18th workshop on Computational Models of Natural Argument (CMNA)

In conjunction with the 2018 Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB 2018)

6th April 2018

Table of Contents

Program Committee

Andrew Aberdein	Florida Institute of Technology
Michał Araszkiewicz	Jagiellonian University
Trevor Bench-Capon	University of Liverpool
Floris Bex	Utrecht University
Katarzyna Budzynska	Polish Academy of Sciences and University of Dundee
Elena Cabrio	Université Côte d'Azur, CNRS, Inria, I3S, France
Reva Freedman	Northern Illinois University
Massimiliano Giacomin	University of Brescia
Floriana Grasso	University of Liverpool
Nancy Green	University of North Carolina Greensboro
Davide Grossi	University of Groningen
Randy Harris	University of Waterloo
Helmut Horacek	Saarland University
Nir Oren	University of Aberdeen
Fabio Paglieri	ISTC-CNR Rome
Rudi Palmieri	
Saint Dizier Patrick	IRIT-CNRS
Andrea Pazienza	Università degli Studi di Bari
Paul Piwek	The Open University
Henry Prakken	Department of Information and Computing Sciences, University of Utrecht & Faculty of Law, University of Groningen
Marko Tkalcic	Free University of Bolzano Faculty of Computer Science
Serena Villata	CNRS - Laboratoire d'Informatique, Signaux et Systèmes de Sophia-Antipolis
Doug Walton	University of Windsor
Simon Wells	Edinburgh Napier University
Adam Wyner	University of Aberdeen
Tommy Yuan	University of York

On Scientific Enquiry and Computational Argumentation

Federico Cerutti[1]

Abstract. In this speculative paper we discuss how existing work in formal argumentation can support the creation of a *Regulæ Philosophandi Ratiocinator*, i.e. a machinery implementing general principles of formal science. In particular, we review two research projects in this light, one aimed at supporting intelligence analysis—CISpaces.org—and one aimed at assessing natural language interfaces to formal argumentation.

1 INTRODUCTION

Epistemology is central to western philosophy: the pessimistic cave story of Plato as well as the optimistic of *anamnesis* are examples of it. When it comes to using computer science to support epistemology we cannot avoid to look at Leibniz and his *Caluculs Ratiocinator*—e.g. [3, p. 654]—as a precursor of several approaches aimed at creating a language for representing every piece of available knowledge and then applying logical reasoning to infer new knowledge. While extremely powerful in specific contexts, those approaches are not widely adopted in scientific enquiry due to their general lack of robustness against highly uncertain and only partially observable phenomena.

This paper is in favour of a *Regulæ Philosophandi Ratiocinator* (cf. [9, p. 387]) that implements modern and widely adopted theories of epistemology. In particular, according to Popper, the advancement of scientific knowledge is based upon a process of conjecture and refutation [10], an inherently argumentative process. In Section 2 we summarise the CISpaces.org project [13] and show how existing theories of computational argumentation can already provide (limited) support for scientific enquiry in real domains. While we abstain from discussing approaches to argument mining [8], language clearly plays a role in formulation of theories. However, "although clarity is valuable in itself, exactness or precision is not [. . .]. Words are significant only as instruments for the formulation of theories, and verbal problems are tiresome: they should be avoided at all cost" [10, p. 28]. In Section 3 we summarise an experiment assessing natural language interfaces to formal argumentation [1] and criticise it.

2 SCIENTIFIC METHODS OF INTELLIGENCE ANALYSIS: CISPACES

In [13] we discuss how the process of representing reasoning lines using argumentation schemes [15] and structured argumentation techniques—in particular a simplified version of ASPIC [7]—supports, with the help of efficient algorithms for computing semantics extensions [2], the process of sensemaking, by complementing human expertise in the generation of intelligence products.

[1] Cardiff university, UK, CeruttiF@cardiff.ac.uk

CISpaces.org, based upon [13], facilitates the core phase of sensemaking within the intelligence analysis process in a declarative format. Intelligence analysis is an iterative process of foraging for information and sensemaking in which the analysis structure increases incrementally from a shoebox of information, through evidence files, to the generation and evaluation of hypotheses.

CISpaces.org therefore supports—yet not guide—the analyst in a process of conjecture—of hypotheses—and refutation, based on critical questions and other known arguments/facts through an assessment of their acceptability status. Although CISpaces.org has been developed for addressing tactical situational understanding problems—in particular answering questions associated to *who* did *what*, *when*, *where*, and *to which purpose*, hence linking causes to effects and evaluating competing hypotheses—in [5] we showed its flexibility by analysing the case of Prosecutor v. Karadžić (MICT-13-55-A) in front of the UN International Criminal Tribunal.

Let us consider here a simpler case. In [14] (now retracted), Wakefield et al. present an early report investigating the case of 12 children experiencing a loss of acquired skills, including language, where "onset of behavioural symptoms was associated, by the parents, with measles, mumps, and rubella [MMR ed.] vaccination in eight of the 12 children" (conjecture).

This paper triggered a larger study (a cohort of 537,303 children) summarised in [4] where it it is shown that "There was no association between the age at the time of vaccination, the time since vaccination, or the date of [MMR ed.] vaccination and the development of autistic disorder" (refutation).

While in principle those arguments can be formalised and handled by CISpaces.org, thus supporting in part the claim that computational argumentation can be of benefit for scientific enquiry, they also highlight the need for further studies. The class of argumentation schemes used in scientific enquiry is only partially overlapping with those analysed in [15]. For instance, the findings discussed in [4] heavily rely on results from statistics that should be further represented in form of arguments. While in some cases arguing on the basis of the results of statistical tests or probabilistic inferences without further discussions can be acceptable [16], in other cases a deeper analysis [12] might be necessary. This in general depends on the audience of the analysis, which manifests the need for communicating arguments.

3 NATURAL LANGUAGE INTERFACES TO FORMAL ARGUMENTATION

The use of graphical models is the most common approach used in the formal argumentation community to capture argumentative structures. However they require a significant levels of training and resource to be produced and consumed. Instead of training users on another (graphical) language for representing argumentative structures, we can leverage our societal model, through which we are trained in reading and writing; that is, using natural language.

It is germane to consider one specific communicative goal, namely explaining the acceptability status of arguments in a given argumentation framework. In an experiment described in [1], we investigate this communication goal with a Wizard of Oz approach—hence manually generating different pieces of texts—albeit not in a scientific enquiry context.

The experiment consists of presenting participants with texts, written in natural language, followed by a questionnaire. Texts provide natural language interfaces to the following knowledge base formalised using [11]: $\Gamma = \langle S, D \rangle$ with $S = \{s_1 :\Rightarrow s_{AAA};\ s_2 :\Rightarrow s_{BBB}\}$ and $D = \{r_1 : s_{AAA} \wedge \sim ex_{AAA} \Rightarrow c;\ r_2 : s_{BBB} \wedge \sim ex_{BBB} \Rightarrow \neg\, c;\ r_3 :\sim ex_{pref} \Rightarrow r_1 \prec r_2\}$. According to [11] Γ gives rise to the following set of arguments: $\mathcal{A}rgs = \{\mathbf{a}_1 = \langle s_1, r_1\rangle, \mathbf{a}_2 = \langle s_2, r_2\rangle, \mathbf{a}_3 = \langle r_3\rangle\}$ where $\mathbf{a}_2$ $\mathcal{A}rgs$-defeats $\mathbf{a}_1$, hence the set of justified arguments is $\{\mathbf{a}_2, \mathbf{a}_3\}$.

We generate texts for four different domains, including weather forecast (derived from an example discussed in [6]):

> The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow. Meanwhile, the forecast service of the broadcasting company BBB says that it will be cloudy tomorrow but that it will not rain. It is also well known that the forecasting service of BBB is more accurate than the one of AAA.

Participants then are asked to determine which of the following positions they think is accurate: *Tomorrow will rain* ($\mathcal{P}_{\mathbf{A}}$); *Tomorrow will not rain* ($\mathcal{P}_{\mathbf{B}}$); *I cannot conclude anything tomorrow weather* ($\mathcal{P}_{\mathbf{U}}$).

Our hypothesis is that the majority of participants would agree with BBB's statement (*position* $\mathcal{P}_{\mathbf{B}}$). While such an hypothesis is supported in general, in the case of weather forecast the actual percentages of agreement are: 5% for $\mathcal{P}_{\mathbf{A}}$; 50% for $\mathcal{P}_{\mathbf{B}}$; and 45% for $\mathcal{P}_{\mathbf{U}}$. When asked to justify their position, the majority of participants supportive of $\mathcal{P}_{\mathbf{U}}$ refers to the fact that weather forecast naturally carries high uncertainty.

An interesting question—that we hope to answer soon—is whether significantly different results could be obtained using modified textual representations, such as the following:

> Tomorrow will be cloudy but not rainy according to BBB; opposite forecast is provided by AAA, known to be untrustworthy.

Indeed, critiques to this experiment can be: (1) different formulations might very well serve the same communicative goal; (2) different audience will require different level of explanation.

Regarding the first aspect, in CISpaces.org we adopt a rather pragmatic approach. The output of its natural language generation system in the given case would be[2] the following:

> We have reasons to believe that:
> - tomorrow should be cloudy but not rainy
>
> Here the pieces of information we received:
> - The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow.
> - The forecast service of the broadcasting company BBB says that it will be cloudy tomorrow but that it will not rain.
> - It is well known that the forecasting service of BBB is more accurate than the one of AAA.

[2]Currently CISpaces.org does not handle preferences (yet) hence the hypothetical sentence.

Regarding the second aspect, it is worth noticing that the above three natural language interfaces expose neither the reasoning lines linking premises to conclusion, e.g. "The weather forecasting service of the broadcasting company AAA says that it will rain tomorrow, *therefore tomorrow should rain*;" nor the reasoning lines leading to computing the acceptability status of arguments, e.g. "*Since AAAs and BBBs conclusions are incompatible, and since BBB is more accurate than AAA, it is reasonable to tentatively conclude that tomorrow should not rain.*"

4 CONCLUSION

In this speculative paper we argue that scientific enquiry can be supported by formal argumentation, that is uniquely equipped to implement the process of conjecture and refutation discussed by Popper [10]. While we base our speculation only on anecdotal evidence, notably adaptations of CISpaces.org and of previous experiments on natural language interfaces to formal argumentation, they seem convincing enough to suggest that we can soon be equipped to build a *Regulæ Philosophandi Ratiocinator*, a machinery implementing general principles of formal science.

REFERENCES

[1] F. Cerutti, N. Tintarev, and N. Oren, 'Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation', in *ECAI 2014*, pp. 207–212, (2014).

[2] F. Cerutti, M. Vallati, and M. Giacomin, 'jArgSemSAT: An Efficient Off-the-shelf Solver for Abstract Argumentation Frameworks', in *KR 2016*, pp. 541–544, (2016).

[3] G. W. Leibniz, *Philosophical Papers and Letters*, Springer Netherlands, Dordrecht, 1976.

[4] K. M. Madsen, A. Hviid, M. Vestergaard, D. Schendel, J. Wohlfahrt, P. Thorsen, J. Olsen, and M. Melbye, 'A Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism', *New England Journal of Medicine*, **347**(19), 1477–1482, (2002).

[5] Y. McDermott Rees and F. Cerutti. Request for leave to make submissions as amicus curiae. `https://goo.gl/PKHRHM`, 2018.

[6] S. Modgil, 'An argumentation based semantics for agent reasoning', in *LADS2008*, pp. 37–53, (2008).

[7] S. Modgil and H. Prakken, 'A general account of argumentation with preferences', *Artificial Intelligence*, **195**, 361–397, (2013).

[8] M.-F. Moens, 'Argumentation mining: Where are we now, where do we want to be and how do we get there?', in *FIRE '13*, (2013).

[9] I. Newton, *Philosophiæ naturalis principia mathematica*, Apud G. & J. Innys, 3rd edn., 1726.

[10] K. R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge*, Routledge, 5th edn., 1989.

[11] H. Prakken and G. Sartor, 'Argument-based extended logic programming with defeasible priorities', *Journal of Applied Non-Classical Logics*, **7**(1-2), 25–75, (1997).

[12] S. T. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, and B. Verheij, 'A two-phase method for extracting explanatory arguments from bayesian networks', *International Journal of Approximate Reasoning*, **80**, 475 – 494, (2017).

[13] A. Toniolo, T. J. Norman, A. Etuk, F. Cerutti, R. W. Ouyang, M. Srivastava, N. Oren, T. Dropps, J. A. Allen, and P. Sullivan, 'Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis', in *AAMAS 2015*, pp. 781—-789, (2015).

[14] A. Wakefield, S. Murch, A. Anthony, J. Linnell, D. Casson, M. Malik, M. Berelowitz, A. Dhillon, M. Thomson, P. Harvey, A. Valentine, S. Davies, and J. Walker-Smith, 'Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *The Lancet*, **351**(9103), 637–641, (1998).

[15] D. Walton, C. Reed, and F. Macagno, *Argumentation schemes*, Cambridge University Press, NY, 2008.

[16] M. Williams, Z. W. Liu, A. Hunter, and F. Macbeth, 'An updated systematic review of lung chemo-radiotherapy using a new evidence aggregation method', *Lung Cancer*, **87**(3), 290 – 295, (2015).

Argumentation patterns and strategies in financial communication genres

Rudi Palmieri[i]

Ever since its origins in Ancient Rhetoric, the study of argumentation has been characterised by an explicit and strong interest in the actual uses of reasoning in contexts of public interaction, particularly politics and law. In recent years, the range of contexts investigated by the argumentation community has significantly extended to include domains like health, science and education, journalism and finance. In this talk, I will present financial communication as an inherently argumentative context and an emergent field of argumentation research. I will explain why argumentation is so relevant for finance and what benefits argumentation research can gain from studying the strategic uses of arguments in the financial context (Palmieri, 2017). Then, I will discuss some significant results obtained from the dialectical and rhetorical analysis of argumentative interactions in different financial market activities. More in particular, I will review and discuss recent studies which attempt to bring together qualitative and quantitative analytic approaches through computer-aided manual and semi-automatic annotation of corpora referring to different financial communication genres, like entrepreneurial pitches, earnings conference calls, profit warnings and financial news stories and comments.

REFERENCES

[1] Palmieri, R. (2017). The Role of Argumentation in Financial Communication and Investor Relations. In: A. Laskin (Ed.), *The Handbook of Financial Communication and Investor Relations* (chapter 4, pp. 45-60). John Wiley & Sons.

[i] Department of Communication and Media, University of Liverpool.
Rudi.Palmieri@liverpool.ac.uk

Arguments in Gradatio, Incrementum and Climax; a Climax Ontology

Cliff O'Reilly[1] and Yetian Wang[2] and Katherine Tu[3] and Sarah Bott[4] and Paulo Pacheco[5] and Tyler William Black[6] and Randy Allen Harris [7]

Abstract. Climax is a compound rhetorical figure that includes Incrementum and Gradatio; Gradatio, in turn, is a series of Anadiploses. We report on a novel suite of ontologies that describe these figures and their interconnections. With influence ranging from ancient analysts to a particular study by Jeanne Fahnestock we model the figural structure and aspects of argumentation and cognitive affinities. The key structures for the purpose of argumentation are two overlapping ordered *series* that give *support* to *claims* argued by the coalescing figures. Incrementum has a uni-directional series that exhibits a semantic increase whereas Gradatio shows an overlapping series where semantic properties are distributed less evenly. The resultant Climax *comprises* these two constituent figures and produces a complex argument structure where overlapping series generate multiple, reinforcing claims. Our ontologies are developed in the Web Ontology Language (OWL), validated for consistency and published online.

1 INTRODUCTION

The subject of this report is the rhetorical figure of **Climax**. It is a figure long in history and deep in complexity. We describe the history and background of research in this field—both linguistic and computational—and follow that with an outline of our own investigations. The resulting ontology output is available online.[8]

The importance of rhetorical figures for Argumentation generally ([1], [11], [7], [17], [16], [24], [47]) and Computer Argumentation in particular ([12], [13], [20], [27], [25]) is increasingly clear. Whenever we deal with rhetorical figures, however, we cannot overlook the fact that the traditional literature presents considerable challenges. The tradition is a long and multiplex one—multicultural, multilinguistic, multidisciplinary, and multifactorial—full of riches, but also inconsistent and occasionally even contradictory. The terminology can be especially troublesome. We focus on a small cluster of related figures we call **Climax, Gradatio, Anadiplosis, and Incrementum**, names drawn from the tradition but fixed more precisely by the Waterloo Rhetorical Figure Ontology (sketched briefly, at different stages, in [16] and [25]). Climax is the central figure in this project; the other three all occur as independent figures, but when combined they realize the figure Climax. Our research is therefore a contribution to computational argument studied and the study of rhetorical figures, especially their combinatorics. The term Climax has an ancient provenance—from the Greek κλιμαξ (meaning ladder or staircase). Aristotle may reference it, in his defining treatise *Art of Rhetoric*, discussing the "the manner of Epicharmus"—a comedic dramatist known for exaggeration. Aristotle uses the terms *combination* (συντιθεναι) and *building up* (εποικοδομειν, e.g. as in a house). "[C]ombination," he says in this context, "is an exhibition of great superiority and appears to the origin of great things" and notes that "of two things that which is nearer the end proposed is preferable" [1]. But Rhys Roberts, in his popular translation, in fact renders εποικοδομειν as Climax, and the great 19th century Classicist, Edward Meredith Cope, glosses Aristotle's comments on the style of Epicharmus here as "the building up of one phrase upon (επι) another, one rising above another step by step, like the rounds of 'a ladder' (κλιμαξ), or the stages of a building" [11]:1.142.

The term Gradatio is from the Latin for step (*gradus*). It appears in the early handbook erroneously attributed to Cicero, the *Rhetoric ad Herennium*, where Caplan translates it as Climax, and translates the definition as "the figure in which the speaker passes to the following word only after advancing by steps to the preceding one" [10], along with numerous, not fully uniform examples, such as the following:

> I did not conceive this without counselling it; I did not counsel it without myself at once undertaking it; I did not undertake it without completing it; nor did I complete it without winning approval of it. (1)

Some of the examples, though not all, include a semantic incline (a 'ladder'). The author notes that the defining characteristic of Gradatio is "the constant repetition of the preceding word," over phrase or clause breaks, adding that this repetition "carries a certain charm" [10].

The two words, then, are frequently treated as Greek/Latin synonyms. But we can see at least two processes at work, and the Waterloo Rhetorical Figure Ontology definitions isolate these processes. We use *Gradatio* for the step-wise advancement, phrase-to-phrase or clause-to-clause signaled by the repetition across phrase or clause boundaries. We use *Climax* for those cases where this movement 'rises up' as on a ladder, a classic example being:

> The industry of Africanus brought him excellence, his excellence glory, his glory rivals. (2)

Here we see not just the repetition across clause boundaries, but an 'increase,' a semantic 'rising up' with the relevant terms: excellence is surpassed by glory (note, too, that industry is surpassed by

[1] Anglia Ruskin University, UK, email: cliff.oreilly@anglia.ac.uk
[2] University of Waterloo, Canada, email: yetian.wang@uwaterloo.ca
[3] University of Waterloo, Canada, email: ktu@uwaterloo.ca
[4] University of British Columbia, Canada, email: sarah.bott@alumni.ubc.ca
[5] University of Waterloo, Canada, email: ppacheco@uwaterloo.ca
[6] University of Waterloo, Canada, email: twblack@edu.uwaterloo.ca
[7] University of Waterloo, Canada, email: raha@uwaterloo.ca
[8] http://artsresearch.uwaterloo.ca/rhetfig/climactic_ont.owl

excellence). But there are two further decompositions we need for full precision of Climax, one lexico-syntactic, the other semantic—respectively, they are the rhetorical figures Anadiplosis and Incrementum.

The term *Anadiplosis* is from the Greek (αναδιπλωσις), meaning 'doubling.' It is defined by Susenbrutos as "when the last word of a previous clause is repeated at the beginning of the following clause," giving examples such as:

> Then follows wondrously beautiful Astyr, Astyr, relying on his steed. [7]:50 (3)

The term *Incrementum* is perhaps the only self-evident one among our figural terms, since English has absorbed it into ordinary language, for quantitative or qualitative increases. The Early Modern rhetorician, Henry Peachum defines it "as is a form of speech, which by degrees ... we make our saying grow, and increase by an orderly placing of wordes making the latter word alwaies exceede the former in the force of signification," with examples like this:

> Neither silver, gold, nor precious stones might be compared to her vertues. (4)

The distinguishing characteristic of Incrementum is a series of words (three or more) from the same semantic domain, in which each subsequent word in the series increases along some metric (size, beauty, intensity, status, and so on).

As English scholar Michael Ullyot has noted, "[G]radatio's admixture with other figures and tropes makes its edge cases more difficult to detect" [6] and this phenomenon of interweaving figures was a key theme of our research from conception through to implementation; indeed, that is what makes our project so interesting. More precise definitions and further examples follow in the body of our paper, but the relation among our figures is as follows:

- Anadiplosis features repeating elements on either side of a phrase or clause boundary.
- Gradatio is a series of Anadiploses.
- Incrementum is a series of same-domain words in which each subsequent word marks an increase along some semantic scale.
- Climax is an amalgam of Gradatio and Incrementum, such that each word featured in a phrase- or clause-boundary repetition-pair marks an increase along some semantic scale.

We use the word *ontology* (another linguistic import from the Greeks) in two ways. First, simply as a means of describing elements within a domain such as those of rhetorical figures, argumentation and cognition. The second is more formal and necessitates a representation scheme and a framework of formal logic, both within a computational system. We choose the Web Ontology Language (OWL)[9] for its flexibility, freely-available tools and interoperability within the Semantic Web movement [3].

Our research goal is manifold: to elucidate in more detail the structures within the figure of **Climax** and others related to it; to record these structures in a formal way with a clear, controlled output to be used in computational research into rhetoric and natural language of higher order; and finally to delve deeper into the workings of these figures and illuminate aspects that cross over into related fields such as Natural Argumentation and Cognitive Science.

Our implementation of a suite of ontologies came after much deliberation and discussion of the ways in which the various schemes and tropes act and our methods and resultant output are discussed in subsequent sections. So far our implementation is limited to the few figures we have studied closely. Although ultimately intended for automatic figure detection, this is a current future goal and we bear in mind the experience of others in pigeon-holing the works of writers across the ages: "Authors of literary texts take license with the formal conventions of rhetorical figures; their departures from convention are (as we have argued) a hallmark of individual literary style. Our task as human readers is to judge whether the form is a sufficient and necessary condition for the function" [6].

[9] `https://www.w3.org/2001/sw/wiki/OWL`

2 BACKGROUND

2.1 Rhetorical Figures

The two most common types of rhetorical figures are tropes, which concern meaning, and schemes, which concern form. Tropes, such as Metaphor, Metonymy, or Synecdoche, are based on semantics, whereas Schemes, such as Rhyme, Alliteration, and Anadiplosis, are based on form. Our ontology involves both tropes (Incrementum), schemes (Anadiplosis, Gradatio) and combinations of the two (Climax). Being subsumed by the semantic category for matters of taxonomy, we consider Climax to be a trope. Compared to tropes like Metaphor and Metonymy, the figures in our ontology have not been as thoroughly studied. In her 1996 article [15], Fahnestock develops a place in argumentation for Incrementum and Gradatio. In their manuscript [6], Bradley and Ullyot use regular expressions to find instances of Gradatio.

Anadiplosis is the repetition of the last word or word string of one colon (a clause which is grammatically, but not logically, complete) at the beginning of the subsequent colon ("sleet" in (5)). When multiple Anadiploses occur in succession, this is known as **Gradatio**. Gradatio then is a sequence of Anadiploses. Examples (6) and (7) are instances of Gradatio.

