumentation (CMNA) 16, a Workshop at the 2016 International Joint Conference on Artificial Intelligence (IJCAI)*, New York, 2016.
- [59] E. Shutova, S. Teufel and A. Korhonen, Statistical metaphor processing, *Computational Linguistics.* **39**(2) (2013), 301–353. doi:10.1162/COLI_a_00124.
- [60] C. Strommer, Using rhetorical figures and shallow attributes as a metric of intent in text [supervised by Chrysanne DiMarco; Randy Allen Harris, Committee Member], PhD dissertation, Cheriton School of Computing, University of Waterloo, 2011.
- [61] W.C. Tindale, *Rhetorical Argumentation: Principles of Theory and Practice*, Sage Publications, 2004.
- [62] M. Utvić, Construction of a reference corpus of contemporary Serbian language, PhD dissertation, University of Belgrade, Faculty of Linguistics, 2014.

- 1 [63] T. Veale, Detecting and generating ironic comparisons: An application of creative information retrieval, in: *Artificial* 1
2 *Intelligence of Humor, Papers from the 2012 AAAI Fall Symposium*, Arlington, Virginia, USA, 2012, pp. 2–4, available 2
3 from: <http://www.aaai.org/ocs/index.php/FSS/FSS12/paper/view/5557>. 3
4 [64] D. Walton, What is reasoning? What is an argument?, *Journal of Philosophy* **87**(8) (1990), 399–419. 4
5 [65] D. Walton and F. Macagno, A classification system for argumentation schemes, *Argument & Computation*. **6**(3) (2015), 5
6 219–245. doi:10.1080/19462166.2015.1123772. 6
7 [66] D.N. Walton, *Argumentation Schemes for Presumptive Reasoning*, L. Erlbaum Associates, 1996. 7
8 [67] T. Zidrasco, S. Shiramatsu, J. Takasaki, T. Ozono and T. Shintani, Building and analyzing corpus to investigate appro- 8
9 priateness of argumentative discourse structure for facilitating consensus, in: *International Conference on Industrial,* 9
10 *Engineering and Other Applications of Applied Intelligent Systems*, LNAI, Vol. 6097, Springer, 2010, pp. 575–584. 10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

UNCORRECTED PROOF