> Snow turned to sleet, sleet to rain. [14]:124 (5)

> Out of joy strength came, strength that was fashioned to bear sorrow; sorrow brought forth joy. [2]:257–258 (6)

> Be secret then, trust not the open air, for air is breath, and breath blown words raise care. [29]:372 (7)

Incrementum—a figure of semantic increase—is often contrasted with another figure, Decrementum, a figure of semantic *decrease*. For the purposes of our ontology, the figures of Incrementum and Decrementum have been combined into Incrementum, as both figures contain a succession of words with scalar, absolute-value increase—whatever the direction of this increase, we argue, depends on one's point of view. (4) is an instance of Incrementum (with "silver", "gold", "precious stones" and "her vertues" increasing semantically); (8) arguably contains a Decrementum and then an Incrementum, but we consider them both Incrementa (with semantic increase or decrease between the pairs of objects, from "proud man" and "Lucifer" to "flowers in medowes" and "stars").

> In dispraise. Thus a proud man is called Lucifer, a drunkard a swine, an angry man mad. In praise. Thus a fair virgin is called an Angel; good musick celestial harmony; and flowers in medowes, stars. [42] (8)

The figure of **Climax** is a Gradatio with semantic increase, where the elements of the Anadiploses of the Gradatio are the same as the elements of the Incrementum. Again, when Gradatio occurs with *Decrementum*, it is often known as *Anti-climax*, but for the purposes of our ontology, we call both these of figures Climax. Examples (9) and (10) are instances of Climax (where in (9), the repeating elements are "hours", "days" and "year"; in (10), the repeating elements are "designer" and "person").

> Minutes are hours there, and the hours are days, / Each day's a year, and every year an age. [44] (9)

> Design must have had a designer. That designer must have been a person. That person is GOD. [35] (10)

The form of each of these figures is paired with a function that in turn renders each of these figures cognitive. Anadiplosis contains lexical repetition in salient positions (namely colon boundaries). Gradatio contains the same lexical repetition and positioning, as well as succession. Incrementum contains succession and semantic increase. And Climax contains all of these—lexical repetition, positioning, succession, and semantic increase.

Another important characteristic of rhetorical figures is their tendency to occur simultaneously. For example, (9) is an instance of Climax, but also contains Rhyme; (10), also a Climax, contains Polyptoton or a repetition of words with derivational changes ("design" and "designer"). Note that the very definition of Climax has Gradatio and Incrementum occurring simultaneously, and that the definition of Gradatio includes Anadiplosis.

2.2 Argumentation

Argumentation theory "is a rich, interdisciplinary area of research straddling the fields of philosophy, communication studies, linguistics, and psychology" and involves many theoretical constructions such as *Argument*, *Arguer* (or *Proponent*), *Audience*, *Rebuttal*, *Contradiction* etc. [4]. For our models we make use of two key ideas—*Claim* and *Support*. A Claim is a central point in an argument that is being assumed or conveyed by an Arguer. Supports are assumptions that attempt to convince the audience that the Claim is valid.

Despite the fact that, in the modern computational period, "rhetoricians and argumentation scholars have been very slow to catch on to the role of rhetorical figures in argumentation" [25], there are a number of studies which influence this growing field and our project within it.

A major influence on our research is Jeanne Fahnestock's work on Incrementum and Gradatio [15]. Fahnestock covers authoritatively the rhetorical background of these and related figures and then, beginning with Aristotle's *Topics*, expounds on their relation to the making of arguments. Firstly, rhetorical figures can be used to *epitomize* arguments; that is, an argument is often conveniently and memorably summed up in a phrase containing rhetorical figures. This is especially true for Incrementum, which can serve diverse argumentative functions. One key function is the *graded series*—analogous to the "dialectical tradition of arguing from the more or the less". Incrementa can be defined by the way they create a set of related elements that vary over the length of a figure. This can be done to bridge antithetical points (for example, without the use of Incrementum, there is a large conceptual gap between "minutes" and "age" in (9), "design" and "God" in (10), etc.). Most often this variation is uni-directional and, as she says of Kenneth Burke's assessment: "it invites the audience's participation in its construction or completion, a participation that amounts to a kind of identification with the formal device" [15]. A commonality of genus or category must be perceived in the arguer or audience and our ontologies wrap this complexity in the concept of *Idea*. Important also is the ordering by increase or decrease of the common quantity or attribute—which we model in Incrementum as an *Increase* property on the *Idea* entity. This is even more effective in Climax, where the repetition of Gradatio can be used to link the increasing ideas more strongly. Incrementum can also be used to argue that the value of something is greater than another. For example, in (4), "her vertues" is placed after "precious stones", suggesting that "her vertues" are the most valuable of the objects given.

Argued by Aristotle (and Fahnestock) is that the continuity of genus is present "in different degrees" and (after Piaget) predicated on "the ability to draw analogies between members of different categories". The complexity of this seemingly basic cognitive function is currently computationally intractable when we look to examples of Incrementum and Climax.

Gradatio, as compared with Incrementum, has a slightly different argumentative form. Its chief argumentative function is to chart a chain of influence (at its strongest, a chain of causation), as we can see in the example below. Rather than a figure-wide series moving from origin to end point, according to a "teleological principle", Gradatio creates an overlapping, smooth series of steps that doesn't necessarily aim to bring out an end point argument [15]. Instead it either brings together or pushes apart two ends of a conceptual spectrum by virtue of evoking in the audience a continuum or a fragmentation across the series. We can see this in this argument about the importance of large predators to the overall health of an ecosystem:

1. Large predators create carcasses of large prey.
2. Carcasses of large prey add nutrients and humus to the soil.
3. Enriched soil creates lush vegetation.
4. Lush vegetation attracts small herbivores, such as snowshoe hares.
5. Snowshoe hares attract mid-size predators, such as foxes.
6. Foxes displace smaller predators, such as weasels.
7. Displaced weasels become prey for avian predators, such as owls. [43]

We have idealized this argument somewhat, from a 1995 *New York Times* article, and mapped it more tightly to the Gradatio structure than in the original ([16]:109), to show how the Gradatio can chart the chains of influence that might be asserted in argumentation.

The Gradatio form is perhaps most familiar to argumentation scholars in context of informal logic:

A leads to *B*
B leads to *C*
C leads to *D*
D leads to . . .
. . . which leads to HELL.
We don't want to go to HELL
———
Don't take that first step *A*. [43]

The classic form of the slippery slope argument, since it is ostensibly a causal- or influence- chain argument, is the Gradatio. Of course, as we know, most slippery slopes are rarely furnished with a long articulated chain. They often go from *A* straight to HELL, as in this example:

> Once a man is permitted on his own authority to kill an innocent person directly, there is no way of stopping the advancement of that wedge. ... Once the exception has been admitted, it is too late; hence, the grave reason why no exception may be

allowed. That is why euthanasia under any circumstances must be condemned. [23][45]

Douglas Walton calls arguments like these, "compressed slippery slope" arguments ([5]:281f). More frequently, some of the steps are filled in, but we only get a single instance of each alleged causal link (meaning that the Anadiploses are left out), as in:

> Jeff! You know what happens when people take drugs! Pretty soon the caffeine won't be strong enough. Then you will take something stronger. Then, something stronger. Eventually, you will be doing cocaine. Then you'll be a crack addict! [43]

Walton schematizes this structure with what he calls a Sequential Premise, namely: "carrying out A_0 would lead to A_1, which would in turn lead to carrying out A_2, and so forth, through a sequence $A_2, \ldots, A_x, \ldots, A_y, \ldots, A_n$". Among this sequence for slippery slope arguments is a subsequence, he says, manifesting a "gray zone where x and y are indeterminate points" where a loss-of-control premise occurs, and the escalation continues to A_n, the "catastrophic outcome premise," going to HELL. These characteristics are definitional of the basic slippery slope argument ([5]:288).

Whether each step is spelled out, insinuated, or adumbrated, the schematic structure of the argument always follows the $A \rightarrow B, B \rightarrow C, C \rightarrow D$, etc. of the Gradatio. In fact, the "*A* leads to *B*" passage above, is Bradley Harris Dowden's explication of the "Jeff!" passage in his Logical Reasoning textbook.

But there is also a semantic aspect to slippery slope arguments that is not apparent in the $A \rightarrow B, B \rightarrow C, C \rightarrow D$, etc, Gradatio form alone. This semantic aspect can be seen in both the compressed euthanasia argument and the stepwise stronger-drugs argument: a scalar increase. In the euthanasia argument, the increase is presumably in a series of steps from some acceptable life-taking act (A_0) toward arbitrary and heinous acts of murder (A_n). In the stronger-drug argument, the increase is in the strength (and danger) of the drugs, from Red Bull (A_0) to crack (A_n). Dowden suggests this increase with his catastrophic end-term, HELL (A_n), but the defining semantics of a slippery slope are not just the endpoint itself but stepwise increases toward it. In short, the ideal form of the slippery slope argument, fallacious or reasonable, is a Climax.

The association of rhetorical figures and argumentation has been reported in a number of other recent works such as Yuan [49], Mehlenbacher [31], Lawrence et al. [30] and Mitrović et al. [32]. In an editorial by Harris and Di Marco [25], from the same journal, the importance is placed on "Repetition... such a fundamental aspect of neurocognition that we literally could not think without it", and included also is a description of the usage of Antimetabole in various US General Elections where repetition and symmetry take on important argumentative aspects (if you say something enough, it sticks). More generally the case is made that "since arguments are all the products of human minds engaging other human minds", which exhibit "important patterns of commonality", "Figures provide a way to see those patterns of commonality in argumentation".

As recounted by Mitrović et al., "argument schemes... can be seen as historical descendents of Aristotle's *Topics*" [32]. Fahnestock creates the same connection and states, "Arguments... require, first of all, that subjects being ordered by degree seem to belong to the same genus or category, at least in the perception of the arguer and audience" [15]. We have developed this fundamental principle in our work by the inclusion of the ontological entity *Similarity*. Our concept of comparison (a prerequisite for arguers' and audiences' ability to determine genus or category similarity) is enabled by including this ontology class with relationships to other entity classes such as *Idea* or *Token*. *Similarity* has properties that give it a type and an amount so that comparisons can be nuanced and multi-dimensional.

In their 2017 paper Lawrence, Visser and Reed conducted a pilot study into argument mining over a set of rhetorical figures (six schemes and two tropes). Their goal is to "test the established form-function pairings of the figures on quantitative empirical grounds". Pre-annotated texts are "segmented into the constitutive dialogue units and associated propositional units – in the AIF ontology". The Argument Interchange Format (AIF) is an ontology for argument-related concepts representing an argument as a set of nodes in a directed graph [9] and aims to "consolidate the work that has already been done in argumentation mark-up languages and multi-agent systems frameworks" [37]. In our ontologies we create argument-theoretic commonalities with AIF by including concepts for Claim and Support. Developing this connection is an area for future work.

An area of interest that we have not developed is Rhetorical Structure Theory (RST) [46] which incorporates a theory of "semantic organisation of text" [25], but does not focus on rhetorical figures in the sense that our work intends. Overlap between RST and argumentation has been investigated previously ([21]) and we would imagine this could be an area for further ontological research.

2.3 An Ontological Approach

Ontologies are a way of representing and organizing concepts and relationships between them. The largest previous work done in ontologies for rhetorical figures by Mitrović et al. [32] ontologically models many rhetorical figures in Serbian, including Anadiplosis (*palilogija/anadiploza*) and an Incrementum-like figure called "climax" (*amplifikacija*), but Harris et al. [26] explain that it is an intuitive but still "surprisingly novel" field of study.

As explained further by Harris et al ([26]), ontologies are ideal for representing ideas as complex as rhetorical figures, which have specific properties and are often interrelated. Anadiploses, for example, have the property of repeating strings between word boundaries. Multiple occurrences of Anadiploses create Gradatio, so we can say Gradatio *comprises* Anadiploses. When Incrementum and Gradatio occur together, we have Climax; hence, Climax *comprises* Incrementum and Gradatio. Furthermore, as we saw, rhetorical figures have a tendency to co-occur. In the future, if more figures are to be modelled, they could be more easily combined with other models using an ontological approach.

Because the Climactic Suite can be neatly described with a relatively small group of figures—Anadiplosis, Incrementum, Gradatio, and Climax—it was an ideal suite to model. Previous ontological work has been done on the "Chiastic Suite" of figures, which includes but is not limited to figures such as Antimetabole. This proved to be difficult because, compared to the Climactic Suite, Chiasmi involve more cognitive affinities, have a greater range of rhetorical functions, and involve more complex combinatorics (frequently co-occurring with other figures of repetition and parallelism, as well as with the Trope, Antithesis. Hence, the Climactic Suite offers a relatively isolated group of rhetorical figures to work with as a starting point to modelling other groups of figures in the future.

In effect, ontologies are simply descriptions of related concepts. The process of modelling a concept like a rhetorical figure itself is a large part of the ontology; with each property and relationship, we make a deliberate choice regarding the concepts we model. Furthermore, ontologies are the stepping stone to automating the detection of rhetorical figures; they can reason about the objects they represent, as we will see later in the discussion of our ontology.

We deliberately underspecify our ontology. The variation in patterns of figures in real text means that, as we attempt to specify and hone the elements and properties we find, instances of figures cease to conform to the model. We also do not describe constraints and property characteristics such as symmetry, reflexivity and cardinality. For the purposes of clarity in this report we wish to maintain a high-level view of the entities and relationships, however we recognise that further specification of these details will be necessary in order to improve the efficiency of inference in light of assumptions such as the Open World Assumption.

3 METHOD

We began with a bottom-up approach, starting with the simplest figure, Anadiplosis. Note that this is somewhat different from the bottom-up approach used to model Antimetabole in Harris et al. [26], in that we begin with the figure, whereas they began with instances of (multiple) figures. Instead of looking at the figures present within an instance, we focused on the syntactic constituents of each instance. We began with a single instance of Anadiplosis and annotated it, resulting in a list of constituents that we needed to be able to capture in our representation: words or groups of words, positions, cola, passages, equality relations, and part-of relations. From this, we modelled Anadiplosis, then Gradatio and Incrementum in parallel, before modelling the compound figure, Climax.

We refined many of these concepts, such as changing "sentence" to "clause" to "bag" to "colon"; "sentence" and "clause" were too restrictive, but "bag" appeared too vague and non-standard. We struggled with defining a word's proximity to a colon boundary, as often the words making up an Anadiplosis do not occur precisely at the ends and beginnings of cola. Furthermore, we introduced the concept of "tokens" to the model, which represent the words or groups of words repeating across boundaries, the epitomizing elements of an Anadiplosis. Since we hoped that our ontology could be used to perform automated detection in the future, it was important that we demarcate the defining characteristics of each figure, even if these characteristics may be obvious for human annotators.

We also considered a cognitive model of Anadiplosis, which uses concepts and ideas rather than words and tokens. Although we abandoned this idea because it runs counter to a rhetorician's definition of the figure, we borrowed from it the concepts of "closing" and "opening" tokens. We say that tokens can "close" a colon or "open" a colon, implying their positions relative to colon boundaries but that they may not be the only words in that position.

Modelling Incrementum and Climax came with difficulties as well. Incrementum, unlike Anadiplosis and Gradatio, is a semantic figure, and semantic figures are trickier to represent than syntactic ones. We debated over whether the epitomizing elements of an Incrementum were the semantically increasing words or the *ideas* behind those words. In the end, we decided it was indeed the ideas, since words can be replaced with synonyms but still make up an Incrementum, as one would expect in the definition of a trope. We also discussed the nuances of the semantic change, and whether this change should be represented as its own class (i.e. as class Gradient) with properties or as a relationship between ideas. It was at this stage that we also considered Incrementum and Decrementum to be the same figure. We decided on representing semantic increase as a relation, not only because it was simpler but because relations allow for properties that are true of semantic increase, such as transitivity (e.g. if C semantically increases from B, and B semantically increases from A, then C semantically increases from A). Furthermore, we modelled Climax as the intersection of Gradatio and Incrementum, when the epitomizing elements of Gradatio and those of Incrementum are the same. This makes sense because, by definition, a Climax must have an Incrementum, must have a Gradatio, and only occurs when these figures overlap.

4 ONTOLOGY

Our research brings forth a set of novel ontologies of the rhetorical figures described previously. The primary focus is on the form of constituent elements, but the main purpose of developing these ontologies is to provide a structural reference for future analyses, including computational approaches and to generate greater understanding of the subject figures by the process of ontological analysis itself.

Harris et al. [26] describe the RhetFig project that has analysed the figure Antimetabole for the composition of an OWL ontology. The authors describe the various approaches to modelling a complex linguistic structure such as Antimetabole. Their analysis calls the conceptual elements involved in rhetorical figures Cognitive Affinities such as CONTRAST, SIMILARITY, SEQUENCE, REPETITION and POSITION. Antimetabole, by their analysis, utilizes the affinities of REPETITION, SEQUENCE and CONTRAST. Research into the ontology representations of the cognitive aspects of rhetorical figures is a growing field [26] [34].

An important aspect of our models is that they describe the domain accurately for as many examples of a particular figure as possible. Evaluating that there are very few exception figures that cannot conform to our models is important. However, we recognize that the attribution of figures to real text can be subjective and so not all figures that are labeled as Anadiplosis, Gradatio, Incrementum or Climax will validate in their form and function to our model. We accept this and will take these exception cases as future work to validate that our model is accurate.

An important goal with ontology engineering in the domain of language analysis is to be able to share and re-use the work of others. By publishing ontologies and knowledge bases to the internet we hope to encourage others both inside and outside of academia to benefit from agreed definitions for shared concepts. Our ontologies are represented in OWL (Ontology Web Language) and therefore useful for integration into the Semantic Web or other computational applications that can utilise XML representations.

Evaluating our approach takes a number of forms; we consider both the output and the process. The process of analysis has brought new understanding and shone a light on new pathways of discovery yet to be followed. The ontology as an output is evaluated in section 5.2 of this paper. Future work includes using the ontologies in action for computational processing.

We now take each figure in turn to describe its elements and then summarize the combined model.

The top level entity in our models is Passage. We take this to mean any span of written words that can range in size from an entire book to a simple sentence. The allocation of a figure within a passage is a subjective operation, especially in the case of tropes which are more dependent on semantics than linguistic surface structure.

4.1 Anadiplosis

The central entities involved in the figure of Anadiplosis are the Closing-Token and Opening-Token (both sub-classes of Token) that are related by a Similarity measurement (with type and amount recorded by the of-type and of-amount properties respectively) that are located in Adjacent Cola and therefore repeat across a grammatical boundary closely proximated. The Closing Token appears in the closing section of the initial colon and the Opening Token appears in the opening section of the second colon. This naming is perhaps unintuitive, but reflects the location of the Token in line with its role across the cola. We choose to model the repeating entity as a Token rather than the more usual Word because of our assumption that Anadiplosis can act across elements that are not simply words, e.g. phrases. By abstracting the thing that repeats into a Token we can then ascribe any particular individual item in a text to that class of thing.

Anadiplosis is not characterised only by words that are identical, but can be reflected in a repetition of words in different inflections. Therefore the similarity measurement captures orthographic differences where a property is shared.

We model the concept of a boundary by including the class of Colon, an instance of which must be adjacent to and Precedes another instance of Colon. Where these two cola are next to one another and contain Tokens that are similar we can say that Anadiplosis is evoked. We introduce the concept of Idea in the Anadiplosis model. This is a vague concept aimed at capturing something subjective about the thing(s) that evokes it. We might imagine a concept space driven by the Arguer and populated by intended and unintended ideas from which the audience may or may not conceive. Each Anadiplosis Token evokes the same idea.

Our model for Anadiplosis is shown in Figure 1.

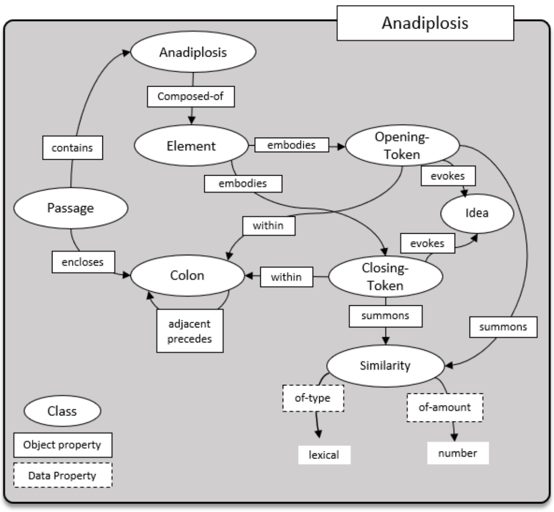


Figure 1. Our model for the scheme of Anadiplosis

4.2 Gradatio

The figure of Gradatio is in essence simply a repetition of a number of instances of Anadiplosis. Here we introduce the idea of proximity through the object property of proximal. One or more instances of the class of Anadiplosis can be proximal to another instance of the same class. This is a subjective measure (we describe problems and future work in relation to this later in this paper). A Passage contains an instance of Gradatio which comprises more than one instance of Anadiplosis that are Proximal to one another.

In this ontology we also reflect the Similarity between adjacent Anadiploses which acts in concert with the intra-similarity in each Anadiplosis which can have the effect of "spanning a conceptual gap" [15].

In the Gradatio ontology we introduce the concepts of Series-Position, Series, Claim and support. We model each instance of Anadiplosis as having a position in the series (either Initial, Mid or Final) and joining a Series which, in concert, supports a Claim.

Our model for Gradatio is shown in Figure 2.

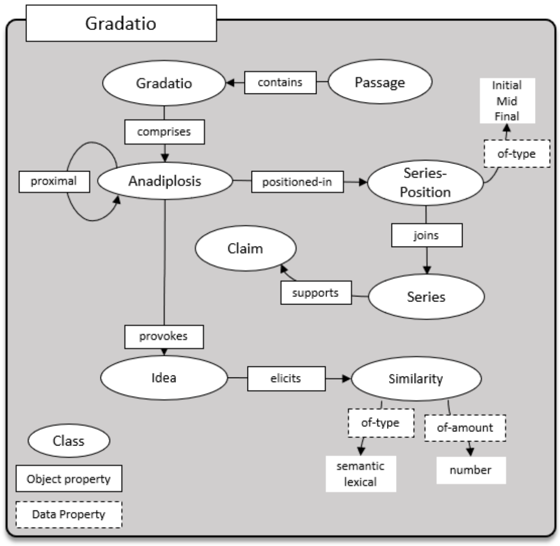


Figure 2. Our model for the trope of Gradatio

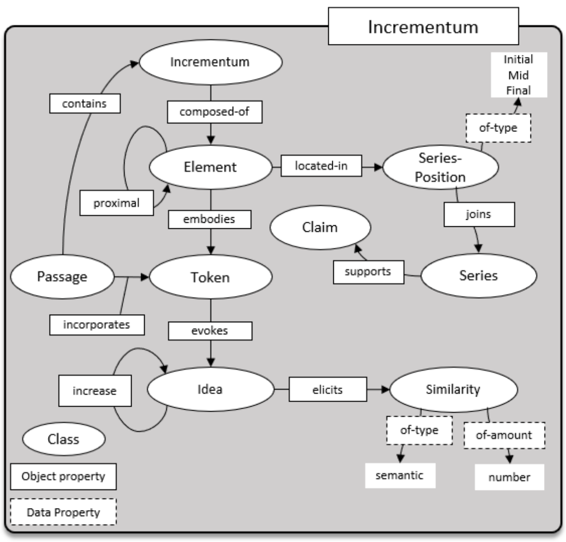


Figure 3. Our model for the trope of Incrementum

4.3 Incrementum

The Incrementum model makes use of the Idea class again. This is our method of capturing the particular facet of semantic content—the variation of which over a Passage can be said to be Incrementum (or Decrementum). We say that an instance of a Token evokes an instance of an Idea that has a similarity relationship to another instance of an Idea in the same Passage and where this exists that Incrementum exists too.

We also model the concepts of a Series in support of a Claim. In Incrementum a directed graded series is created, the elements in which are located in either Initial or Mid position and the Final position being the effective end-point all of which drive the Claim via a supports property.

Our model for Incrementum is shown in Figure 3.

4.4 Climax

Our final model is for the figure of Climax. This is modelled as a combination of the previously-described figures. We say that when Gradatio Comprises Anadiploses with Elements that are the same Elements as those with which an instance of Incrementum has a Composed-of relationship, an instance of Climax is evoked and is contained in the Passage. The argumentation aspects of Climax (beyond those already described for the sub figures) are not modeled yet as these are more complex to elucidate. It is an area of future work for us to develop this.

Our model for Climax is shown in Figure 4.

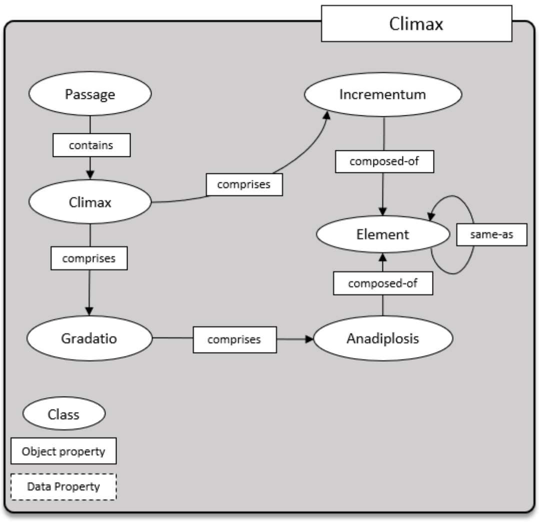


Figure 4. Our model for the trope of Climax

5 IMPLEMENTATION

The ontology was implemented in OWL using the Protégé editor [33]. The implemented OWL ontology captures the concepts denoted in the figures from previous sections using a set of axioms in terms of classes, object and datatype properties, and individuals (instances). The axioms can be represented using RDF triples in the form of (Subject, Property, Object). For example, the axiom that represents the fact "a figure is composed of an element" can be denoted as the triple `(Figure, composedOf, Element)` where **Figure** and **Element** are classes and **composedOf** is an object property from the ontology. In other words, the domain and range of the property composedOf are the classes Figure and Element respectively. For readability purpose, the first appearance of each class and property name is represented using boldface text where class names are capitalized. Individuals and RDF triples are represented using teletype text, e.g., `passage1`.

5.1 Example

We use the following passage as an example to demonstrate our ontology.

> Design must have had a designer. That designer must have been a person. That person is GOD. [35] (10)

The components in this passage are labelled as instances of classes **Passage**, **Colon** and **Token**. The RDF triple `(passage1, rdf:type, Passage)` states that the instance `passage1` is an instance of the class Passage where the property `rdf:type` indicates the instanceOf relation [8]. The selection of constructed instances and their relations are listed in *Appendix1*.

As denoted by *Appendix1*, the sample passage consists of instances representing each colon and token that appeared in the passage. A token is a word or phrase that occurs in a figure and **evokes** some idea. For example, the two appearances of "designer" are represented by the individuals `t1` and `t2` where both are instances of the class Token. That is, tokens `t1` and `t2` both have a string value "designer" and **evoke** an instance `idea1` of the class **Idea**. The similarity between the tokens is captured by the instance `sim1` of class **Similarity**. Tokens `t3` and `t4` have value "person" and both evoke `idea2` and **summons** the instance `sim2` of class Similarity. Token `t5` has the value "God" and evokes `idea3`. Furthermore, `t1` and `t3` are instances of **ClosingToken**, `t2` and `t4` are instances of **OpeningToken** which are subclasses of the class Token. A closing token appears near the end of a colon which "closes" the colon. Similarly, an opening token is placed near the beginning of the colon which "opens" the colon. Therefore the figure Anadiplosis is characterized by instances of ClosingToken and OpeningToken where the tokens refer to the same word as demonstrated by `t1` and `t2`. The repetitions created by the words "designer" and "person" created two instances of the class **Anadiplosis** in the sample passage. These instances of Anadiplosis are represented by the instances `anadiplosis1` and `anadiplosis2` which are connected to instances `ana_element1` and `ana_element2` of the class **ElementOfAnadiplosis**, which in turn **embodies** individuals of `t1` through `t4` which are instances of the class Tokens that represent the words "designer" and "person" respectively.

Gradatio is a figure consisting of two anadiploses, and is represented by the individual `gradatio1` which is an instance of the class **Gradatio** that is connected with `anadiplosis1` and `anadiplosis2` via the property comprises. The instance `anadiplosis1` and `anadiplosis2` are connected to instances of class **Series-Position** named `ana_sp1` and `ana_sp2` via the property **positioned-in**, which **joins** together to form a series represented by an instance `series1` of the class **Series** which **supports** an instance `claim1` of the class **Claim**. Each pair of Anadiploses **provoke** a common Similarity instance that has a **type** and **amount**. This similarity reflects the fractured continuation across the whole Gradatio, identified by Fahnestock [15].

An Incrementum is characterized by the semantic increase that exists among the words “designer”, “person”, and “God”. An instance of **Incrementum** named `incrementum1` was used to represent this rhetorical figure which is linked to three elements of Incrementum (i.e., `inc_element1`, `inc_element2`, and `inc_element3`) via the property composedOf. Each element embodies tokens `t1` to `t5` representing the words “designer”, “person”, and “God” where each token evokes a certain idea which is represented by the instances `idea1`, `idea2`, `idea3` respectively. As shown in *Appendix1*, `idea3` is linked to `idea2` which in turn links to `idea1` via the property **increases**. The instances of elements of Incrementum are also connected by the properties **proximal** and **precedes** which denote the proximation and ordering of the tokens embodied by the elements. Similar to the instances of class Anadiplosis discussed above, the elements (i.e., `inc_element1`, `inc_element2`, and `inc_element3`) of the instance `incrementum1` are connected to instances of class Series-Position named `inc_sp1`, `inc_sp2`, and `inc_sp3` respectively via the property **located-in**, which are connected to an instance `series2` of the class Series via property joins. The instance `series2` uses the property supports to link to an instance `claim2` of the class Claim.

The individuals discussed are representations of four rhetorical figures within the sample passage, i.e., two Anadiploses (`anadiplosis1, anadiplosis2`) which forms one Gradatio (`gradatio1`) and an Incrementum (`incrementum1`). The figures Gradatio and Incrementum form the last rhetorical figure which is a Climax represented by the individual `climax1` which is an instance of the class **Climax**. This instance is connected to `gradatio1` and `incrementum1` via the property **comprises**. A property called **same-as** that linked `inc_element1` and `ana_element1` indicates that these figure elements embody tokens with the same idea. Therefore it is the same tokens that repeat in Anadiplosis which also possess semantic increase in the Incrementa which in turn form the Climax. The instances `inc_element2` and `ana_element2` are connected in a similar manner. With all components within the sample passage represented using OWL instances, it is now possible to infer relations among each token, colon, and rhetorical figures within the passage.

5.2 Validation and Evaluation

The implemented OWL ontology and instances constructed are validated to be logically consistent by the HermiT 1.3.8 reasoner [39]. We evaluate the implemented OWL ontology and sample instances by following the methodologies for ontology evaluation discussed by Gruninger and Fox [22], where a set of competency questions were developed as requirements for the ontology. The implemented ontology must be able to represent concepts and relations within the competency questions and infer the results [19]. Our ontology was evaluated by answering the following competency questions:

1. List all cola.
2. List all tokens.
3. What are the tokens related to figure X?
4. Does figure X consist of other rhetorical figures?
5. Display the semantic increase that formed Incrementum X.

The competency questions were then translated into SPARQL queries [36] where results can be retrieved from the example instances developed in the previous section. Note that prefix namespaces are omitted to increase readability.

1. List all cola.

```
SELECT ?x WHERE { ?x rdf:type Colon }
```

Table 1. List all cola result

?x
colon1
colon2
colon3

This query simply retrieves all instances of the class Colon.

2. List all tokens.

```
SELECT ?x WHERE {
   {?x rdf:type Token} UNION
   {?x rdf:type ?c.
   ?c rdfs:subClassOf Token}}
```

Table 2. List all tokens result

?x
token1
token2
token3
token4
token5

Similarly, this query simply retrieves all instances of the class Token. This includes instances of OpeningToken and ClosingToken which are subclasses of Token.

3. What are the tokens related to figure X

```
SELECT ?y WHERE {
   ?x rdf:type ?a.
   ?a rdfs:subClassOf Figure.
   ?x composedOf ?e.
   ?e rdf:type ?b.
   ?b rdfs:subClassOf Element.
   ?e embodies ?y.
   {?y rdf:type Token.} UNION
   {?y rdf:type ?c.
   ?c rdfs:subClassOf Token.}}
```

Table 3. Tokens related to figure x result

?x	?y
anadiplosis1	token1
anadiplosis1	token2
anadiplosis2	token3
anadiplosis2	token4
incrementum1	token1
incrementum1	token2
incrementum1	token3
incrementum1	token4
incrementum1	token5

This query returns all instances that are related to an instance `?x` of the class Figure. Specifically, it returns instances of Element and Token linked by the properties composedOf and embodies respectively. We can replace the variable `?x` with a specific figure such as `anadiplosis1` in which only `token1` and `token2` are returned.

4. Does figure X consist of other rhetorical figures?

```
SELECT ?x ?y
   WHERE {
      ?x rdf:type ?a.
      ?a rdfs:subClassOf Figure.
      ?x comprises ?y.
      ?y rdf:type ?b.
      ?b rdfs:subClassOf Figure.}
```

Table 4. Does figure x consists of other figures result

?x	?y
gradatio1	anadiplosis1
gradatio1	anadiplosis2
climax1	incrementum1
climax1	gradatio1

This query returns all instances of Figure connected by the property comprises. If we replace the variable `?x` by a specific instance of Figure, e.g. `climax1`, then the query returns all instances of Figure that `climax1` comprises, i.e., `gratadio1`, `incrementum1`, `anadiplosis1`, and `anadiplosis2`.

5. Display the semantic increase that formed Incrementum X.

```
SELECT ?idea ?increasedIdea
   WHERE {
         ?x rdf:type Incrementum.
         ?x composedOf ?e.
         ?e rdf:type ElementOfIncrementum.
         ?e embodies ?y.
         ?y rdf:type ?c.
         ?c rdfs:subClassOf Token.
         ?y evokes ?idea.
         ?idea increases ?increasedIdea.}
```

Table 5. Semantic increase in Incrementum x result

?idea	?increasedIdea
idea3	idea2
idea2	idea1

An Incrementum is linked to an element which embodies some tokens that evoke an idea. This query first finds instances of Idea that are related to the instances of Token connected to the Incrementum `?x`. Then the query evaluates the property increases between the instances of Idea and returns the result. We can replace the variable `?x` with a specific instance of the class Incrementum, e.g., `incrementum1`, which yields the same result.

6 CONCLUSION

Our project has several goals. One is to elucidate structure to the understanding of Anadiplosis, Gradatio, Incrementum and Climax. We attempt this through the analysis and development of an ontology suite. We model the fine detail of each figure and consistently apply it to real-world examples of figures. This part of our project was successful, but until we extend the reach of our ontology output and include rhetoricians and users from other backgrounds we cannot be sure that either the descriptions are sensible or that the model is coherent for all the diverse variations in figures to which the descriptions are yet to be applied. The ontology is modeled and developed in OWL which enables us to utilize the power of logical inference and validation with tools such as Protégé. It also enables significant reuse within the Semantic Web movement and by publishing the files ontologies online we enable others to benefit from our work.

We hold to the idea that ontology engineering often brings benefits both in terms of the eventual output (e.g. as an XML representation to be shared and utilised), but also for the process itself of analysing a particular domain. This has been a theme of the work in analysing rhetorical schemes where significant insights have arisen from ontologically-driven knowledge engineering. This is not the only goal, however, and we aspire to take all ontologies forward into computer models that do a number of different tasks from describing, quantifying, discovering and elucidating what is a fascinating and important domain of artificial intelligence research.

The suite of ontologies contains an individual model for each figure that are combined into a single OWL file. Care was taken to name classes and properties so that no overlap would occur. Many hours were spent discussing the various aspects of the figures and issues that were raised include the idea of proximity in language and how to model this. Obviously any figure must have elements that are in some way aligned closely to each other, usually in the same sentence or group of clauses. To express this in a formal, flexible and constrained way was too difficult given the wide variety of examples in existence. We settle on the concept of Proximal which evades the issue somewhat, but captures the essence. For automated categorization of figures we are aware that this issue will need to be addressed for computational purposes and that a variation in word distance or equivalent would be suitable for defining this.

Our main influence was Fahnestock's analysis from 1996 [15] in which she discusses the argumentation aspects of Gradatio and Incrementum. Considering them as series-formative structures with variation in the direction and fragmentation of the perceived semantic properties gives us enough to build models where we merge the surface features with underlying properties of meaning (which is ultimately why we are studying rhetorical figures). Going further we relate the elements in the sequence, from their sequence positions, to the support they give to the claim that the figures are making towards an audience.

We hope that by encoding these entities in relation to these important rhetorical figures we can provide some benefits to computational rhetoric especially in the area of argumentation analysis. The burgeoning field of argument mining is another area where, because of connections we have included to some basic concepts of argument theory, some benefits can be drawn out in support of applications that highlight automatically argument schemes and supporting claims.

In parallel to the benefits to computational applications, we share our findings into the inner workings of these figures. Through our abstraction and aggregation process coupled with testing against real-world examples of each figure, we are confident that our model is accurate, however this is yet to be tested on a large set of figures.

7 FUTURE WORK

The initial goals of our project have been achieved and we are continuing with our analyses and intending to put our output to computational use in various areas such as figure detection, but we also highlight many areas on which we would like to work in the future and list them below with descriptions of the purpose and context.

1. The current state-of-the-art technologies for classification employ Machine Learning techniques, including Neural Networks, which require much data for training. One possible effective way to acquire this data is through a Gamesourcing project, which would considerably speed up populating our database, increasing the training precision on the task of automated rhetorical figure recognition by gathering the players to annotate new rhetorical figure examples upon a rewarding system offered by the Game.
2. In addition, the level of proximity (Proximal) of a Token to a Colon boundary may strengthen or weaken the presence of a rhetorical figure, such as Anadiplosis, or even invalidate its occurrence; another example to consider is the Proximal value for more than one instance of Anadiplosis, to compose a Gradatio. Therefore, several ways of measuring proximity using both true positives and negatives about the presence of a given rhetorical figure must be evaluated, ranging from word count distance to semantic distance, this latter referring to how much (maybe subjectively) the rhetoric figure effect is affected.
3. Another important metric that must be investigated is the level of semantic similarity between words, Tokens or Ideas. This metric is important to automatically detect increasing Ideas; for example, such as those that must occur in Incrementum, e.g. the words "person", "designer" and "God" invoke ideas that have semantic similarity concerning the notion of agency, but this is difficult to define precisely and an open problem in AI.
4. The Argument Interchange Format ontology is an established consolidating tool for conceptualizing argument structures. Our work creates only minimal linkage to this area via the concepts of Claim and Support, but we believe that there is room for growth to both *Information Nodes* (relating to argument content and representing claims) and *Scheme Nodes* (domain-independent patterns of reasoning) [9]. We envisage more associations being drawn out here and perhaps the development of an extension to AIF.
5. A clear goal for future work on these ontologies is to extend the conceptualizations of the argument structures so far outlined. We only model the surface features of an argument (Claims, and Support) yet there is a wealth of other elements that could be brought into the support structures behind these features. For example, the premise that Climax is a figure built up from Incrementum and Gradatio seems well established. However, the argument structures at play—a contrasting set of uni-directional graded series (Incrementum) and overlapping staggered series (Gradatio)—can be thought of as developing an even more complex combined argument that could be modelled. We also mentioned Rhetorical Structure Theory previously and another interesting area to develop would be to extend our work by including references to RST's elements similar to Mitrović et al [32]. Anadiploses may be key figures involved in Coherence Relations such as Elaboration, Circumstance, and Background, for instance, and the trope Antithesis is surely related to the relation Antithesis.
6. In many of our project meetings the go-to activity was to draw up the ontology in question on the board designed deliberately to help each of us understand what was going on in the ontology—to create a visual argument for the proposals. Visualizations in argument is an existing area of research ([5] [40] [41]), but we believe that the interplay of visual elements from the perspective of ontologies of rhetorical arguments is a novel area for research. A graded series that develops an argument applies both to words and images and is used in the example given by Fahnestock when she discusses Gradatio—George Gaylord Simpson's "Horses: The Story of the Horse Family in the Modern Worlds and through Sixty Million Years of History" [15]. Similarly the famous image of the March of Progress (from Howell [28]) shows a uni-directional graded series with a clear enough argument behind it [41]. We believe this is an interesting area to explore further.
7. Consistency-checking for data that comply with our ontology is also important in later research. This requires an automated process to determine whether instances constructed are consistent with each other. For example, if a passage contains a figure, then the tokens embodied by the figure should be within the same passage. A consistency checker was implemented by Wang and Fox [48] for city performance indicators represented using ontologies [18] where similar approach could be adopted for our Climax ontology.
8. Lastly, we want to expand the ontology analysis to include the cognitive aspects of these figures tackled on this research. When we perceive the patterns we describe as Climax, our brains do pre-determined tasks that can be put under the banners of Cognitive Affinities (like Repetition, Symmetry, Balance and Scale) and Image Schemata ([26] [34]). These are driven by neurological structures (not yet understood) that manifest as types of understanding/cognitive processing of input. Though we have touched on this a little, it has been far from comprehensive, and has left this opportunity for further inquiry. The benefits of this work would be a greater understanding of how the figures actually work "under the hood" so to speak and would increase our abilities to develop computational approaches to the management of figures for example, in text, but also to peek behind the curtain of how our brains work and thereby contribute to cognitive science in general. Rohrer [38] discusses experimental studies that connect the sensorimotor cortex to linguistic expression and Metaphor.

ACKNOWLEDGEMENTS

We would like to thank Marie Dubremetz for support with background material.

REFERENCES

[1] Aristotle, *Art of Rhetoric, trans. John Henry Freese, Loeb Classical Library*, Cambridge, MA: Harvard University Press, 1926.

[2] James Baldwin, *Go Tell It On the Mountain*, Mass Market, 2013 [1952].

[3] Tim Berners-Lee, James Hendler, Ora Lassila, et al., 'The semantic web', *Scientific American*, **284**(5), 28–37, (2001).

[4] Philippe Besnard and Anthony Hunter, *Elements of argumentation*, volume 47, MIT press Cambridge, 2008.

[5] David S Birdsell and Leo Groarke, 'Toward a theory of visual argument', *Argumentation and advocacy*, **33**(1), 1, (1996).

[6] Adam James Bradley and Michael Ullyot, *Human Creativity vs Machine Requirements: Computational Rhetoric and the Search for Gradatio*, Unpublished manuscript, n.d.

[7] Joseph Xavier Brennan, 'The epitome troporum ac schematum of joannes susenbrotus: Text, translation, and commentary', *unpublished thesis, University of Illinois*, (1953).

[8] Dan Brickley, Ramanathan V Guha, and Brian McBride, 'RDF schema 1.1', *W3C recommendation*, **25**, 2004–2014, (2014).

[9] Carlos Chesñevar, Sanjay Modgil, Iyad Rahwan, Chris Reed, Guillermo Simari, Matthew South, Gerard Vreeswijk, Steven Willmott, et al., 'Towards an argument interchange format', *The Knowledge Engineering Review*, **21**(4), 293–316, (2006).
[10] Ad Herennium Cicero, *trans. Harry Caplan, Loeb Classical Library*, Cambridge, MA: Harvard University Press, 1954.
[11] Edward Meredith Cope et al., *The Rhetoric of Aristotle: with a commentary*, volume 2, University Press, 1877.
[12] James Crosswhite, 'Rhetoric and computation', *Symposium on argument and computation, Bonskeid House, Perth-shire, Scotland.*, (2000).
[13] Jim Crosswhite, John Fox, Chris Reed, Theodore Scaltsas, and Simone Stumpf, 'Computational models of rhetorical argument', in *Argumentation Machines*, 175–209, Springer, (2003).
[14] Don DeLillo, *White Noise*, Penguin Books, 1985.
[15] Jeanne Fahnestock, 'Series reasoning in scientific argument: Incrementum and gradatio and the case of Darwin', *Rhetoric Society Quarterly*, **26**(4), 13–40, (1996).
[16] Jeanne Fahnestock, *Rhetorical Figures in Science*, Oxford University Press, New York, 1999.
[17] Jeanne Fahnestock, 'Figures of argument (OSSA 2005 keynote address)', *Informal Logic*, **24**(2), (2004).
[18] Mark S Fox, 'Polisgnosis project: Representing and analysing city indicators', in *Enterprise Integration Laboratory, University of Toronto Working paper*, (2015).
[19] Mark S Fox, 'The role of ontologies in publishing and analyzing city indicators', *Computers, Environment and Urban Systems*, **54**, 266–279, (2015).
[20] Floriana Grasso, 'Towards computational rhetoric', *Informal Logic*, **22**(3), (2002).
[21] Nancy L Green, 'Representation of argumentation in text with rhetorical structure theory', *Argumentation*, **24**(2), 181–196, (2010).
[22] Michael Grüninger and Mark S Fox, 'Methodology for the design and evaluation of ontologies', (1995).
[23] Bradley Dowden Harris, *Logical reasoning*, Wadsworth Publishing, 1993.
[24] Randy Allen Harris, 'Figural logic in Gregor Mendel's "Experiments on Plant Hybrids"', *Philosophy & Rhetoric*, **46**(4), 570–602, (2013).
[25] Randy Allen Harris and Chrysanne Di Marco, 'Rhetorical figures, arguments, computation', *Argument & Computation*, **8**(3), 211–231, (2017).
[26] Randy Allen Harris, Chrysanne Di Marco, Ashley Rose Mehlenbacher, Robert Clapperton, Insun Choi, Isabel Li, Sebastian Ruan, and Cliff O'Reilly, 'A cognitive ontology of rhetorical figures', in *Proceedings of AISB Annual Convention 2017*, pp. 228–235, (2017).
[27] Randy Allen Harris and Chrysanne DiMarco, 'Constructing a rhetorical figuration ontology', in *Persuasive Technology and Digital Behaviour Intervention Symposium*, pp. 47–52, (2009).
[28] Francis Clark Howell, 'Early man', Technical report, (1966).
[29] Thomas Kyd, *The First Part of Jeronimo*, c. 1604.
[30] John Lawrence, Jacky Visser, and Chris Reed, 'Harnessing rhetorical figures for argument mining', *Argument & Computation*, **8**(3), 289–310, (2017).
[31] Ashley Rose Mehlenbacher, 'Rhetorical figures as argument schemes–the proleptic suite', *Argument & Computation*, **8**(3), 233–252, (2017).
[32] Jelena Mitrović, Cliff O'Reilly, Miljana Mladenović, and Siegfried Handschuh, 'Ontological representations of rhetorical figures for argument mining', *Argument & Computation*, **8**(3), 267–287, (2017).
[33] Mark A Musen, 'The Protégé project: a look back and a look forward', *AI matters*, **1**(4), 4–12, (2015).
[34] Cliff O'Reilly and Randy Allen Harris, 'Antimetabole and image schemata: Ontological and vector space models', in *Proceedings of the Joint Ontology Workshops 2017 (JOWO 2017)*, eds., Stefano Borgo, Oliver Kutz, Frank Loebe, and Fabian Neuhaus.
[35] William Paley, 'Natu'. Lincoln Edmands & Company, (1833).
[36] Eric Prud, Andy Seaborne, et al., 'SPARQL query language for RDF', (2006).
[37] Iyad Rahwan and PV Sakeer, 'Towards representing and querying arguments on the semantic web', *FRONTIERS IN ARTIFICIAL INTELLIGENCE AND APPLICATIONS*, **144**, 3, (2006).
[38] Tim Rohrer, 'Image schemata in the brain', *From perception to meaning: Image schemas in cognitive linguistics*, **29**, 165–196, (2005).
[39] Rob Shearer, Boris Motik, and Ian Horrocks, 'Hermit: A highly-efficient owl reasoner.', in *OWLED*, volume 432, p. 91, (2008).
[40] Cameron Shelley, 'Rhetorical and demonstrative modes of visual argument: Looking at images of human evolution', *Argumentation and Advocacy*, **33**(2), 53, (1996).
[41] Cameron Shelley, 'Aspects of visual argument: A study of the march of progress', *Informal Logic*, **21**(2), (2001).
[42] Gent. Smith, John, *The mysterie of rhetorique unveil'd wherein above 130 the tropes and figures are severally derived from the Greek into English*, 1665.
[43] William K. Stevens, 'Wolf's howl heralds change for old haunts', *The New York Times*, (January).
[44] Sir John Suckling, *Aglaura*, c. 1637.
[45] Joseph V. Sullivan, 'The immorality of euthanasia, in beneficent euthanasia', (1975).
[46] Sandra A Thompson and William C Mann, 'Rhetorical structure theory', *IPRA Papers in Pragmatics*, **1**(1), 79–105, (1987).
[47] Christopher W Tindale, *Acts of arguing: A rhetorical model of argument*, SUNY Press, 1999.
[48] Yetian Wang and Mark S Fox, 'Consistency analysis of city indicator data', in *Planning Support Science for Smarter Urban Futures*, 355–369, Springer, (2017).
[49] Ying Yuan, 'The argumentative litotes in the analects', *Argument & Computation*, **8**(3), 253–266, (2017).

Appendix 1—Figural Instances example

Table 6. Passage Instances

Instance Name	Property	Object
passage1	rdf:type	Passage
passage1	hasValue	Design must have had a designer. That designer must have been a person. That person is God.

Table 7. Figural Instances

Instance Name	Property	Object
passage1	contains	anadiplosis1
passage1	contains	anadiplosis2
passage1	contains	gradatio1
passage1	contains	incrementum1
passage1	contains	climax1
passage1	encloses	c1
passage1	encloses	c2
passage1	encloses	c3
c1	rdf:type	Colon
c1	hasValue	Design must have had a designer.
c1	precedes	c2
c2	rdf:type	Colon
c2	hasValue	That designer must have been a person.
c2	precedes	c3
c3	rdf:type	Colon
c3	hasValue	That person is God.
t1	rdf:type	ClosingToken
t1	hasValue	designer
t1	within	c1
t1	evokes	idea1
t2	rdf:type	OpenToken
t2	within	c2
t2	hasValue	designer
t2	evokes	idea1
t3	rdf:type	ClosingToken
t3	hasValue	person
t3	within	c2
t3	evokes	idea2
t4	rdf:type	OpenToken
t4	hasValue	person
t4	within	c3
t4	evokes	idea2
idea1	rdf:type	Idea
idea2	rdf:type	Idea
sim1	rdf:type	Similarity
sim2	rdf:type	Similarity
t1	summons	sim1
t2	summons	sim1
t3	summons	sim2
t4	summons	sim2
sim1	of-type	lexical
sim1	of-amount	1.0
sim2	of-type	lexical
sim2	of-amount	1.0
anadiplosis1	rdf:type	Anadiplosis
anadiplosis1	composedOf	ana_element1
anadiplosis2	rdf:type	Anadiplosis
anadiplosis2	composedOf	ana_element2
ana_element1	rdf:type	ElementOfAnadiplosis
ana_element1	embodies	t1

Table 8. Figural Instances (cont.)

Instance Name	Property	Object
ana_element1	embodies	t2
ana_element2	rdf:type	ElementOfAnadiplosis
ana_element2	embodies	t3
ana_element2	embodies	t4
gradatio1	rdf:type	Gradatio
gradatio1	comprises	anadiplosis1
gradatio1	comprises	anadiplosis2
anadiplosis1	proximal	anadiplosis2
idea3	rdf:type	Idea
idea4	rdf:type	Idea
anadiplosis1	provokes idea3	
anadiplosis2	provokes idea3	
sim3	rdf:type	Similarity
idea3	elicits	sim3
idea4	elicits	sim3
sim3	of-type	semantic
sim3	of-amount	0.5
seriespos1	red:type	Series-Position
seriespos2	red:type	Series-Position
anadiplosis1	positioned-in	seriespos1
anadiplosis2	positioned-in	seriespos2
seriespos1	of-type	Initial
seriespos2	of-type	Final
series1	rdf:type Series	
seriespos1	joins	series1
seriespos2	joins	series1
series1	supports	claim1
claim1	rdf:type	Claim
incrementum1	rdf:type	Incrementum
incrementum1	rdf:type	Incrementum
incrementum1	composedOf	inc_element1
incrementum1	composedOf	inc_element2
incrementum1	composedOf	inc_element3
inc_element1	rdf:type	ElementOfIncrementum
inc_element1	embodies	t1
inc_element1	embodies	t2
inc_element1	proximal	inc_element2
inc_element1	precedes	inc_element2
inc_element1	same-as	ana_element1
inc_element2	rdf:type	ElementOfIncrementum
inc_element2	embodies	t3
inc_element2	embodies	t4
inc_element2	proximal	inc_element3
inc_element2	precedes	inc_element3
inc_element2	same-as	ana_element2
t5	rdf:type	Token
t5	hasValue	God
t5	within	c3
t5	evokes	idea5
inc_element3	rdf:type	ElementOfIncrementum
inc_element3	embodies	t5
sim4	rdf:type	Similarity
sim5	rdf:type	Similarity
idea5	elicits	sim4
climax1	rdf:type	Climax
climax1	comprises	gradatio1
climax1	comprises	incrementum1

From Psychological Persuasion To Abstract Argumentation: A Step Forward

Jean-Baptiste Corrégé[1] and **Emmanuel Hadoux**[2] and **Ariel Rosenfeld**[3]

Abstract. Developing argumentation-based persuasive agents that leverage human argumentative techniques is an open challenge in the computational argumentation field. In this paper, we propose a computational perspective on the psychological techniques people tend to follow during persuasion interactions drawing on psychological evidence. We focus on four well-established psychological techniques, model and investigate them using a recently proposed argumentative computational framework. Our investigation reveals both similarities and gaps between the two which can be either leveraged or addressed in the design of argumentation-based persuasive agents and future theoretical developments.

1 Introduction

A key human skill, used across many domains and activities, is the ability to *persuade*. Politicians strive to persuade their constituents, parents try to persuade their children to eat healthier food, etc. People use many different techniques for persuading others. These *human persuasive techniques* have been thoroughly investigated in the *real world* by psychology researchers. Surprisingly, despite the major advancements of the computational **argumentation theory**, providing grounded techniques and models analysed and tested in theoretical settings, the study of the possible connections between human persuasive techniques and computational models has yet to be properly examined.

In this work we provide a novel investigation of the connections between psychological persuasion literature and argumentation theory. Through this tentative investigation we are able to identify the potential use of psychological persuasive principles in argumentation-based systems and find potential directions for future work in adapting and/or extending current argumentative principles to correctly account for psychological persuasion literature. *Our findings contribute an additional stage in the greater challenge of bridging the gap between argumentation theory and people.*

The paper is structured as follows: In Section 2 we survey related works which tried to bridge the gap between psychology and argumentation, coming from both sides. We also review the *Weighted Attack/Support Argumentation graphs* [14] and the necessary definitions used in this paper. In Section 3, we discuss four well-established psychological persuasive techniques. For each technique, we present the idea underlying the technique along with supportive evidences validating the technique. Finally, in Section 4, we model the psychological persuasive techniques discussed in Section 3 using abstract argumentation and evaluate the resulting model.

[1] LIMSI, CNRS, Université Paris-Saclay, F-91405 Orsay
[2] Department of Computer Science, University College London, London, UK
[3] Weizmann Institute of Science, Rehovot, Israel

2 Background

Within the computational argumentation field, a significant effort has been placed on proposing and evaluating models and techniques aimed at allowing an automated agent (*i.e.*, *persuader*) to persuade a person (*i.e.*, *persuadee*). Theoretically, an agent would seek to deploy an *optimal* persuasive policy, mapping each possible state of a dialogue to the *best* argument for the agent to present. This persuasive policy may strive to maximize different objectives:

- likelihood of having a specific set of arguments (*i.e.*, target arguments) accepted at the end of the dialogue, [11, 3].
- persuadee's valuation of a specific point of view (represented as a single target argument) [19].
- belief of the persuadee in the target arguments [12],
- plausibility of the target arguments [14].

However, while different computational argumentative techniques have been proposed and investigated in theoretical settings, *human persuasive techniques* have been thoroughly investigated in the *real world* by psychology researchers. These studies have identified the psychological grounds and characteristics of the different techniques that people actually use. The apparent gap between the notion of persuasion in argumentation theory and human persuasive techniques prevents automated persuasive agents from building upon proven psychological persuasive evidence and thus reduces the potential impact of such agents.

A handful of previous works have examined different facets of the connections between argumentation theory and human behaviour. For example, Rahwan et al. [16] have studied the reinstatement argumentative principle in questionnaire-based experiments, Cerutti et al. [5] examined humans' ability to comprehend formal arguments and Rosenfeld and Kraus [17, 18] have established that the argumentation theory falls short in explaining people's choice of arguments in synthetic and real world argumentative settings. To the best of our knowledge, in this recent line of research, no work has used *psychological evidences* to investigate the computational argumentation theory applicability and its possible adaptation.

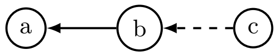


Figure 1: Example of bipolar argument graph where plain arrows mean attacks and dashed arrows mean supports.

In order to perform reasoning in a persuasive context, an argumentation framework needs to be defined (see [4] for a recent review). In its most basic form, an argumentation framework consists of a set of arguments A and an attack relation R over $A \times A$ [7]. In previous investigations of human argumentative behaviour (*e.g.*, [17, 18]), it was noticed that people often use supportive arguments rather than attacking ones, which necessitates the addition of the support relation as suggested in [1]. Furthermore, it is shown that people associate different belief levels in arguments, as suggested in [2], and different strength levels with interactions between arguments, as suggested in [8]. Interestingly, a framework named *Weighted Attacks/Support Argument* [14] embedding all these components has recently been proposed. We review this framework below.

2.1 Weighted Attacks/Support Argument

Weighted Attacks/Support Argument (WASA) graphs [14] are able to model argument graphs with attacks, supports, initial plausibility and strength of interactions between arguments taken into account. This framework merges several concepts: First of all, it is bipolar [1], allowing an additional support relation. Moreover, it uses initial weights as the plausibility for the arguments. In this work, we interpret the plausibility as an initial strength given to an argument.

Definition 1 *A WASA graph is characterized by a triplet* $\mathbb{A} = \langle \mathcal{A}, \mathcal{G}, w \rangle$*, where,*

- $\mathcal{A}$ *is a vector of size* n *ordering a set of arguments,*
- $\mathcal{G}$*, the* transposed adjacency matrix*, a square matrix of order* n*, with* $g_{ij} \in \{-1, 0, 1\}, \forall i, j \in \{1, \ldots, n\}$*, where* $g_{ij} = -1$ *(resp. 1) represents an attack (resp. a support) from* j *to* i *and 0 means no relation,*
- w *is a weight vector in* $\mathbb{R}^n$*.*

Example 1 *Example of WASA graph.*

The bipolar argument graph depicted in Figure 1 can be represented as a WASA graph as follows:

$$\left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} w_0^a \\ w_0^b \\ w_0^c \end{pmatrix} \right\rangle$$

The acceptability of an argument a is calculated using the *Direct Aggregation Semantics* and is called the *acceptability degree* $Deg_{\mathbb{A}}(a)$ in a WASA $\mathbb{A} = \langle \mathcal{A}, \mathcal{G}, w \rangle$.

In order to calculate it, we first need to define a *damping factor* $d \geq 1$, acting as a decreasing effect the further the arguments are from a in the argument graph. Then, we can calculate the propagation matrix

$$Pr^{\mathcal{G},d} = \sum_{i=0}^{\infty} \left(\frac{1}{d} \mathcal{G} \right)^i .$$

Note that the sum is defined at the infinite. However, for most applications, the sum converges to a stable propagation matrix, *i.e.*, a matrix that does not change after further propagation steps as long as $d > m$ with m the maximum indegree in $\mathcal{G}$. When the propagation does not exactly converges, we can stop the process when the matrices before and after the additional propagation step are ϵ-close. In this case the propagation matrix can be approximated by $(I - \frac{1}{d}G)^{-1}$ [14].

Finally, we can calculate the acceptability degree vector

$$Deg_{\mathbb{A},d} = Pr^{\mathcal{G},d} \times w$$

for all the arguments. We denote $Deg_{\mathbb{A},d}(a)$ the acceptability degree of argument a.

In this work, we extend the traditional argument graph depiction as presented in Figure 1 to take into account temporal aspects of the dialogue. Namely, we add additional information to the graph: the *step* at which the argument has been/to be played. This also allows us to represent duplicate arguments that may be played several times in a given dialogue. Specifically, each argument is amended with a subscript, denoting at what step it is presented. A subscript of zero denotes that an argument is not presented at all.

3 Psychosocial Persuasion Principles

We focus on four well-established techniques commonly used by professional in, for instance, sales or marketing, which have been formalized by psychologists. These technique are aimed at persuading other people. Within the field of persuasive technologies, Fogg [9] defined persuasion as "an attempt to shape, reinforce, or change behaviours, feelings, or thoughts about an issue, object, or action".

Each of the four techniques is presented along with the psychological intuition standing behind it and one or two human studies from the literature that corroborated the benefit of the technique. Following psychological terminology, we define the *target request* to be a designated argument in the argumentation framework which represents the persuader's aim or goal – namely, having the persuadee doing or believing something. This target request is equivalent to the goal argument in Rosenfeld and Kraus's framework [19]). Positing only the target argument would probably not suffice to persuade the persuadee in many setting. Therefore, it is necessary to posit additional arguments which interact with the target request or interact with other arguments that may attack the target request. Each technique prescribes a procedure of how and when to posit these additional arguments.

3.1 Foot in the Door

3.1.1 The Premise

The *foot in the door (FITD)* technique has been first described by Freedman and Fraser [10]. This technique consists in asking a small favor before asking for the target behavior (*e.g.*, asking someone for direction before asking for money). Individuals who have been asked a small request before the target one generally tend to answer more favorably compared to individuals who have straightforwardly been asked the target request. This effect is due to the fact that accepting a small, initial request leads individuals to see themselves as being social – "agreeing to requests made by strangers". Consequently, when confronted with a second request, individuals tend to comply with the above perception and accept more willingly a bigger request. Failure to conform to the self-image generated by the first request generates a cognitive dissonance, which can explain compliance.

3.1.2 Studies

In their paper, Freedman and Fraser [10] report two studies. In the first one, the target request was to ask housewives to allow a survey team to come into their homes for two hours to conduct a study about the household products they use. Participants were assigned to one of four experimental groups, depending on the first contact (*i.e.*, the initial request) before asking the target request:

1. They were asked to answer some short questions about the kinds of soaps they use (*FITD*).
2. They were asked if they would be willing to answer different questions but the questions themselves were not asked (*agree only*).
3. They were *merely approached* but not asked anything.
4. There was no initial contact (*control group*).

Results show that the compliance rate is:

1. 52.8% when the *FITD* was used,
2. 33.3% when *agree only*,
3. 27.8% when *merely approached*,
4. 22.2% for the *control group*.

In their second study, the target request was to ask participants to put a very large sign which said "Drive Carefully" in their front yard. The authors designed several types of initial requests (*e.g.*, participants were initially asked to either put a small sign in their garden or sign a petition). A control condition was added, in which participants were not initially approached. In the control condition, only 16.7% of the participants complied with the target request. The highest compliance rate was obtained by asking something that was similar (*i.e.*, put a small sign) and on a similar issue (*i.e.*, safe driving), in which case 76% of the participants agreed to the target request. In the three other configurations, 47% complied with the target request, which remains higher than in the control condition.

Theses results show that making a small initial request before a larger one brings about an increased compliance rate with the target request. This effects holds whether both requests focus on the same behaviour or not and whether both requests target the same issue or not. However, the best compliance rate is achieved when both requests target the same type of behaviour, focused on the same issue.

3.2 Door in the Face

3.2.1 The Premise

The *door in the face, (DITF)* principle has been first theorized by Cialdini et al. [6]. This technique is almost symmetrical to the *FITD* technique discussed above. Using the *DITF* technique, one asks an "unreasonable" request before proposing a smaller one – the target request. The mechanism behind it is that, after the big request have been rejected, proposing a smaller request is perceived as a concession that the persuader has made from her original request. Thus, in order to maintain a certain level of reciprocity in the relation, the persuadee will tend to comply more with the target request than if it was made without the preparatory action.

3.2.2 Studies

In their original paper [6], the authors report three studies. In one of these studies, the target request was to have students accompany a group of juvenile delinquents on a two-hour trip to the zoo. They grouped the students into three conditions:

1. students with whom they engaged the interaction by asking them first to act as counselors to juvenile delinquents for a period of two years (*big request*),
2. students without any other request except for the two-hour trip (*control group*),
3. students where *both options* were presented and subjects were to choose which (if any) of the two options to take.

The results are as follows:

1. 50% agreed when the door in the face technique was used (first asking the *big request*).
2. 16.7% of the subjects of the *control group* complied with the request.
3. 25% agreed on the small request when *both options* were presented.

3.3 Repetition

3.3.1 The Premise

The *repetition* principle has been developed and tested by Petty and Cacioppo [15]. Simply put, the technique calls for the reformulation of arguments presented multiple times.

This repetition is not endless. Although it has been show that repeating an argument two or three times, under different formulations, may be beneficial, this effect tends to decrease as the number of repetition increases.

3.3.2 Studies

Petty and Cacioppo [15] conducted two studies in which participants heard the same argument (in different formulations) zero (for control), one, three and five times in succession. They were then asked to rate their agreement with the target argument and list the arguments they could recall.

The results show that participants' agreement increases for the first three conditions and decreases when the argument is presented five times.

3.4 Anchoring

3.4.1 Principle

The *anchoring* technique has been described by Tversky and Kahneman [21]. It refers to the tendency of people to generate judgements and estimations based on an initial reference point, an *anchor*. It is thus quite simple to manipulate this anchor by providing it in the argument itself, for instance. In such case, subsequent judgement made by the persuadee are

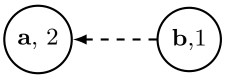


Figure 2: Foot in the door

expected to be biased toward the anchor initially provided. While this technique is more commonly used in numerical settings (*e.g.*, providing an anchor for a value of a product), it can also be applied to any type of arguments that could be ranked. In a sense, anchoring can be seen as a generalization of both the *foot in the door* and the *door in the face* principles. Indeed, in the former case, the persuader uses a smaller request first while in the latter she uses a bigger request first.

3.4.2 Studies

Tversky and Kahneman [21] provided several examples of the use of the anchoring technique. In one study, experimenters asked participants to give a series of estimations (in percentages), such as the percentage of African countries in the United Nations (UN). Before the estimation was provided, a random number between 0 and 100 was presented to participants by spinning a wheel in the participants presence. The results show that different initial number presented on the wheel led participants to generate different estimations: the group that received the number 10 estimated that 25% of African countries were in the UN (on average), whereas the group that received the number 65 estimated it at 45%.

In another study reported, experimenters asked two groups of students to estimate, within five seconds, the product of a numerical expression:

1. $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$, or,
2. $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$, *i.e.*, the exact same sequence but in reverse order.

The median estimate for the first group (with the descending sequence) was 2250, whereas it was 512 for the second group (with the ascending sequence). This result can be simply explained by the fact that subjects based their estimation on the results of the first operations they were able to make, which are obviously higher in the descending sequence.

4 From Psychosocial Persuasion Principles to Argumentation Frameworks

In this work, we will use the following interpretations of the plausibility of an argument and its acceptability degree. We consider that the plausibility w_a of an argument a corresponds to the acceptance ratio of the control conditions of the different psychological experiments. The degree of acceptability $Deg_{\mathbb{A},d}(a) \in [0,1]$ of an argument a is used as threshold. When the value is bigger (resp. smaller) than 0.5, we expect the acceptance ratio to be bigger (resp. smaller) than 50%. This is consistent with the threshold in the epistemic approach to probabilistic argumentation (see, *e.g.*, [13]).

4.1 Foot in the Door

4.1.1 Argumentation Framework:

Before the application of the *FITD* technique, the persuader wants to have a single argument **a** accepted at the end of the debate (it is the target request). Figure 2 depicts the application of the *FITD* technique on this simple graph. Note that we show the modification in an isolation context. However, in general the arguments that we want to apply the techniques on are usually part of a bigger graph.

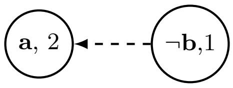


Figure 3: Door in the face.

Using the *FITD* principle, the persuader starts by playing a small argument **b** in order to have it accepted by the persuadee. She then plays argument **a** as in the original graph. However, this time, the argument **a** is supported by argument **b**. Therefore, the acceptance of **b** benefits the potential acceptance of **a**.

Example 2 *Figure 2 2 can be instantiated as follows:*

a *We come to your house to ask you questions.*
b *You answer some questions over the phone.*

The strategy of the persuader is to increase the chance of acceptance of **a** *by triggering the acceptance of* **b** *before.*

4.1.2 Analysis

The WASA associated to the single argument before the application of the *FITD* principle is trivial. Therefore we show below the procedure directly on the modified graph of Figure 2. The WASA is defined as follows:

$$\mathbb{A}' = \left\langle \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} w_a \\ w_b \end{pmatrix} \right\rangle$$

The propagation matrix after convergence, with $d = 2$ is:

$$\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$$

When calculating the degree of acceptability for both **a** and **b**, we have $Deg_{\mathbb{A},d}(a) = w_a - \frac{w_b}{2}$ and $Deg_{\mathbb{A},d}(b) = w_b$.

This means that argument **a** is accepted iff $w_a - \frac{w_b}{2} > 0.5$. Referring to the study presented in Section 3.1.2, we see that the acceptance ratio in the control condition is 22.2%. We consider $w_a = 0.222$ and that argument **a** is thus accepted iff $w_b > 0.556$. According to [10], about two third of the participants agreed with the smaller request. Therefore argument **a** should be accepted and indeed, 52.8% of the participants agreed with the target request when the *FITD* was used.

4.2 Door in the Face

4.2.1 Argumentation framework

In the same idea than previously, the initial graph is a single target argument. However, for the transformation, we consider in this case that the support of the target request is the negation of the second argument. Note that this is an abuse of notation where we really mean that it is the non-acceptance of argument **b** that reinforces argument **a**.

Example 3 *Let us instantiate the arguments of Figure 3, in a context of the zoo trip experiment, as follows:*

a *Look after juvenile delinquents for a two-hour trip.*
b *Look after them for two years.*

4.2.2 Analysis

The WASA associated with Figure 3 is:

$$\mathbb{A}' = \left\langle \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} w_a \\ 1 - w_b \end{pmatrix} \right\rangle$$

The WASA is the same as previously, except for the initial plausibility part for argument **b**.

In this case, argument **a** is accepted iff $w_a + \frac{1-w_b}{2} > 0.5$.

Following the study presented in Section 3.2.2, $w_a = 0.167$. Therefore, we need to have $1 - w_b > 0.666$ in order to have the target request accepted. In [6], the authors state that no participants accepted the bigger request prior to be presented with the second, smaller one. Following the same intuition, we consider $w_b = 0$. We can then conclude that the target request should be accepted. The psychological study suggests the same conclusion.

4.3 Repetition

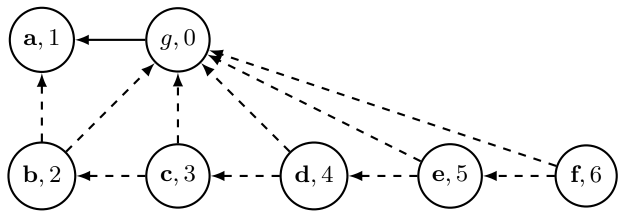


Figure 4: Repetition

Figure 4 depicts how the repetition technique can be represented. Argument a is the target request while arguments b to f are the **same** argument (in different formulations) repeated a certain number of times. As stated before, this differ from traditional abstract argumentation where arguments are represented in the graph irrespectively of the way they are used. In this case, by representing the timestep at which the argument has been played allows us to represent that an argument has been played several times in the dialogue. Argument g is a fictitious argument reinforced each time a repetition is made.

4.3.1 Analysis

The WASA associated is as follows for 5 repetitions. The WASA for the other steps can be easily deduced from this one.

$$\mathbb{A} = \left\langle \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} w_a \\ w_b \\ w_c \\ w_d \\ w_e \\ w_f \\ w_g \end{pmatrix} \right\rangle$$

If we take a damping factor of exactly one plus the maximum indegree in the graph induced by each step (*i.e.*, 2 for 0 repetition, 3 for 1 repetition, 4 for 4 repetitions and 6 for 5 repetitions), we obtain the following equations for the acceptance of argument **a** in each condition.

0 repetition $w_a - \frac{w_g}{2}$
1 repetition $w_a + \frac{2*w_b}{9} - \frac{w_g}{3}$
3 repetitions $w_a + \frac{3*w_b}{16} - \frac{w_c}{64} - \frac{17*w_d}{256} - \frac{w_g}{4}$
5 repetitions $w_a + \frac{5*w_b}{36} - \frac{w_c}{216} - \frac{37*w_d}{1296} - \frac{253*w_e}{7776} - \frac{1549*w_f}{46656} - \frac{w_g}{6}$

In this case, multiple solutions are possible. According to [15] $w_a = 0.372$. We decide to associate the following values with the different arguments: $w_b = 0.9, w_c = 0.5, w_d = 0.2, w_e = 0.05, w_f = 0.01$ and $w_g = 0$. We assume that the argument that is being repeated is an argument with a high plausibility value at the beginning but that its strength decreases as it keeps being used. As argument **g** is a fictitious one, we give it an initial value of 0.

With these values, the acceptability degree for each condition is: 0.372 for 0 repetition, 0.572 for 1 repetition, 0.520 for 3 repetitions and 0.487 for 5 repetitions. Therefore, argument **a** is accepted in the conditions 1 and 3 and rejected for 0 and 5. These results agree with the original study.

4.4 Anchoring

As a generalization of the foot in the door and the door in the face principle, the argument graph for the anchoring principle is also a generalization. However, the initial plausibilities in the WASA are no tied to the actual arguments this time but rather to their position in the ranking and the objective in the persuasion problem.

For instance, if the objective is to sell a car at the highest price possible, the first argument should be a price above the actual price and then, in a second time, the actual price. On the other hand, if the objective is to buy the very same car, it is better to give a very low price first and then converge towards the price we were willing to pay from the beginning. Therefore, in the former case, the initial value should be low for the extremely high price, increasing the more it closes the gap with the price, and then decreasing again as is goes further down, past the price and vice-versa in the latter case.

5 Discussion

In this paper we have made a modest step towards bridging the gap between abstract argumentation and psychological evidence for persuasion. This can be viewed as part of a larger effort to investigate what drives human decision-making in the argumentative context [20].

We have explained how to design argument graphs modelling four different psychological techniques, commonly used by people, and we have shown how they can be used to theoretically explain the observed results in human studies.

In future works we plan to include additional psychological techniques and a deeper analysis of the WASA and other abstract argumentation frameworks. A comparison with the traditional semantics in bipolar argumentation frameworks is another interesting direction. Finally, performing user studies based on the new representation is crucial to validate this new hybrid formalization.

Acknowledgements

This research is part funded by EPSRC Project EP/N008294/1 (Framework for Computational Persuasion).

REFERENCES

[1] Leila Amgoud, Claudette Cayrol, Marie-Christine Lagasquie-Schiex, and Pierre Livet, 'On bipolarity in argumentation frameworks', *International Journal of Intelligent Systems*, **23**(10), 1062–1093, (2008).

[2] Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza, 'Automatic evaluation of design alternatives with quantitative argumentation', *Argument & Computation*, **6**(1), 24–49, (2015).

[3] Elizabeth Black, Amanda J Coles, and Christopher Hampson, 'Planning for persuasion', in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pp. 933–942. International Foundation for Autonomous Agents and Multiagent Systems, (2017).

[4] Gerhard Brewka, Sylwia Polberg, and Stefan Woltran, 'Generalizations of dung frameworks and their role in formal argumentation', *IEEE Intelligent Systems*, **29**(1), 30–38, (2014).

[5] Federico Cerutti, Nava Tintarev, and Nir Oren, 'Formal arguments, preferences, and natural language interfaces to humans: an empirical evaluation', in *Proceedings of the Twenty-first European Conference on Artificial Intelligence*, pp. 207–212. IOS Press, (2014).

[6] Robert B Cialdini, Joyce E Vincent, Stephen K Lewis, Jose Catalan, Diane Wheeler, and Betty Lee Darby, 'Reciprocal concessions procedure for inducing compliance: The door-in-the-face technique.', *Journal of personality and Social Psychology*, **31**(2), 206, (1975).

[7] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial intelligence*, **77**(2), 321–357, (1995).

[8] Paul E Dunne, Anthony Hunter, Peter McBurney, Simon Parsons, and Michael Wooldridge, 'Weighted argument systems: Basic definitions, algorithms, and complexity results', *Artificial Intelligence*, **175**(2), 457–486, (2011).

[9] Brian J Fogg, 'Persuasive computers: perspectives and research directions', in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 225–232. ACM Press/Addison-Wesley Publishing Co., (1998).

[10] Jonathan L Freedman and Scott C Fraser, 'Compliance without pressure: the foot-in-the-door technique.', *Journal of personality and social psychology*, **4**(2), 195, (1966).

[11] Emmanuel Hadoux, Aurélie Beynier, Nicolas Maudet, Paul Weng, and Anthony Hunter, 'Optimization of probabilistic argumentation with markov decision models', in *International Joint Conference on Artificial Intelligence (IJCAI)*, (2015).

[12] Emmanuel Hadoux and Anthony Hunter, 'Strategic sequences of arguments for persuasion using decision trees', in *Proceedings of the AAAI Conference on Artificial Intelligence*. AAAI Press, (2017).

[13] A. Hunter and M. Thimm, 'Probabilistic reasoning with abstract argumentation frameworks', *Journal of Artificial Intelligence Research*, (2017). (in press).

[14] Till Mossakowski and Fabian Neuhaus, 'Bipolar weighted argumentation graphs', *arXiv preprint arXiv:1611.08572*, (2016).

[15] Richard E Petty and John T Cacioppo, 'Effects of message repetition and position on cognitive response, recall, and persuasion', *Journal of personality and Social Psychology*, **37**(1), 97–109, (1979).

[16] Iyad Rahwan, Mohammed I Madakkatel, Jean-François Bonnefon, Ruqiyabi N Awan, and Sherief Abdallah, 'Behavioral experiments for assessing the abstract argumentation semantics of reinstatement', *Cognitive Science*, **34**(8), 1483–1502, (2010).

[17] Ariel Rosenfeld and Sarit Kraus, 'Argumentation theory in the field: An empirical study of fundamental notions', in *Proceedings of the Workshop on Frontiers and Connections between Argumentation Theory and Natural Language Processing, Forlì-Cesena, Italy, July 21-25, 2014.*, (2014).

[18] Ariel Rosenfeld and Sarit Kraus, 'Providing arguments in discussions on the basis of the prediction of human argumentative behavior', *ACM Transactions on Interactive Intelligent Systems*, **6**(4), 30:1–30:33, (December 2016).

[19] Ariel Rosenfeld and Sarit Kraus, 'Strategical argumentative agent for human persuasion', in *22nd European Conference on Artificial Intelligence (ECAI)*, volume 285, p. 320. IOS Press, (2016).

[20] Ariel Rosenfeld and Sarit Kraus, *Predicting Human Decision Making: From Prediction to Intelligent Agent Design*, Morgan & Claypool, 2018.

[21] Amos Tversky and Daniel Kahneman, 'Judgment under uncertainty: Heuristics and biases', in *Utility, probability, and human decision making*, 141–162, Springer, (1975).

Policy generalisation in reinforcement learning for abstract argumentation

Sultan Alahmari[1] and **Tommy Yuan** [2] and **Daniel Kudenko**[3]

Abstract. Policy generalisation is an important attribute for an argumentative learning agent to apply the learned solutions to different environments. Learning agent needs to know the specific argumentation patterns which can help to identify optimal argument in different argumentation graphs. This paper demonstrates some difficulties in identifying patterns for learning in abstract argumentation systems. We propose to look into the internal structure of the arguments in order to facilitate the identification of useful argument patterns.

1 Introduction

Argumentation is a type of communication between agents with the purpose of reaching an agreement on what to believe [14]. There has been increasing research in agent argumentation over the past decade [18]. For an agent to be an effective dialogue participant, it needs to have a set of dialogue strategies in order to make high quality dialogue contributions. By reviewing the literature in computerised dialogue systems, such as [19] and [16], it was noted that the dialogue strategies for most implemented systems are hardwired into the agent [2]. However, given the dynamic nature of argumentation, pne problem with the hardwired strategy is that because the heuristics are fixed, it is not possible to refine or extend the dialogue strategy especially when dealing with newly arising dialogue situations. One way to address this is to make the agent search for the optimal strategies based on each situation, for instance using trial and error, the agent with the best strategy wins the dialogue [6].

Machine learning is believed to be able to meet this challenge [1] since it is flexible for an agent to learn dialogue strategies through past experiences. In addition, learning makes an agent easier to adapt to not only a deterministic environment but also a stochastic environment [1]. A common approach for machine learning in an agent context is reinforcement learning (RL) [10]. RL maps each state with an action by interacting with the environment, the agent can then learn what to do and how to connect a different situation with an action in order to maximise the cumulative reward [10]. However, the agent does not know which action to take initially so it needs to explore all actions by randomly trying them out. Figure 1 illustrates how reinforcement learning works. Some work in the literature combines reinforcement learning with dialogues, for instance [5]. Their focus, however, is on negotiation as opposed to persuasive argumentation, which is a different kind of dialogue [12]. Our chief interest is reinforcement learning for argumentation.

In [1] and [2], we present the ***ARGUMENTO+*** system, named after its predecessor *ARGUMENTO* as reported in [19]. ***ARGUMENTO+*** allows an RL agent to play an argument game against different baseline agents. The result is promising when an agent learns and plays in the same argumentation graph. It would be ideal if the knowledge learned from one argument graph could be applied to a different argument graph. This technique is called policy generalisation in the area of reinforcement learning. The key challenge here is to identify state action patterns in abstract argumentation that can be effectively applied to different argumentation graphs. [11].

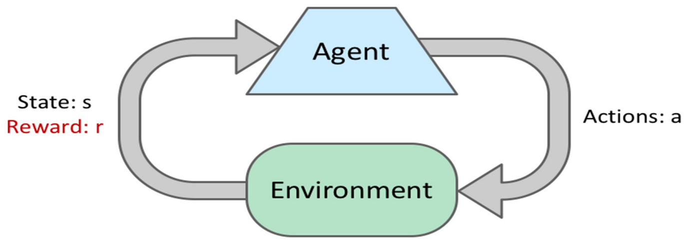


Figure 1. Reinforcement learning

The aim of this paper is to report our ongoing work in policy generalisation. The remainder of this paper is organised as follows. We first introduce the work done so far, we then propose a generalisation approach and discuss the result. Finally, we discuss our intended future work in this area.

2 Argumento+

We have built a reinforcement learning argumentation test-bed, ***ARGUMENTO+***, using the Java programming language. An Abstract Argumentation Framework [4] is used to represent the argumentation process. The argument game presented in [13] was adapted for reasons of simplicity and flexibility. The details of the argument game are as follows:

The argument game can be represented as a tuple of: G=<A, D, R, P> where: A is the argumentation system, D is the dialogue history which contains a set of moves made by the players, R is the set of rules that players need to follow when making a move, P is the set of players, normally 2 denoted as 0, 1. In [13], Wooldridge defines six rules that each participant must follow in a simple argument game and they are:

1. First move in D is made by player0 e.g. $P_0 = 0$

[1] University of York, UK, email: smsa500@york.ac.uk
[2] University of York, UK, email: tommy.yuan@york.ac.uk
[3] University of York, UK, email: daniel.kudenko@york.ac.uk

2. Players take turns making moves (one move per turn). $P_i = P_{imod2}$.
3. Players cannot repeat a move $\forall a_i, a_j \in D, a_i \neq a_j$.
4. Each move must attack (defeat) the previous move $a_i \rightarrow a_{i-1}$
5. The game is ended if no further moves are possible $\forall a_i \in A \wedge \notin D, a_i a_n$
6. The winner of the game is the player that makes the final move $G_{winner} \doteq P_{nmod2}$

In ***ARGUMENTO+***, the learning agent adopts the most commonly used reinforcement learning method, the Q-Learning algorithm. The formula for the Q-learning algorithm is: $Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma \max_a Q(S_{t+1}, A_t) - Q(S_t, A_t)]$ The aim of this algorithm is to make an agent learn from experience and map each state with an action by choosing the maximum value from the Q-table, which is updated after each episode (episode is a number of repeated game between players, each episode is one game). To enable Q-learning, we need to identify state, action and reward function.

The state representation in the literature (e.g.[14];[3]) is adapted where states are nodes in the argumentation graph and actions are the attack relation between arguments. The aim of reinforcement learning is to allow the agent to learn how to act in the environment to maximise the long term cumulative reward, and to explore the optimal actions for each state to achieve the agents goal. It is supported in [15] that learning will occur iteratively and through a trial-and-error method, depending on the experience of interaction between the agent and the environment and the reward it received.

In this research, the reward for the agent is designed as the number of acceptable arguments in the grounded extension. The reason for the adoption is that the grounded extension contains a set of acceptable argument that have been put forward by the dialogue participants, and each individual agent wishes to maximise the acceptability of their own arguments in each episode [1].

After performing an initial experiment to investigate whether the learning agent can learn to argue against the baseline agents [1]; [2], we found that it was generally encouraging to apply reinforcement learning to argumentation. However, we discovered issues with state representation, where a state is defined as the argument itself in the argumentation graph. An argument sometimes appears in different dispute lines and therefore cannot represent a unique dialogue state. As a result of the confusion over state representation, a learning agent picking an argument with a high value may sometimes lose the game. This issue has an negative impact on the agent's performance. To tackle this issue, we proposed and experimented with a more sophisticated state representation, that is (levelOfTree, agentID, currentState, previousState). The results are promising and it can clearly be seen that the learning agents perform better against different baseline agents as demonstrated in Figure 2, 3 and 4 where the performance of the learning agent is in green and the baseline agent in blue.

The agent so far learns and performs in the same argumentation graph. When facing a new argument graph, the agent has to learn from scratch. It would be ideal if the learning agent could transfer what has been learned in one graph to a different argument graph. This relates to RL policy generalisation which will be discussed next.

3 Policy generalisation for abstract argumentation

Policy generalisation intends to generalise the policy that have been learned by a RL agent. As a result, a leaned policy that has been learned from one argument graph should be able to applied to a different argument graph. One solution is to identify possible argument patterns (e.g. state-action pairs) that can be effectively applied across a range of different argument graphs.

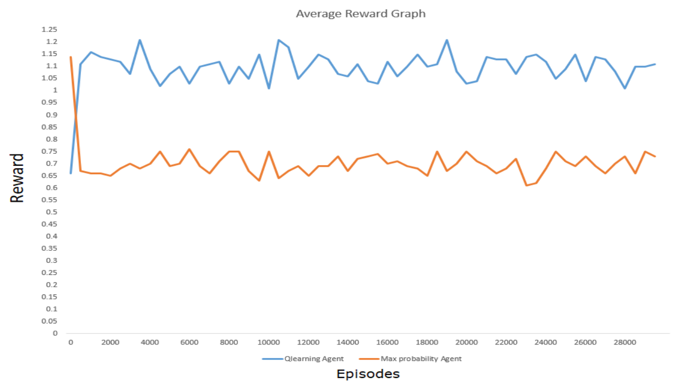


Figure 2. RL agent against Max-Probability agent

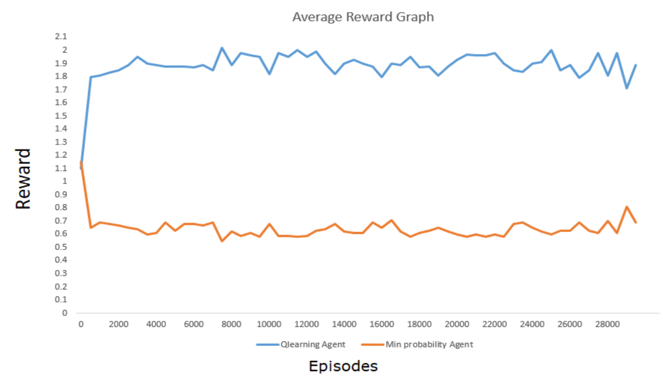


Figure 3. RL agent against Min-Probability agent

Since we are dealing with abstract argumentation system where only the arguments and the attacking relations are known, this leads us to looking into the attacking relations between the arguments which might form useful features for argument representation. We propose to take the feature of the number of attackers and the number of immediately winning attackers for use to represent an argument action. As an example shown in Figure 5, argument C, D has zero attackers, argument B has one immediately winning attacker and argument a has two attackers and one immediately winning attacker.

Number of attackers provides the number of possibilities that an argument can be attacked. Number of immediately winning attackers provides the number of immediately successful attackers. A further feature (currently named as category) can be derived by using the formula of (number of immediately winning attackers)/(number of attackers). This number provides a short term view on the proportion of the winning attackers. The value for category ranges from 0 to 1, it can therefore be further classified into different intervals {0, (0,0.25], (0.25,0.5), 0.5, (0.5,0.75], (0.75,1), 1}. The smaller the number is, the argument is likely win from a short term of view. As a result, the categories might be qualified as definite win, high likely win, likely

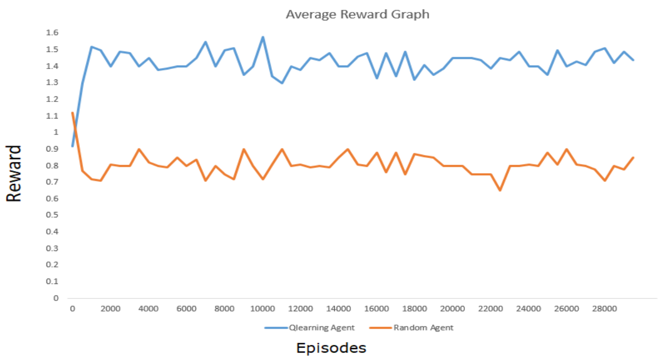


Figure 4. RL agent against Random agent

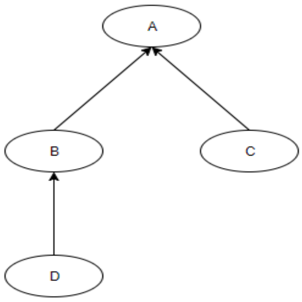


Figure 5. Argumentation graph

win, maybe win, unlikely win, high unlikely win, definitely lose from a short term view.

Number of attackers and category have been applied in representing argument actions and implemented in ***ARGUMENTO+***. To make the state representation unique for the current game, we use the following state representation: (depthOfTree, Argument, Category, NumOfAttackers). Two Q-tables have been maintained: one for the current argument game and the other is general that can be used by other argumentation graphs. After finishing each argument game we transfer the values to the general Q-table which contains only (Category, NumOfAttackers). However, if two arguments in the current game have the same category and number of attackers it will take an average of these two values then transfer that to the general Q-table.

In order to evaluate whether the generalisation method works, we needed to identify a data set in order to test the agent's performance. Initially, we tested three different graphs and found that the policy converged in episode 50 in all three, but there was a problem because the performance was not stable. We suspect that a big data set is needed in order to achieve the stability. Therefore, we randomly generated 50 different graphs, they are fully connected with a number of nodes between 5 to 10. To ensure a uniform choice, we chose Leave One Out Cross Validation in 50 graphs, and take an average at the end.

We ran the experiment over 50 games with each game having 50 graphs. The agent was trained on 49 graphs and then tested with the 50th graph. We decided to encourage our learning agent to take two different approaches based on the number of arguments in the grounded extension. We examined whether the RL agent was interested in winning the game with the minimum or maximum number of arguments. Indeed the reward shaping needs to be modified, as in the flowchart in Figure 6. The experiment also envisages to have one agent with knowledge and one without. After 50 games, an average of the rewards after every 5 episodes was taken as shown in Figures 7 and 8.

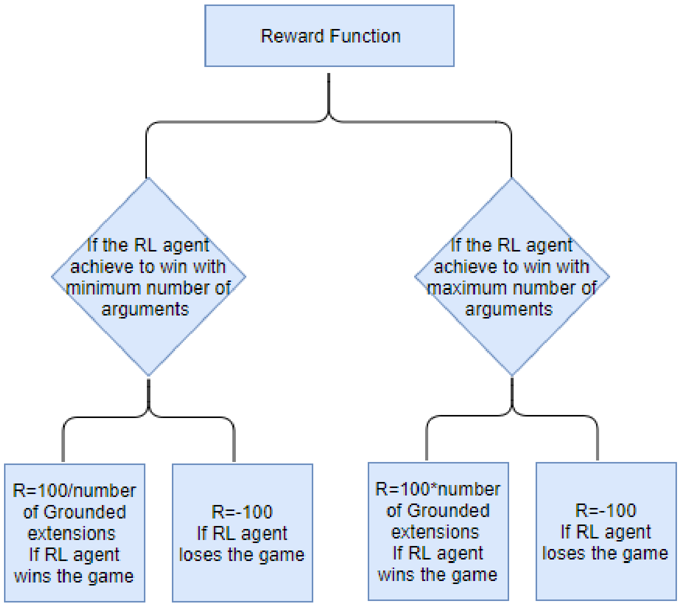


Figure 6. Reward shaping

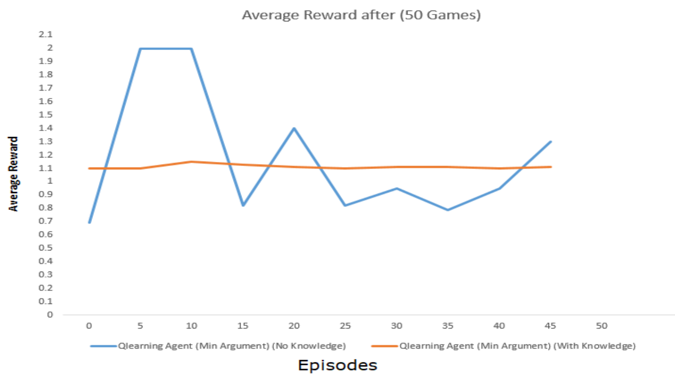


Figure 7. Cross validation for RL agent with and without knowledge with minimum numbers of arguments

In the first few episodes of both cases, the learning agent (in red line) demonstrates some advantage of the learnt knowledge but soon the advantage was overtaken by the agent learning from scratch (in blue line) most of the time. This is an unexpected result though the usefulness of the learned knowledge is merely demonstrated at the

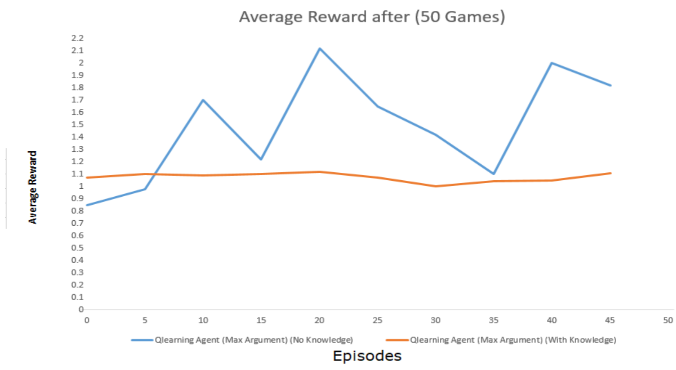


Figure 8. Cross validation for RL agent with and without knowledge with maximum numbers of arguments

start. By comparing both cases in Figures 7 and 8, the learning performs better when trying to win with minimal number of arguments. By inspecting the Q-tables, the only consistent finding is that the arguments with zero number attackers attract highest value in the winning with minimal number of argument scenario. On reflection of the experimental result, the argumentation graph in Figure 9 is used an an example to facilitate the analysis. A uniformed distribution is assumed where the winning possibility of an argument is 50/50 For example,the current state for the learning agent is argument 'A' and the agent needs to decide which argument to choose from 'B' or 'C' or 'X'. In our proposal, the agent can see the next level of the tree arguments 'D', 'E' and 'F' respectively. Therefore the chance of winning for argument 'B', 'C', 'X' are 0.25, 0.5 and 1 respectively.

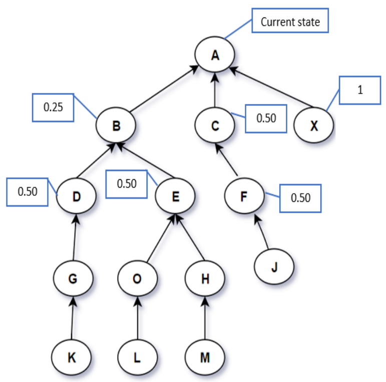


Figure 9. Argumentation graph with possibility of winning

We would normally expect the argument with a lower number of attackers to do better. This is the case for argument 'X' (with value 1) in a winning with the minimum number of arguments scenario. However when the learning agent is looking to maximise the number of grounded extensions, 'C' (with value 0.5) is the best choice. A further example can be seen from the argumentation graph in Figure 10. Although argument 'Q'(with value 0.5) has higher winning possibility value than 'R' (with value 0.125), the learning agent will choose R (with a lower value) because it is higher long term reward. We believe that this is the main reason why the learning agent cannot identify useful patterns to generalise the policy for different graphs.

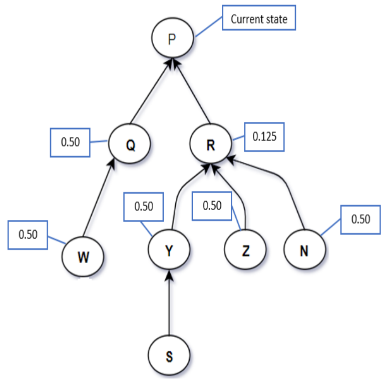


Figure 10. Different scenario of argumentation graph

4 Conclusions and Future work

We have designed an RL agent for abstract argumentation and the agent performed well in a single argument graph. We also reached the conclusion that in abstract argumentation it is hard to capture useful argument patterns that can be reused in different argument graphs. It might be sensible to move from the abstract argumentation to proposition-based argumentation where the internal structure of an argument is considered.

There are many dialogue games in the area of informal logic and computational dialectics that are operated at the propositional level (e.g. [12]; [8]; [17]). Informal logic dialogue games possess rich features in dialogue states and ample room for strategic formation where various argument patterns can be identified.

The first step of investigation could be to study the representation of goal, state, actions and reward functions for such dialogue so that reinforcement learning can effectively take place. For a persuasive dialogue the dialogue goal can be specified as converting each others' view point. Dialogue history, commitment stores and agent knowledge base contribute to the formulation of the dialogue state which should provide sufficient information for an agent to make decision for an action.

Reward functions for dialogue games are complicated to design. Ground extensions, which we have been used successfully in abstract argument games, can be applied here, with extra facilities to transform the pool of proposition-based commitments to abstract argument systems. ASPIC+ by Prakken and Modgil [7] and its implementation - TOAST [9] by Reed and Snaith will be useful here. Further reward functions can also be explored in order to capture the naturalness of a dialogue e.g. argument flows. Upon the successful learning of a single argumentation topic, the learned policies can be tested on a different topic or even a different game to see whether it is general.

ACKNOWLEDGEMENTS

We would like to thank the referees for their comments which helped improve this paper.

REFERENCES

[1] Sultan Alahmari, Tommy Yuan, and Daniel Kudenko, 'Reinforcement learning for abstract argumentation: Q-learning approach', in *Adaptive and Learning Agents workshop (at AAMAS 2017)*, (2017).

[2] Sultan Alahmari, Tommy Yuan, and Daniel Kudenko, 'Reinforcement learning for argumentation: Describing a phd research', in *Proceedings of the 17th Workshop on Computational Models of Natural Argument (CMNA17)*, (2017).

[3] Heriberto Cuayáhuitl, Simon Keizer, and Oliver Lemon, 'Strategic dialogue management via deep reinforcement learning', *arXiv preprint arXiv:1511.08099*, (2015).

[4] Phan Minh Dung, 'On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games', *Artificial intelligence*, **77**(2), 321–357, (1995).

[5] Takuya Hiraoka, Kallirroi Georgila, Elnaz Nouri, David Traum, and Satoshi Nakamura, 'Reinforcement learning in multi-party trading dialog', in *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, p. 32, (2015).

[6] Piotr S Kośmicki, 'A platform for the evaluation of automated argumentation strategies', in *International Conference on Rough Sets and Current Trends in Computing*, pp. 494–503. Springer, (2010).

[7] Sanjay Modgil and Henry Prakken, 'The aspic+ framework for structured argumentation: a tutorial', *Argument & Computation*, **5**(1), 31–62, (2014).

[8] Henry Prakken, 'Formal systems for persuasion dialogue', *The knowledge engineering review*, **21**(2), 163–188, (2006).

[9] Mark Snaith and Chris Reed, 'Toast: Online aspic+ implementation.', *COMMA*, **245**, 509–510, (2012).

[10] Richard S Sutton and Andrew G Barto, *Reinforcement learning: An introduction*, volume 1, MIT press Cambridge, 1998.

[11] Matthew E Taylor and Peter Stone, 'Transfer learning for reinforcement learning domains: A survey', *Journal of Machine Learning Research*, **10**(Jul), 1633–1685, (2009).

[12] Douglas Walton and Erik CW Krabbe. Commitment in dialogue, 1995.

[13] Michael Wooldridge, *An introduction to multiagent systems*, John Wiley & Sons, 2002.

[14] Michael Wooldridge, *An introduction to multiagent systems*, John Wiley & Sons, 2009.

[15] Erfu Yang and Dongbing Gu, 'Multiagent reinforcement learning for multi-robot systems: A survey', Technical report, tech. rep, (2004).

[16] Tangming Yuan, David Moore, and Alec Grierson, 'A human-computer dialogue system for educational debate: A computational dialectics approach', *International Journal of Artificial Intelligence in Education*, **18**(1), 3–26, (2008).

[17] Tangming Yuan, David Moore, Chris Reed, Andrew Ravenscroft, and Nicolas Maudet, 'Informal logic dialogue games in human–computer dialogue', *The Knowledge Engineering Review*, **26**(2), 159–174, (2011).

[18] Tangming Yuan, Jenny Schulze, Joseph Devereux, and Chris Reed, 'Towards an arguing agents competition: Building on argumento', in *Proceedings of IJCAI2008 Workshop on Computational Models of Natural Argument*, (2008).

[19] Tangming Yuan, Viar Svansson, David Moore, and Alec Grierson, 'A computer game for abstract argumentation', in *Proceedings of the 7th Workshop on Computational Models of Natural Argument (CMNA07)*, (2007).

ArguMessage: A System for Automation of Message Generation using Argumentation Schemes

Rosemary J. Thomas, Nir Oren and **Judith Masthoff**[1]

Abstract. This paper describes a system that uses argumentation schemes and limited user input to automatically generate persuasive messages that encourage behaviour change. We have used this system in the domain of healthy eating, but are also exploring its use in other domains such as behaviour change for cyber-security. The argumentation schemes used have been selected and amended by mapping them to Cialdini's principles [5].

1 Introduction

Individuals are increasingly recognising the importance of healthy eating and its effects on well-being. However, many find it difficult to eat healthily, leading to negative outcomes such as diabetes and obesity. Personalised messages have previously been shown to impact on positive health behaviour, and so may be used to promote healthy eating habits [8]. Researchers have investigated the personalisation of messages by adapting which of the widely used Cialdini principles of persuasion should be applied [4, 6]. The number of Cialdini principles is limited, and so the question arises as to whether the far more detailed and structured logical statements commonly used in everyday dialogue, i.e., argumentation schemes, could be used to provide finer-grained personalisation.

In our previous studies [4], we manually created and validated[2] messages for each Cialdini principle (which was extremely time consuming). Since argumentation schemes have a definite structure with easily modifiable variables, it may be easier to automate the process of message creation after the initial validation of message types. In addition, variables can be substituted with alternatives that can help in building a large corpus of messages that can be used by, for example, intelligent healthy eating trainer software. Our primary research objective is to automate personalised persuasive messages that will be able to sustain behaviour change. This could be achieved by incorporating Cialdini's principles of persuasion [1] and argumentation schemes [7, 11]. In this paper, we illustrate the system build on the basis of the mapped argumentation schemes.

2 Related work

Cialdini's Principles and Argumentation Schemes. The six principles of persuasion formulated by Cialdini [2] were Reciprocation; Commitments and Consistency; Social Proof; Liking; Authority; and Scarcity. In our previous studies [4] we decided to exclude Reciprocation and Scarcity from the follow-on studies. Only 2 Reciprocation messages validated with Kappa $\geq$ 0.4, and these were positive and negative framings of different message contents, making them hard to use for comparison in follow-on studies. Whilst 4 Scarcity messages validated with reasonable agreement (Kappa $\geq$ 0.4), none validated with Kappa $\geq$ 0.7. Additionally, both these principles are difficult[3] to use in a healthy eating persuasive context. Table 1 illustrates the four remaining Cialdini principles.

Table 1: Four Cialdini's Principles [3]

Cialdini's Principles	Description
Commitments and Consistency (COM)	"It is easier to resist at the beginning than at the end". When a person makes a dedication, he or she will experience individual and social strains to act in accordance with that initial choice.
Social Proof (SOC)	"Where all think alike, no one thinks very much". People confirm what is acceptable by knowing what others believe as acceptable.
Liking (LIK)	"The main work of a trial attorney is to make a jury like his client". We are likely to comply to requests put forward by the ones we recognise and like.
Authority (AUT)	"Follow an expert". The symbol of power linked to a person will make people adhere to their advises.

Argumentation schemes [11] are rules leading from assumptions to conclusions that are often found in everyday dialogues. Some schemes provide extremely strong support for their conclusion (such as deductive inference). However, many schemes are defeasible; if the assumptions hold, then the scheme conclusions are *probably* true, but exceptions to the conclusion do exist. This latter type of scheme is increasingly used in artificial intelligence and intelligent system applications [10].

3 Implementation

3.1 Background

The mapping of Cialdini's principles to the argumentation schemes is summarised in Table 3. We developed a message generation system using this mapping as its foundation. Given below is an explanation of one of the argumentation schemes [5].

[1] University of Aberdeen, U.K., email: [r02rj15, n.oren, j.masthoff] @abdn.ac.uk

[2] Over 150 participants classified the messages into the six principles and the Free-Marginal Kappa [9] was used to validate how effectively (1 complete agreement, 0.7-1 exceptional agreement and 0.4-0.7 reasonable agreement) our messages were classified. A message's Kappa had to be greater than 0.4 for a reasonable classification.

[3] Reciprocation is hard to apply in a system, as it requires a plausible favour and Scarcity may not be plausible in real life.

Major Premise	**Actor A** is committed to **Commitment C** according to **Goal G**.
Minor Premise	When **Actor A** is committed to **Commitment C**, it can be inferred that **Actor A** is also committed to **Action N** which contributes to **Commitment C**.
Conclusion	**Actor A** is committed to **Action N**.
Message Structure	As **Actor A** wants to achieve **Goal G**, **Actor A** is committed to **Commitment C**. So, **Actor A** is also committed to **Action N** as it helps **Actor A** achieve **Commitment C**.

Table 2: Argument from commitment with goal, and corresponding message.

Table 3: Cialdini's Principles Mapping to Argumentation Schemes [5]

Cialdini's Principles	Argumentation Schemes
Commitments and Consistency	Argument from commitment with goal Practical reasoning with goal Argument from waste with goal Argument from sunk cost with action Argument from values with goal
Social Proof	Argument from popular opinion with goal Argument from popular practice with action
Liking	Practical reasoning with liking Practical reasoning with goal and liking Argument from position to know with goal and liking
Authority	Argument from expert opinion with goal Argument from rules with goal Argument from position to know with goal Argument from memory with goal

Argument from commitment with goal. This scheme states that the proposed "action" supports the "actor" in fulfilling a "goal" they committed to previously. In the domain of healthy eating, this scheme can be used to encourage users to commit to a positive healthy eating "action" backed by their previous "commitment". The generated message is developed using a message structure created for each argumentation scheme, as demonstrated in Table 2 for the "argument from commitment with goal" argumentation scheme.

To create automated messages for the argument from commitment with goal scheme, we needed to describe a specific"commitment", "goal" and "action" for the "actor" who would be the intended subject of the message. Our aim is to crowd-source such messages, and our system therefore — as shown in Figure 1 — presents a user with a sample message using the message structure, and poses questions to instantiate the scheme's variables. In this argumentation scheme (see Figure 1), we asked three questions:

Q1. What is the goal of the user?
A. The goal of the user is to _________. This provides the input for Goal G.
Q2. What is the user therefore committed to do?
A. The user is committed to _________. This provides the input for Commitment C.
Q3. What specific action contributes to achieve this commitment?
A. The user should ________. This provides the input for Action A.

To instantiate the variables appropriately, the user's answers are required to be in a verb form. To achieve this, we provided the user with the first part of the answer (e.g., stating that "The goal of the user is to . . . " for Question 1).

The Appendix illustrates the remaining 13 argumentation schemes, and the questions for the users along with the answer structures that we have developed.

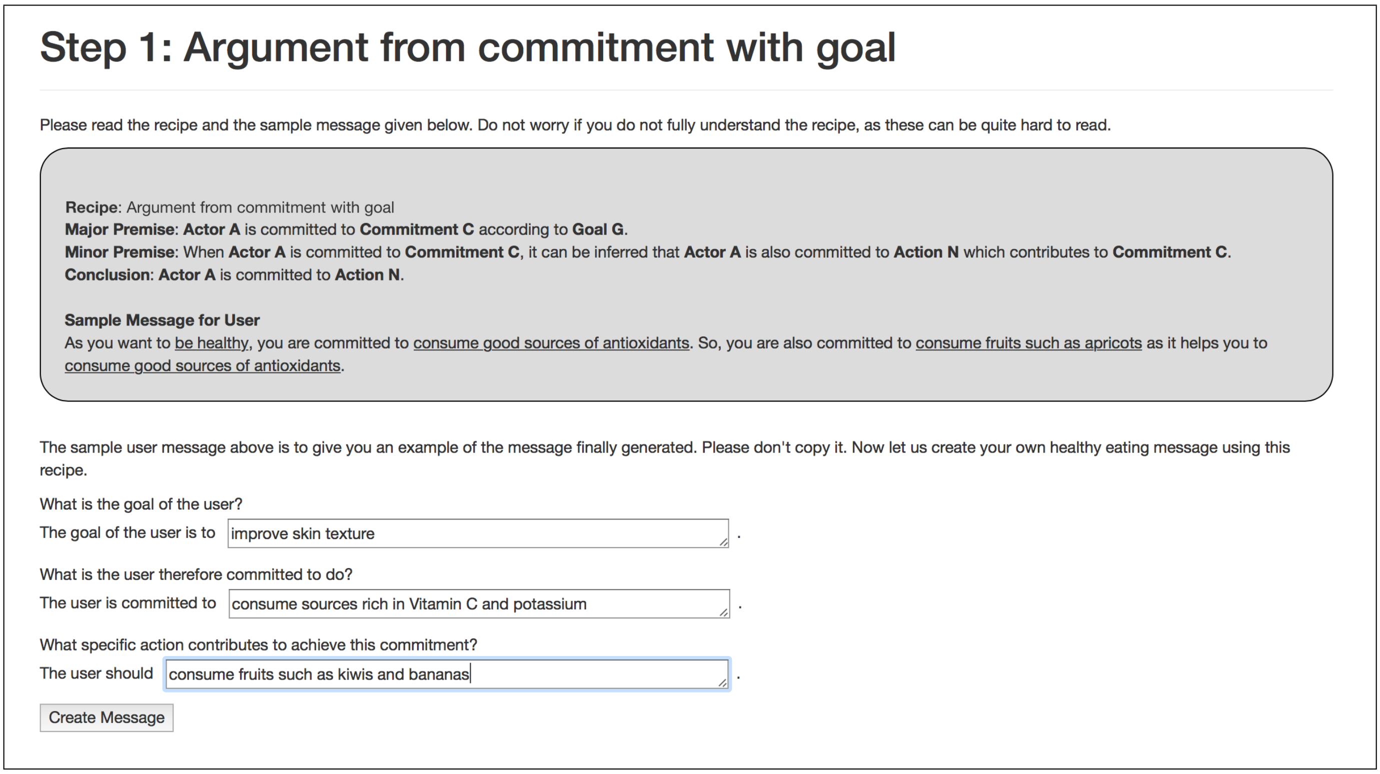
3.2 Using the system

We intend to use the system within a set of user studies. The participant is presented with the summary of the study instructions which states that they required to generate a total of three messages with three "recipes" (argumentation schemes) by answering some questions that provide the input for generating messages. Next, they are shown the explanation of a "recipe" . This is followed by a set of questions which require a small amount of participant input to generate the message. An example of the completed participant inputs is shown in Figure 1. Then, the participant presses the 'Create Message' button, which takes them to the second step which shows the generated message. In this case the message generated would be "As you want to improve skin texture, you are committed to consume sources rich in Vitamin C and potassium. So you're also committed to consume fruits such as kiwis and bananas as it helps you to consume sources rich in Vitamin C and potassium". The system uses template-based natural language generation to produce these messages. Participants provide their level of satisfaction with the message generated on a 5-point Likert scale that ranges from not satisfied to totally satisfied. In addition, they may provide detailed feedback, as input to further improve the system. When the participant presses the 'Submit' button, they are taken to the next randomly selected recipe. The same process is repeated to generate a set of three messages per participant in total.

4 Future work

We will conduct studies with lay people; argumentation scheme experts; and domain experts (e.g., dieticians) to generate a corpus of messages using the developed system, and investigate the extent to which the system makes it easy to produce good messages. We will validate the messages produced with argumentation scheme experts, to check they correspond to the argumentation schemes used to generate them. Next, the pre-validated messages will be validated as 'well-advised' or appropriate in discussions with the domain experts. Finally, we will investigate the perceived persuasiveness of these messages with respect to different types of user, to form the basis of personalized message algorithms. The latter extends the work we conducted in [4] to investigate the impact of personality on persuasiveness of messages produced from Cialdini's principles.

Whilst our initial research was focussed on the healthy eating domain, the system and the messages it generates can also be used in other domains. For example, we have started to apply it in the behaviour change for cyber-security domain [3]. The argumentation schemes used in the system are all adapted from [11]. Given Walton et al.'s schemes are mostly developed for general purposes, it is likely that domain specific argument schemes can be proposed for use by



Step 1: Argument from commitment with goal

Please read the recipe and the sample message given below. Do not worry if you do not fully understand the recipe, as these can be quite hard to read.

Recipe: Argument from commitment with goal
Major Premise: **Actor A** is committed to **Commitment C** according to **Goal G**.
Minor Premise: When **Actor A** is committed to **Commitment C**, it can be inferred that **Actor A** is also committed to **Action N** which contributes to **Commitment C**.
Conclusion: **Actor A** is committed to **Action N**.

Sample Message for User
As you want to be healthy, you are committed to consume good sources of antioxidants. So, you are also committed to consume fruits such as apricots as it helps you to consume good sources of antioxidants.

The sample user message above is to give you an example of the message finally generated. Please don't copy it. Now let us create your own healthy eating message using this recipe.

What is the goal of the user?
The goal of the user is to [improve skin texture] .

What is the user therefore committed to do?
The user is committed to [consume sources rich in Vitamin C and potassium] .

What specific action contributes to achieve this commitment?
The user should [consume fruits such as kiwis and bananas] .

Create Message

Figure 1: Explanation of argumentation scheme and questions

the proposed system. So, schemes specifically for healthy eating and cyber-security could be developed and incorporated.

The system is currently only used to generate individual persuasive messages. These messages could then be used by a dialogue system. This raises interesting questions on how to pick the best next argument.

ACKNOWLEDGEMENTS

This work is partly supported by the EPSRC award EP/P011829/1.

REFERENCES

[1] Robert B. Cialdini, 'Harnessing the science of persuasion', *Harvard Business Review*, **79**(9), 72–79, (2001).
[2] Robert B. Cialdini, *Influence: The Psychology of Persuasion*, HarperCollins e-books, 2009.
[3] Rosemary J. Thomas, Matthew Collinson, and Judith Masthoff, 'Caught by phishing emails? How can argumentation schemes be used to protect users?', in *Proceedings of AISB Annual Convention 2018*, Symposium on Digital Behaviour Intervention for Cyber Security.
[4] Rosemary J. Thomas, Judith Masthoff, and Nir Oren, 'Adapting healthy eating messages to personality', in *Persuasive Technology. 12th International Conference, PERSUASIVE 2017, Amsterdam, The Netherlands, April 4-6, 2017, Proceedings*, (2017).
[5] Rosemary J. Thomas, Nir Oren, and Judith Masthoff. How can argumentation schemes be used for persuasion? In Progress.
[6] Maurits Kaptein, Panos Markopoulos, Boris de Ruyter, and Emile Aarts, 'Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles', *International Journal of Human-Computer Stud*, **77**, 38 – 51, (2015).
[7] Irene Mazzotta and Fiorella de Rosis, 'Artifices for persuading to improve eating habits', in *AAAI Spring Symposium: Argumentation for Consumers of Healthcare*, (2006).
[8] Seth M. Noar, Christina N. Benac, and Melissa S. Harris, 'Does Tailoring Matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions', *Psychological Bulletin*, **133**(4), 673–693, (July 2007).
[9] Justus J. Randolph, 'Free-marginal multirater kappa (multirater k[free]): An alternative to fleiss fixed- marginal multirater kappa', in *Joensuu Learning and Instruction Symposium 2005*, (2005).
[10] Douglas Walton, 'Abductive, presumptive and plausible arguments', *Informal Logic*, **21**(2), 141–169, (2001).
[11] Douglas Walton, Chris Reed, and Fabrizio Macagno, *Argumentation schemes*, Cambridge University Press, 2008.

Appendix

The tables below illustrates the remaining argumentation schemes and the questions.

Table 4: Practical reasoning with goal

Major Premise	**Actor A** has **Goal G**.
Minor Premise	Carrying out **Action N** is a means to realise **Goal G**.
Conclusion	Therefore, **Actor A** ought to carry out **Action N**.
Message Structure	If **Actor A** performs **Action N**, it helps **Actor A** to achieve **Goal G**. So, **Actor A** ought to do this.

Table 5: Questions

Practical reasoning with goal
1. What is the goal of the user? A. The goal of the user is to ________. 2. What is the desired action from the user? A. The user should ________.

Table 6: Argument from waste with goal

Major Premise	If **Actor A** stops trying to realise **Goal G** now, all of **Actor A**'s previous efforts to realise **Goal G** will be wasted.
Minor Premise	If **Actor A**'s previous attempts to realise **Goal G** are wasted, that would be a bad thing.
Conclusion	Therefore, **Actor A** ought to continue trying to realise **Goal G**.
Message Structure	If **Actor A** stop trying to achieve **Goal G** now, all **Actor A**'s previous efforts will be wasted. Therefore, **Actor A** ought to continue trying to do that.

Table 7: Questions

Argument from waste with goal
1. What is the goal of the user? A. The goal of the user is to ________.

Table 8: Argument from sunk cost with action

Time T1	Time of **Actor A**'s commitment to **Action N**.
Time T2	Time of **Actor A**'s confrontation with the decision whether carry out **Action N** or not.
Major Premise	There is a choice at **Time T2** between **Action N** and not-**Action N**.
Minor Premise	At **Time T2**, **Actor A** is pre-committed to **Action N** because of what **Actor A** did or committed to at **Time T1**.
Conclusion	Therefore, **Actor A** should choose **Action N**.
Message Structure	**Actor A** has a choice whether or not to perform **Action N**, however **Actor A** was committed to do so earlier. So, **Actor A** should choose to **Action N**.

Table 9: Questions

Argument from sunk cost with action
1. What did the user commit to do? A. The user is committed to ________.

Table 10: Argumentation from values with goal

Major Premise	**Value V** is positive as judged by **Actor A**.
Minor Premise	The fact that **Value V** is positive affects the interpretation and therefore the evaluation of **Goal G** of **Actor A**.
Conclusion	**Value V** is a reason for **Actor A** retaining commitment to **Goal G**.
Message Structure	If **Actor A** achieves **Goal G**, it will help **Actor A** to realise **Value V**, which is regarded as positive by **Actor A**. This justifies that **Actor A** should achieve **Goal G**. Therefore, **Actor A** should retain **Actor A**'s commitment to it.

Table 11: Questions

Argument from values with goal
1. What does the user regard as important in their life? A. The user regards to ________ as important in their life. 2. What is the goal of the user that is related to the above? A. The goal of the user that is related to the above is to ________.

Table 12: Argument from popular opinion with goal

Major Premise	**Actor A** has **Goal G**. **Action N** is generally accepted as contributing to **Goal G**.
Minor Premise	If **Action N** is generally accepted as contributing to **Goal G**, that gives a reason for **Actor A** to do **Action N**.
Conclusion	There is a reason for **Actor A** to do **Action N**.
Message Structure	It is generally agreed that if **Actor A** performs **Action N**, this will help **Actor A** to achieve **Goal G**. So, **Actor A** should perform **Action N**.

Table 13: Questions

Argument from popular opinion with goal
1. What is the goal of the user? A. The goal of the user is to ________. 2. What is the action taken by the user achieve their goal? A. The user should ________.

Table 14: Argument from popular practice with action

Major Premise	**Action N** is a popular practice among **Actor B**.
Minor Premise	If **Action N** is a popular practice among **Actor B**, that gives a reason for **Actor A** to think that **Action N** is acceptable.
Conclusion	Therefore, there is a reason for **Actor A** to accept **Action N**.
Message Structure	**Actor B** performs **Action N**. **Actor A** should therefore do likewise.

Table 15: Questions

Argument from popular practice with action
1. What is a popular good practice? A. A popular good practice is to ________.

Table 16: Practical reasoning with liking

Major Premise	**Actor B** will appreciate it if **Actor A** carries out **Action N**.
Minor Premise	Carrying out **Action N** is a means to realise **Actor A**'s affinity towards **Actor B**.
Conclusion	Therefore, **Actor A** ought to carry out **Action N**.
Message Structure	**Actor A**'s **Actor B** will appreciate it if **Actor A** performs **Action N**. So, **Actor A** ought to do that.

Table 17: Questions

Practical reasoning with liking
1. Who does the user like? A. The user likes their ________. 2. What action should the user undertake to gain appreciation from that person? A. The user should ________.

Table 18: Practical reasoning with goal and liking

Major Premise	**Actor A** has **Goal G**. **Actor B** will appreciate it if **Actor A** realises **Goal G**.
Minor Premise	Carrying out **Action N** is a means to realise **Goal G** and **Actor A**'s affinity towards **Actor B**.
Conclusion	Therefore, **Actor A** ought to carry out **Action N**.
Message Structure	If **Actor A** performs **Action N** it helps **Actor A** to achieve **Goal G** and **Actor A**'s **Actor B** will appreciate it. So, **Actor A** ought to do that.

Table 19: Questions

Practical reasoning with goal and liking
1. What is the goal of the user? A. The goal of the user is to ________. 2. What is the desired action from the user to help achieve their goal? A. The user should ________. 3. Who does the user like? A. The user likes their ________.

Table 20: Argument from position to know with goal and liking

Major Premise	**Actor A** has **Goal G**. **Source S** is in position to know about things in a certain **Domain D** containing **Action N** which contributes to **Goal G**.
Minor Premise	**Source S** asserts that **Action N** will attain **Goal G**.
Conclusion	There is a reason for **Actor A** to do **Action N**.
Message Structure	**Actor A**'s **Source S** suggests that **Actor A** performs **Action N** to achieve **Goal G**. So **Actor A** should follow **Source S**'s suggestion.

Table 21: Questions

Argument from position to know with goal and liking
1. What is the goal of the user? A. The goal of the user is to ________. 2. Who is the experienced person liked by the user to help achieve their goal? A. The experienced person is their ________. 3. What do they recommend? A. The user should ________.

Table 22: Argument from expert opinion with goal

Major Premise	**Actor A** has **Goal G**. **Source S** is an expert in **Domain D** containing **Action N** which contributes to **Goal G**.
Minor Premise	**Source S** asserts that **Action N** will attain **Goal G**.
Conclusion	There is a reason for **Actor A** to do **Action N**.
Message Structure	**Source S** recommends that **Actor A** performs **Action N** to achieve **Goal G**. So **Actor A** should follow **Source S**'s recommendation.

Table 23: Questions

Argument from expert opinion with goal
1. What is goal of the user? A. The goal of the user is to ________. 2. Who is the professional with expertise in this field? A. The professional is a ________. 3. What do they recommend? A. The user should ________.

Table 24: Argument from rules with goal

Major Premise	**Actor A** has **Goal G**. If carrying out types of actions including **Action N** is the established rule for helping to achieve **Goal G**, then, A must carry out **Action N**.
Minor Premise	Carrying out types of actions including **Action N** is the established rule for helping to achieve **Goal G**.
Conclusion	**Actor A** must carry out **Action N**.
Message Structure	**Actor A** should perform **Action N** since it is an established rule that helps to achieve **Goal G**.

Table 25: Questions

Argument from rules with goal
1. What is the goal of the user? A. The goal of the user is to ________. 2. What action according to an established rule helps to achieve the goal of the user? A. The user should ________.

Table 26: Argument from position to know with goal

Major Premise	**Actor A** has **Goal G**. **Source S** is in position to know about things in a certain **Domain D** containing **Action N** which contributes to **Goal G**.
Minor Premise	**Source S** asserts that **Action N** will attain **Goal G**.
Conclusion	There is a reason for **Actor A** to do **Action N**.
Message Structure	**Source S** suggests that **Actor A** performs **Action N** to achieve **Goal G**. So **Actor A** should follow **Source S**'s suggestion.

Table 27: Questions

Argument from position to know with goal
1. What is goal of the user? A. The goal of the user is to ________. 2. Who has personal experience to help the user achieve their goal? A. The experienced person is a ________. 3. What do they recommend? A. The user should ________.

Table 28: Argument from memory with goal

Major Premise	**Actor B** recalls **Action N** contributed to **Goal G**.
Minor Premise	Recalling that **Action N** that contributed to **Goal G** is a clear reason for **Actor A** to believe **Action N** is good.
Conclusion	It is reasonable for **Actor A** to believe **Action N** is good.
Message Structure	**Actor A**'s **Actor B** recalls that **Action N** helped **Actor B** to achieve **Goal G**. So, **Actor A** should believe that **Action N** is good.

Table 29: Questions

Argument from memory with goal
1. Who does the user know? A. The user knows their ________. 2. How did they achieve that goal? A. They achieved that goal by ________. 3. What goal was achieved by that person? A. The goal achieved by that person was ________.

Evaluating the Strength of Arguments on the Basis of a Linguistic Analysis: A Synthesis

Mathilde Janier and **Patrick Saint-Dizier** [1]

Abstract. In this contribution, we present several layers of linguistic analysis, the aim of which is to provide indications on the strength of arguments in context. This contribution proposes a synthesis of existing resources to evaluate strength also used in opinion analysis, then it develops features which are proper to argument strength. Linguistic elements related to (1) the argument contents, (2) the discourse structures associated with this argument (which may introduce restrictions), (3) the nature of argument schemes used, and (4) some rhetoric elements are investigated.

1 INTRODUCTION

There are several ways to measure the strength of arguments. The strength can be measured from a logical and pragmatic perspective or it can be measured from a language point of view. Both approaches are not necessarily coherent but they must be combined to produce a relatively accurate measure of strength. Argument strength may be measured for each argument in isolation or for groups of related arguments, taking into account their relations and structure.

In this contribution, an argument is composed of a claim and of one or more propositions P_i which support or attack the claim. Claims and propositions P_i have their own strength. In this contribution, we first identify linguistic phenomena and their related cues which are a priori marks of strength on propositions P_i taken in isolation. We then integrate this analysis into a larger view where a proposition P_i is associated with discourse structures which may reinforce or weaken its strength. In a subsequent stage, sets of related propositions P_i are considered, so that their relative strength can be characterized on the basis of linguistic factors. Finally, the impact of argument schemes and rhetoric cues is explored to give an overall picture of how argument strength based on linguistic analysis can be measured. Priority is therefore given to linguistic analysis, in which results of lexical semantics are relatively stable and accurate, over a more pragmatic and intuitive analysis of argument strength.

This investigation and analysis is carried out within the framework of argument mining and analysis in which, given a controversial issue, arguments for or against this standpoint are mined in different types of texts (see for example [11], [10]). Besides supporting or attacking an issue, propositions P_i may also attack or support each other. The problem of the relatedness between a claim and propositions P_i has been addressed in [16], it will therefore not be discussed in this contribution which focuses on a crucial and difficult parameter: evaluating the potential strength of an argument. In our perspective, persuasion is a kind of contextual evaluation of the strength of an argument. This will not be addressed here, although it is clear that it should be the ultimate component of such an investigation.

[1] CNRS-IRIT, email: mathildejanier@hotmail.com, stdizier@irit.fr

Quite a large number of investigations, more or less successful and ad'hoc, have been developed within the framework of opinion analysis. This document reviews the main results and develops additional or more specific material proper to argument strength analysis. In opinion analysis, platforms and resources such as Sentistrength (http://sentistrength.wlv.ac.uk/) and the Stanford Sentiment Treebank. Major synthesis on opinion strength are developed in [17] and [8]. However, if some features are shared with argument strength analysis, argument strength is more complex to characterize since an argument is a complex system composed of a claim, one or more justifications, and quite frequently some forms of evidence, backing and warrant and rebuttals. Qualifiers may also be stated. Finally, the nature of the argument scheme that has been used may be crucial. This contribution develops a synthesis of a number of these aspects.

Investigations on argument strength have focused on a few aspects such as (1) teaching how to organize written essays and how to organize arguments and give them an appropriate strength, (2) research on persuasion which is, in our view, an analysis of strength in contexts (domain and listeners are taken into account), and (3) in theoretical analysis of argumentation where graphs of attacks and supports are developed. Let us note for example [6] that deals with an in-depth analysis of persuasion, [23] which investigates the content of persuasive messages. Sensitivity to argument strength of various populations is developed in e.g. [3].

The relation of strength with rhetorical questions has been addressed in e.g. [12]. A number of linguistic factors are analyzed in e.g. [1], and later in [18], [19]. However, to the best of our knowledge little has been done to characterize argument strength from a linguistic point of view, within the perspective of argument mining. This article is a contribution to this perspective, it also outlines the high context sensitivity of linguistic factors.

This paper is organized as follows. In a first stage, the contribution to argument strength of individual lexical items found in propositions P_i is investigated. The hypothesis is that such propositions have an intrinsic strength independently of the claim. Lexical semantics structures to organize linguistic data are introduced. Then, the strength variations induced by the combination of several lexical items in a proposition and the support construction in which it may be embedded are explored and tested experimentally. Since it turns out that contextual effect in its broad sense is crucial to have an accurate estimate of the strength of an argument, several contextual parameters are discussed, in particular the impact of the discourse structures which are adjoined to the argument or to a proposition P_i and the kind of argument scheme on which the argument relies.

2 THE ANALYSIS CORPUS

Several types of corpora are used to carry out this investigation. Documents are oral or written, they are essentially in English with a few of them in French. Our corpus is composed of the following elements:

- a corpus of debates extracted from the BBC Moral Maze, analyzed in conjunction with the university of Dundee, with about 2000 arguments [2], available on their platform,
- a corpus of consumer opinions in the hotel domain [20] and in relation with the services offered by the local French airline HOP! (French texts); this corpus is particularly rich in discourse structures which modify argument strength and scope, which includes about 250 arguments,
- a corpus of short texts, that contains about 150 arguments, used to identify the relatedness problem [15], [16] from the domains of vaccination, nuclear energy (in French), and women's condition in India.

These corpora are quite diverse in nature and linguistic characteristics, they allow an accurate identification of the linguistic elements at stake in the expression of strength. This corpus is used to identify and evaluate the importance of various linguistic constructions and linguistic resources in strength expression, it is therefore difficult to evaluate its relevance a priori. The main point is that it contains a large diversity of types of statements so that most linguistic phenomena can be observed, but probably not quantified, which is not our aim at this stage.

3 STRENGTH FACTORS WITHIN A PROPOSITION

Given a claim, propositions P_i for or against it are essentially evaluative statements. These may be direct evaluations or may require knowledge and inference to identify what is evaluated and how. The terms used in propositions P_i to provide an evaluation of a standpoint induce a polarity for the argument and strength indicators which say whether their attack or support is strong or weak. This section explores the linguistic phenomena and their related cues, within the propositions P_i, which may potentially be strength indicators. The hypothesis is that such propositions have an intrinsic strength independently of the claim, which is explored in this contribution. It is however clear that the strength of the argument is a combination of the strength of the claim and of the proposition that supports or attacks it.

Evaluating the strength entailed by linguistic cues is quite subjective. Our goal is to collect those marks and to structure them according to scales. Evaluating their real impact in context requires measures which go beyond this analysis, but this is a necessary step. Some simple elements are given in section 3.4. Each linguistic cue is investigated in isolation, then the correlation of several cues is investigated.

Two levels of the expression of strength are considered here: (1) the implicit strength conveyed by head terms used in propositions and (2) the strength conveyed by expressions, such as propositional attitudes expressions, of which a proposition P_i is the sentential complement. The propositions P_i considered in this investigation have a simple syntactic structure. They are composed of a main point called the kernel and adjuncts – in general discourse structures – which add e.g. restrictions, justifications, purposes or illustrations to the kernel. These discourse structures may scope either over the proposition or over the entire argument.

The linguistic resources which are used are those of our TextCoop platform, with which discourse analysis and argument mining is realized. Resources considered in this investigation come for a large part from general purpose and domain dependent lexical resources we developed for opinion analysis.

3.1 A categorization of the expression of strength

As also shown in the area of opinion analysis, there are many elements which may have an impact on the strength of a proposition P_i. Those with a higher impact are head elements such as verbs, and elements which are less prominent in the syntax such as evaluative adjectives and adverbs. These latter are analyzed as adjuncts to the noun for adjectives and to the VP or the sentence for adverbs. These linguistic elements are used to determine the orientation of the propositions P_i w.r.t. the claim (support, neutral, attack). In addition, their implicit semantics is an important factor to evaluate the overall strength of an argument.

The main categories of elements internal to a proposition P_i which may impact the strength are:

1. positively oriented verbs, such as:
   improve, benefit, optimize, reinforce, preserve, strengthen, guarantee, consolidate.
   e.g. *vaccination against Ebola is necessary because it guarantees the non-proliferation of the disease.*
   There are many such verbs, the semantic function of which may vary over domains.
2. negatively oriented verbs, such as:
   affect, alter, break, demolish, hurt, lessen, ruin, undermine, damage. For example, given the claim:
   the situation of women in India has improved,
   it is attacked by the proposition:
   the persistent lack of education largely affects their independence.
3. similarly to verbs, a number of adjectives and adjectival compounds contribute to the orientation and strength expression. These are usually found in propositions where the verb is neutral (auxiliary, light verb, verbs such as *allow, enable*, where the orientation of the object is crucial) or is largely underspecified w.r.t. to polarity and strength. Adjectives in this category are, for example:
   useful, capable, consistent, resistant, compliant, beneficial, optimal
   for the positively oriented ones and:
   risky, polluted, dangerous, weak, harmful
   for the negatively oriented ones. A typical example is e.g. :
   vaccination against Ebola is dangerous because the adjuvant is toxic, where *toxic* induces the orientation and the strength.
4. expressions derived from verbs, past participles, and adjectival compounds with an evaluative or scalar dimension such as:
   disappointing, potentially risky.
   For example, a negatively oriented argument in relation with a standpoint on the necessity of nuclear plants is:
   Pipe corrosion in nuclear plants is potentially risky.
5. nouns which appear as NP objects in the proposition which have a positive or negative orientation, e.g.: *risk, disease, reward, success*.

The expression of strength is also mediated by a number of terms which introduce propositions P_i. These are called **control constructions**, they sub-categorize for a proposition or a sentential complement which is here a proposition P_i. These constructions, although found in opinion analysis, are more developed in argumentation. They also appear in dialog analysis in general. The scope of these constructions is the entire argument, not the justification, as it would be the case in opinion expression. Control constructions can be organized according to the following linguistic categories:

1. Propositional attitude verbs and expressions: besides the expression of agreement or disagreement, which is their main aim, most of the elements of this category have an implicit weight. In this class fall verbs and expressions such as:
   think, believe, agree, deny, argue, refute, acknowledge, reckon, disagree, accept, reject.
   The semantics of these verbs is investigated in depth in [22]. These elements have different weights which may depend on the context and personal interpretations, for example, *believe* may be weaker or stronger than *think*. Propositional attitude constructions do not have, a priori, an impact on the argument orientation. Propositional attitude constructions can be modified by a negation or by a modal such as *would, could, have to* as in: *I would argue that, I have to acknowledge that*. These may impact the strength.
2. Psychological expressions or expressions denoting a desire, a position or an experience. These expressions may be at the origin of the expression of a doubt or a weak support; they include verbs and expressions such as:
   I feel, I am worried about, I am intrigued by, dream of, be encouraged by, tend to.
   These terms are often in an initial position or in a final position for constructions such as *worries me* as in: *the obligation of vaccination worries me*, where the nominalized sentence is raised to play the role of the subject.
3. Report verbs and associated constructions. They introduce arguments and propositions P_i in a direct manner or as a reported speech from e.g. other participants in a debate or from external persons, frequently considered as experts (see also Section 4.2). Similarly to the two above categories, these constructions can be modified by a negation or a modal. In this category fall expressions such as: *report, say, mention, stated, announced, discuss, claim* and their morphological variants. Identifying the strength of these terms is difficult: while *report, say, announced* are rather neutral, terms such as *claim, stated* are much stronger. For example, given the claim:
   Ebola is a dangerous disease,
   a strong attack may be:
   the authorities of Guinea claimed that there is no risk of proliferation of Ebola.
4. Epistemic constructions. These also occur quite frequently, they include expressions such as:
   know, my understanding is that, I am convinced that, I suppose, I realize, it is reasonable to assume, infer, implies, I can see.
   While some introduce doubts or uncertainty, others are clear affirmations of a certain knowledge that may contradict or support a standpoint.
5. Modal expressions. These behave as left adjuncts and modify some of the expressions described above or may be adjoined to the head verb of the argument. Most of them either weaken the statement or introduce forms of hypothesis:
   might, would, must, have to, could be, should be possible, it is reasonable to, can mean, may mean.
   For example, for the claim on vaccination against Ebola, a weak support could be:
   a systematic vaccination could define sanitary belts to avoid the proliferation of the disease.
6. Adverbials related to the expression of opinion. In our corpora, they frequently increase the strength of the arguments. They include:
   probably, necessarily, most definitely, definitely, surely, usually, frequently, often, certainly, of course, obviously, generally speaking, of course, indeed.

3.2 Structuring expressions of strength by semantic category

It is obviously impossible to a priori assign strength values to the terms given in the different categories given above, nor is it possible to assign weights to their combinations. A option is to structure these terms along scales [4], as for scalar adjectives in opinion analysis. In this experiment, it turns out that the polarity of about 75% of the adjectives are domain independent. While the adjectives used in opinion expression lend themselves relatively easily to an evaluation of their positive or negative character, this is more complex for verbs, modals or the expressions categorized above. To organize the elements in the different categories, an experiment is made using non-branching proportional series (Cruse 86) which allow to define partial orders over groups of terms w.r.t. a given measurable property. These scales organize terms of a category from those with a strong negative orientation to those with a strong positive orientation. A neutral point is mentioned: it is a term when such a term exists or an abstract point. The partial order introduces some flexibility by allowing several terms to be at a given point on the scale when it is not relevant to make strength distinctions between them.

Our approach is
(1) to classify the terms of each category in a dedicated scale following their standard semantics,
(2) to evaluate the results and to possibly revise the classification according to the results obtained from the experiment reported in 3.4.

For example, the negatively and positively oriented verbs given above (3.1, items 1 and 2) are structured as follows:
`[[ruin] - [break, demolish] - [affect, alter, lessen, undermine, damage] - [hurt] - Neutral - [preserve, guarantee] - [benefit] - [improve, consolidate, strengthen] - [optimize]].`
Terms which are considered to have almost the same strength appear in the same set, represented between square brackets. The neutral point is represented by the constant 'Neutral', the two sets around it have a moderate strength while the extremes sets are the strongest ones.

Adjectives are more difficult to structure because they do not modify in an homogeneous way the same property, for example, *resistant* and *optimal* may not operate on the same concepts, where *optimal* is rather higher-order. A global scale such as the following can however be developed:
`[[dangerous, harmful] - [risky, polluted] - [weak] - Neutral - [useful, capable, consistent, beneficial] - [resistant] - [optimal]].`
In this example, a certain number of adjectives is in the same set since these have a relatively similar impact on strength.

Finally, a scale for propositional attitude verbs is the following:
`[[deny - refute - reject] - [disagree] - Neutral - [believe, think, accept] - [agree,`

`acknowledge, reckon] - [argue]].`

The verbs to the extreme of the scale are more crucial in the acceptance or rejection of the claim than those close to the Neutral point. Adverbials modify these verbs or the VP they head by adding or reducing the strength. These can be classified as follows by increasing strength:

`[[probably] - [indeed, usually, of course] - [often, frequently, generally speaking] - [definitely, surely, obviously, necessarily] - [most definitely]].`

3.3 Strength representation when combining categories: a basic model

It is frequent to have propositions P_i that include terms from the two levels presented in section 3.1: a first level of strength is expressed within the proposition and then the proposition is embedded into a variety of constructions from the second set of categories. For example, given the claim:

Nuclear plants are useful since they pollute less than coal or oil.

a proposition such as:

I am definitely convinced that nuclear plants should be banished.

includes the strong negative term *banished* in its statement, which is somewhat soften by the modal *should*. This proposition is included into an epistemic construction with a strong connotation: a strong verb *convinced* modified by the intensifier adverbial *definitely*. Evaluating the strength of such a proposition compared to e.g.:

I am convinced that nuclear plants must be banished.

is not trivial, even for human experts.

To have an accurate analysis of the strength of propositions P_i, a semantic representation of the elements which contribute to strength expression is developed. It is based on the categories of the elements found in the proposition and on a rough estimate of their strength, as reflected by the non-branching proportional series presented in section 3.2. For example, the proposition:

Nuclear plants should be banished.

has the following semantic representation w.r.t. its strength:

$[_{argument}$ verb(strong negative) $\wedge$ modal(weaken)].

where *banished* is among the strongest negative verbs on the corresponding scale while the modal *should* weakens the strength of this verb. Then, the whole proposition:

I am definitely convinced that nuclear plants should be banished,

which includes an epistemic construction, is represented as follows:

$[_{control}$ verb(epistemic, strong positive) $\wedge$ adverbial(reinforce)]($[_{argument}$ verb(strong negative) $\wedge$ modal(weaken)]).

Let us call this expression the **signature of the strength of the proposition**. Considering the different elements of this representation, the resulting strength is strong with a negative orientation.

A simple way to identify the strength of an proposition is to develop composition equations:

- in the proposition: the head terms are the verbs or the adjectives. They a priori have a polarity and a strength level which is lexically induced. A standard scale with 5 values: [null, weak, average, high, maximal] is used in this first experiment. These lexical structures may be combined with intensifiers which are modals for verbs and adverbs for adjectives. Intensifiers weaken or reinforce the strength of the element they modify. For example, if *banished* has the strength 'high' with a negative orientation, then *should* lower it to 'average' while preserving its orientation.
- The same strategy holds for the structure in which the proposition is embedded. For example, *I am definitely convinced* is composed of a head verb with strength 'high', and the adverbial *definitely* increases its strength to 'maximal'.
- The strength and orientation of a proposition are combined with the control structure in which it is embedded. The resulting strength is a function of the strength of each structure, for example the average. In our example, 'maximal' must be combined with 'average', leading to 'high'.

This model takes into account the different linguistic parameters of an proposition, it is however very simple: it is based only on linguistic considerations and on an a priori strength evaluation of each lexical element. It does not take into account other crucial factors such as the context of the utterance, the argument schemes used, the domain and its style, the intonation and the preceding claims and propositions P_i.

3.4 An experimental evaluation of strength based on annotations

It is difficult to ask annotators to evaluate the strength of propositions P_i without any analytical support. The model provided in the previous section, although quite simple, can be used as a support for annotators who can concentrate on each element separately and then make a global evaluation of the strength. In this section, a new protocol for strength analysis is introduced.

The idea is to automatically annotate propositions P_i with the values described in 3.3, and then to ask human annotators to indicate their own evaluation for (1) a proposition, (2) the embedding structure when it exists, and (3) and the combination of the two. In a subsequent stage, discourse structures will also be annotated using our TextCoop platform. The manual annotations can then be compared to the annotations produced by the system as described in section 3.3. A more accurate model of strength analysis can then be developed from these two evaluations.

Let us now illustrate the annotation structure that the annotator uses. The annotator must specify the strength and possibly the orientation for each of the uninstantiated attributes (strength, orientation, intensity). The above example is annotated by the system, based on a lexical and surface syntactic analysis; values are left open so that annotators can filled them in:

```
<proposition strength= , orientation = >
<support strength= , orientation = >
I am definitely convinced that </support>
<kernel strength= , orientation = >
nuclear plants
<modal intensity = > should </modal>
<verb strength= , orientation = > be banished
</verb> </kernel> </proposition>
```

with strength $\in$ [null, weak, average, high, maximal]
orientation $\in$ [positive, neutral, negative]
intensity $\in$ [lower, increase].

A first, preliminary experiment aims at identifying the strength differences as postulated a priori by the linguistic description and as perceived by humans. Contextual effects, such as the style or the strength of other arguments, are not taken into account in this first experiment in order to concentrate on propositions and arguments strength taken in isolation. The contextual dimension will be considered in a second stage (see section 3).

In this initial experiment, a set of two hundred propositions have been annotated. These are constructed via lexical variation, to accurately evaluate the impact of each lexical item, from 15 original claims. In these propositions, the lexical items which originate the strength are substituted by others, e.g. *convinced* becomes *feel* and then *believe*. Substitutions are the identical, as much as possible to preserve relevance, over the 15 claims to preserve the homogeneity of the results. A total of 38 lexical items are tested in various linguistic realizations. The goal is to validate the protocol and have preliminary results before starting a larger experiment.

The strength values are transformed into numbers to allow numerical computations. The following parameters are investigated:

- the strength $S1$ associated with each lexical term: the different values associated with each lexical item are averaged, to produce their average individual strength estimate. Then, a partial ordering similar to the linguistic ordering presented in section 3.3 is constructed based on these values.
- the strength $S2$ of a proposition P_i, elaborated from the individual strength of each element it is composed of, is then computed. This computation reflects the strength of combinations of several lexical items. Vectors are produced to represent all the lexical combinations, e.g.:
  `[lexical head L1, strength S1, lexical intensifier L2, orientation O, Resulting strength R].`
  For example:
  `[banished, 4, should, lower, 3].`
- the strength $S3$ of the support construction is elaborated in a similar way when it exists, it is equal to 1 when there is none,
- the global strength $S4$ of the proposition including the support construction when it exists. For this level, the following vector is considered:
  `[support strength S3, proposition strength R, global strength G].`

The results of this experiment are not very surprising:

- the individual strength of lexical items taken in isolation is very similar to the series developed independently of any context from a linguistic point of view, only 2 elements are classified differently,
- the proposition strength shows a variation of 15% compared to the linguistic estimate of section 3.3. It is lower in 80% of the cases. Most of the modals are interpreted as lowering the strength and a gradation is expected: moderate lowering and strong lowering, which would decrease the initial strength by 2 instead of just 1. The prominent role played by modals in the strength expression is an important result of this task,
- the support construction strength shows a variation of a maximum of 25% either above or below the linguistic estimate. This can be explained by the difficulty to interpret the strength of terms such as *believe* compared to *think*. These terms are in fact context and speaker dependent,
- the combination of the proposition and its support shows a variation of about 30% around the linguistic estimate, which is relatively large and questions the validity of the linguistic classification taken in isolation.

This simple and preliminary experiment shows that while there is a relative stability on the strength of terms such as verbs and adjective, the strength evaluation is less stable for modals, and needs some important adaptations for support constructions and their combination with or influence on the claim and vice-versa. This motivates the second step of our investigation: taking into account various forms of context, which should allow to have a more reliable estimate of the strength of support constructions and modals.

4 OTHER FACTORS OF STRENGTH

Several other factors, which are essentially contextual, have a major influence on the strength of propositions P_i and on arguments more generally. Their influence is however difficult to accurately analyze. These factors are explored in this section. The results of the previous section (3.4) indeed show that the strength induced by some lexical items depends on the context of the utterance.

The first factor are the discourse structures which may be adjoined to a proposition or an argument that describe e.g. circumstances, conditions, restrictions, etc. This factor has been investigated within the RST framework (http://www.sfu.ca/rst/). The second factor is the argument scheme that has been used. Some have a higher strength or reliability than others. The third factor is the context of the proposition: it may be uttered in isolation or it may be part of a series of propositions P_i and of arguments. As developed in e.g. [18], propositions associated with a claim may be structured as series or in parallel. In the first case, the strength is the strength of the weaker one, and in the second case it is the strength of the strongest one. This type of factor is not found in opinion analysis where statements are in general treated in isolation. The fourth factor is the syntactic structure of the premise-conclusion pair where focus shifts can be observed via for example left-extraposition. The last factor is the linguistic context of the utterance. For example some debates may only use soft arguments in order to remain polite and to avoid strong attacks, whereas others use extremely strong terms even for arguments which are not crucial.

In this section, the impact on argument strength of the first two factors is discussed. The remaining ones require additional investigations.

4.1 Influence of discourse structures on argument strength

As in any form of elaborated discourse, arguments are quite frequently associated with elements such as comments, elaborations, comparisons, illustrations, etc. which can be considered as either forms of explanation or secondary or subordinated arguments. These discourse structures are borrowed from the RST ([9], see also http://www.sfu.ca/rst/) considered within the perspective of argument strength analysis. These structures frequently implement argument schemes [21] and applied to opinion analysis [20], as developed in section 4.2 below. In our view, explanation is not a basic rhetorical relation as introduced in RST, but a very generic construction, a 'proto-relation', which covers a large number of communication and argumentative situations.

For the claim:
Ebola vaccination is necessary,
the statement:
the Ebola vaccine is easy to use for emerging countries (cheap, can be transported without any need for refrigeration, active for a long time)
is argumentative where the expression : '(cheap, can be transported without any need for refrigeration, active for a long time)' can be analyzed (1) either as an elaboration or as an illustration of the head expression ' easy to use for emerging countries' (2) or as a secondary or subordinate proposition which supports the main one. In RST theory, the head expression is a nucleus while the elaboration or illustra-

tion is its satellite. The explanation or secondary proposition which supports the main one increases the strength of 'easy to use'.

However, the role of illustrations w.r.t. to argument strength is not easy to determine. Given the claim:
I do not recommend this hotel,
in a proposition such as:
The bathrooms were in a bad condition: [ILLUSTRATION *the showers leaked, and the plug mechanism in the bath jammed ...*],
the illustrations given to support the diagnosis ('bad condition') do not seem to reinforce or weaken its strength. They are interpreted as reformulations which are another way to say something without altering the initial content. The difference between these two examples is the contribution of the illustration: in the first example 'easy to use' is rather vague and is reinforced by the example, whereas in the second example 'bad condition' is more precise and remains at the same strength level.

Let us consider other types of discourse relations such as the circumstance and justification relations. For example, given the standpoint:
Ebola is a dangerous disease,
a justification may weaken a strong proposition, instead of supporting it:
[JUSTIFICATION *in order to avoid any form of panic or, worse, of bio-terrorism*], *the authorities of Guinea claimed that there is no risk of proliferation of Ebola.* In the following example, possibly with a form of irony, the strength and polarity of 'breakfast is excellent' is largely affected – if not reversed – by the contrast:
The breakfast is excellent [PRECISION *with very imaginative exotic fruit salads*] [CONTRAST *but most of the products are not fresh and most have passed their sell-by date*].

More complex – yet realistic – arguments associated with restrictions of various sorts make the identification of the overall strength quite challenging:
[CONTEXT *We stayed here for a one day conference off-season*], *and the hotel was OK* [CONCESSION *- although the room I had was kind of weird.*] *I think it was the sitting room to the suite on the top floor* [PRECISION *- the bed was a fold-out bed, not comfortable,* [CONCESSION *(slept okay though)*], *and the coffee table was small, dirty and pushed to the side.*] [CONCESSION *It did have a lovely terrace though*] *- shame it was raining cats and dogs.* [RECOMMENDATION *Not a great experience.*]
Depending on customers' preferences, this opinion can be judged to be slightly positive or negative, in spite of the negative polarity of the recommendation, which turns out to be the main argument. Therefore, this opinion may either support of attack the standpoint *I do not recommend this hotel*.

Evaluating the impact of discourse structures is therefore a very challenging task. Even if the polarity and strength of each individual structure can be evaluated, their combination with the main argument and their interactions when there are several structures is complex and highly domain dependent. We are now exploring various types of experimental protocols which could contribute to this analysis. The discourse structures shown in the examples are recognized by our TextCoop platform with an accuracy of about 90% [14]. The challenge is now to go into the semantics of each structure.

4.2 The impact of argument schemes on argument strength

Another component to follow is to explore the inner structure of an argument and the underlying scheme that has been used. [21], [13], have identified and structured a large number of schemes which are used in everyday argumentation. Some of them can be detected via a linguistic analysis [5], [7]. These can provide information on the strength of arguments. A number of schemes among the most frequently encountered are reviewed in this section.

4.2.1 Argument from analogy

The typical form of arguments from analogy is as follows:
Premise 1: Generally, case C1 is similar to case C2.
Premise 2: A is true (false) in case C1.
Conclusion: A is true (false) in case C2.

For example:
It has been shown that vaccinating against malaria can be useless in some cases; similarly, the vaccine against Ebola is not recommended.
This sentence makes an analogy between two serious diseases and tries to show that if the vaccine against one of these diseases is useless then the vaccine against the other is useless too. Some linguistic cues marking analogy are: *similarly, x is like y, doing x is as [adjective useful, dangerous, crucial] as doing y.*

Metaphors can also mark analogy. For instance, *Ebola is a war which has to be fought.* An analogy is made between Ebola and war. This type of construction has often been used in literature, some metaphors are now well-known and used in everyday conversations, which proves that its rhetorical effect is high; as a consequence, arguments from analogy may have a strong impact.

4.2.2 Argument from expert opinion

The typical structure of arguments from expert opinion is:
Premise 1: E is a reliable authority in the domain S.
Premise 2: A is a proposition contained in S.
Premise 3: E asserts that A.
Conclusion: Therefore, A.

An example of argument from expert opinion is :
Depression and anxiety should be taken seriously. The London School of Economy reports that half of all illnesses in the under 65s is mental.
In this example, the conclusions of a group of people who has expertise in the domain of health are used to support the claim that mental illnesses have to be taken seriously.

Arguments from expert opinion are marked by two linguistic cues; first, nouns which name the expert, e.g. *expert, doctor, economist, politician* etc.; second, constructions such as reported speech which allow indicating the expert's opinion, e.g. *claim, warn, explain, indicate*, etc. The strength of these report verbs (as suggested in section 3.2) must be taken into account in the scheme. When there is no explicit cue, additional knowledge may be necessary to determine whether a person is an expert. For instance, *Stephen Hawking warned against risks linked to the development of AI* can only be understood as being an expert opinion if one knows that Stephen Hawking has long been working on Artificial Intelligence.

The opinion of experts is used in many cases to support a claim since it is hard to contradict an expertise. As a consequence, arguments from expert opinion have a strong impact. However, the strength of the argument can be critiqued by questioning the knowledge of the experts. For instance, in the above example, one may wonder whether the London School of Economy definitely has expertise in the health domain (see also Section 4.2.8).

4.2.3 *Argument from negative consequences*

This scheme has the following form:
Premise 1: If an action leads to bad consequences, all else being equal, it should not be brought about.
Premise 2: If action A is brought about, bad consequences will occur.
Conclusion: Therefore A should not be brought about.

Vaccinating people against Ebola has reduced their immune system. This vaccine must not be used anymore.
is an argument from negative consequences.

Negative adjectives and nouns are usually found in the premise(s) (here, *reduce*), while action verbs used in the negative form are used in the conclusion (here, *must not be used*). However, these cues are extremely domain dependent. Warning against negative consequences can have a strong impact, but the nouns and adjectives used can help determining how strong the argument is.

4.2.4 *Arguments from examples*

This scheme has the following form:
Premise 1: Example 1 is an example that supports claim P.
Premise 2: Example n is an example that supports claim P.
Conclusion: Claim P is true.

For example: *It has been shown that the vaccine is not the right solution. For example, two weeks after the injection, an old man died and the foetus of a pregnant woman shown malformations.*

Linguistic cues typical of the illustration discourse relation such as *for example, for instance, in the same manner* can contribute to detect the arguments from example. However, these cues are not always linguistically realized, for instance, the same argument could be presented as follows:
Two weeks after the injection, an old man died and the foetus of a pregnant woman presented malformations. The vaccine is not the right solution.

Evaluating how this form of strength interacts with the others, presented above, requires some experimentation. It is not clear, for example, if they all operate at the same level, or if some have a higher weight.

The strength of the argument can be measured with the number of examples used. The above argument has two premises (two examples) supporting the claim. The conclusion could be supported by many other examples of people who badly reacted to the vaccine, which would reinforce the claim that the vaccine is not the right solution.

4.2.5 *Arguments from position to know*

This scheme has the following form:
Premise 1: Source a is in a position to know about things in a certain subject domain S containing proposition A.
Premise 2: a asserts that A (in Domain S) is true (false).
Conclusion: A is true (false).

For instance: *A British politic visiting Western Africa has revealed that the number of deaths due to Ebola has dropped since the vaccination began. Vaccinating populations must therefore continue.*
In this example, the claim that vaccinating against Ebola must continue is supported by the opinion of a British political person. This type of argument is close to arguments from expert opinion. However, arguments from position to know are weaker that arguments from expert opinion because it is easier to question whether the person who is being quoted has the right information.Similarly to arguments from expert opinion, reported speech can help detecting arguments from position to know.

4.2.6 *Argument from popular opinion*

Arguments from popular opinion take the following form:
Premise 1: Everybody is doing X.
Premise 2: X is a good thing to do.
Conclusions: Therefore, X must be the right thing to do.

As an example:
vaccination in general is a cheap and efficient way to get rid of major diseases, therefore all populations exposed to Ebola must systematically undergo vaccination.

Linguistic cues referring to populations and group of people can help detect arguments from popular opinion, e.g. *the population, people, individuals, everyone, all the persons,* etc. The use of numbers or percentages can also mark the strength of the argument. Similarly to arguments from position to know, arguments from popular opinion have less strength than the ones from expert opinion since the action (or opinions) of groups of people can be discussed.

4.2.7 *Arguments from cause to effect*

This scheme has the following form:
Premise 1: Doing X will cause Y to occur or If X occurs then Y will occur,
Premise 2: X is done or X occurs,
Conclusion: Y will occur.

The statement: *A new vaccine has been developed which will lower the number of deaths. The first vaccinations have begun last week. Less farmers in the vaccinated area will die after its injection.*
is an example of argument from cause to effect. This type of argument can be seen as an anticipation: future effects are foreseen; as a consequence, linguistic cues to detect such arguments are uses of future tenses or conditional. Anticipation has however little credibility in many cases, as a consequence, arguments from cause to effect are weak arguments.

4.2.8 *Organizing schemes w.r.t. their strength*

From the observations above, a tentative classification of argument strength induced by argument schemes can be made. In our case, no domain knowledge is considered, which could affect this classification:
Strong: analogy, expert opinion
Moderate: negative consequences, from examples
Weak: position to know, popular opinion, cause to effect.

In Walton, each scheme is associated with a number of critical questions which allow testing the soundness of the argument; these can be used to attack the argument.For instance, the argument from analogy has the following critical questions:
CQ1: Are there respects in which C1 and C2 are different that would tend to undermine the force of the similarity cited?
CQ2: Is A the right conclusion to be drawn in C1?
CQ3: Is there some other case C3 that is also similar to C1, but in which some conclusion other than A should be drawn?

Here are the critical questions for arguments from position to know:
CQ1: Is a in a position to know whether A is true (false)?

CQ2: Is a an honest (trustworthy, reliable) source?
CQ3: Did a assert that A is true (false)?

The critical questions for arguments from expert opinion are:
C1: Is E a genuine authority?
C2: Did E really assert A?
C3: Is E an authority in the right field?

Evaluating the overall strength of critical questions per scheme can be used to determine the strength of the scheme w.r.t. an argument. An argument which has stronger critical questions could be a weak argument (it can be easily attacked), or, on the contrary, it can be a strong one (it is difficult to defeat it).

Finally, the problem of fallacious arguments can interfere with the strength evaluation. For example, analogy is sometimes classified as fallacious. As (Walton et al. 2008, p 49) note: 'the problems seems to be that argument from analogy is a plausible form of argument only when it is used for guessing; it is not good enough to be used to prove a claim.' Evaluating fallacious arguments is a major concern in argumentation, however, in practical situations like ours, this means considering domain and general purpose knowledge and inferences.

5 CONCLUSION

In this contribution, we have surveyed a number of linguistic factors which contribute to the expression of argument strength. We proposed a categorization and a model to structure lexical items which may convey strength. We have outlined the elements which are proper to argument strength analysis and those which may be shared with opinion analysis. We have outlined, via a short experiment their sensitivity to context, taken in its broader sense: including the utterer, the listeners or readers, the domain and context of the arguments. In a second part, we have explored the impact of discourse structures and argument schemes on the expression of strength.

These different features show that it is difficult to evaluate a priori the strength of a a proposition that supports or attacks an argument. The weight of the different components of strength need a careful experimental analysis, and their interactions with context require the development of a model that includes language aspects as well as knowledge and specific forms of reasoning.

REFERENCES

[1] Jean Claude Anscombre and Oswald Ducrot, *L'argumentation dans la langue*, Mardaga, 1983.

[2] Kasia Budzinska, Mathilde Janier, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yakorska, 'A model for processing illocutionary structures and argumentation in debates', in *Proceedings of LREC14*, (2014).

[3] J Corella, S Spencer, and M Zannab, 'An affirmed self and an open mind: Self-affirmation and sensitivity to argument strength', *Journal of Experimental Social Psychology*, **40(3)**, (2004).

[4] Alan Cruse, *Lexical semantics*, Cambridge University Press, 1986.

[5] VW Feng and Greame Hirst, 'Classifying arguments by scheme', in *Proceedings of 49th ACL: Human Language Technologies*, (2011).

[6] Floriana Grasso, John Ham, and Judith Masthoff, 'User models for motivational systems - the affective and the rational routes to persuasion', in *Advances in User Modeling - UMAP 2011 Workshops*, (2011).

[7] Nancy Green, 'Manual identification of arguments with implicit conclusions using semantic rules for argument mining', in *EMNLP17, workshop on Argument Mining*, (2017).

[8] B Liu, *Sentiment Analysis and Opinion Mining*, Morgan and Claypool Publishers, 2012.

[9] William C Mann and S A Thomson, 'Rhetorical structure theory : Toward a functional theory of text organization', *Text*, **8(3)**, 243–281, (1988).

[10] Raquel Mochales and Marie-Francine Moens, 'Argumentation mining', *Artificial Intelligence and Law*, **19**(1), 1–22, (2011).

[11] Raquel Mochales Palau and Marie-Francine Moens, 'Argumentation mining: the detection, classification and structure of arguments in text', in *Proceedings of 12 ICAIL*, (2009).

[12] J Munch and J. Swasy, 'Rhetorical question, summarization frequency, and argument strength effects on recall', *Journal of Consumer Research*, **15(1)**, (1988).

[13] Simon Parsons, Sue Atkinson, Keith Haigh, K. Levitt, and P. McBurney, 'Argument schemes for reasoning about trust', in *Comma08*. IOS Press, (2008).

[14] Patrick Saint-Dizier, 'Processing natural language arguments with the textcoop platform', *Journal of Argumentation and Computation*, **3(1)**, (2012).

[15] Patrick Saint-Dizier, 'The bottleneck of knowledge and language resources', in *Proceedings of LREC16*, (2016).

[16] Patrick Saint-Dizier, 'Knowledge-driven argument mining based on the qualia structure', *Journal of Argumentation and Computation*, **8(2)**, (2017).

[17] M Taboada, J Brooke, M Tofiloski, K Voll, and M Stede, 'Lexicon-based methodsfor sentiment analysis', *Computational Linguistics*, **37(1)**, (2011).

[18] Frans van Eemeren and Rob Grotendorst, *Argumentation, Communication and Fallacies*, Lawrence Erlbaum, 1992.

[19] Frans van Eemeren, Rob Grotendorst, and Francisca Snoeck Henkemans, *Argumentation, Analysis, Evaluation, Presentation*, Routledge, 2001.

[20] Maria Garcia Villalba and Patrick Saint-Dizier, 'Some facets of argument mining for opinion analysis', in *Proceedings of COMMA12, IOS publishing*, (2012).

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

[22] Anna Wierzbicka, *English Speech Act Verbs: A semantic dictionary*, Academic Press, 1987.

[23] X Zhao, A Strasser, J N Cappella, C Lerman, and M Fishbein, 'A measure of perceived argument strength: Reliability and validity', *Communication methods and measures*, **5(1)**, (2011).