

Contents

Trends in Argumentation Logic (<i>Sofia Almpani, Elena Lisanyuk, Andrew Schumann</i>).....	1
Argumentation: Reasoning Universalis (<i>Antonis Kakas</i>).....	6
Ambiguity in Argumentation: The Impact of Contextual Factors on Semantic Interpretation (<i>Dimitra Serakioti, Petros Stefaneas</i>).....	18
Dynamic Approximation of Self-Referential Sentences (<i>Vladimir A. Stepanov</i>).....	25
Determining Argumentative Dispute Resolution Reveals Deep Disagreement Over Harassment Issue (A Case-Study of a Discussion in the Russian Parliament) (<i>Elena Lisanyuk</i>).....	30
Argumentation-Based Logic for Ethical Decision Making (<i>Sofia Almpani, Petros Stefaneas, Panayiotis Frangos</i>).....	46
Non-Monotonic Reasoning in Medieval Theology: Problems and Assumptions (<i>Marcin Trepczyński</i>).....	53
Public Theology Facing a Planet in Turmoil (<i>Ted Peters, Konrad Szocik</i>).....	67
Intellectual and Ethical Virtues in the Situation of War (<i>Vojko Strahovnik, Andrew Schumann</i>).....	70

Trends in Argumentation Logic

Sofia Almpani

National Technical University of Athens
9 Iroon Polytexneiou str.
15780 Zografou
Attiki, Greece

e-mail: salmpani@mail.ntua.gr
<https://orcid.org/0000-0003-4823-479X>

Elena Lisanyuk

St Petersburg State University
7-9, Universitetskaya nab.,
St Petersburg 199034, Russian Federation;
National Research University Higher School of Economics
21/4 Staraya Basmannaya St.,
Moscow 105066,
Russian Federation

e-mail: e.lisanuk@spbu.ru
<https://orcid.org/000-0003-0135-4583>

Andrew Schumann

University of Information Technology
and Management in Rzeszow,
Sucharskiego 2 Street
35-225 Rzeszow, Poland

e-mail: aschumann@wsiz.edu.pl
<https://orcid.org/0000-0002-9944-8627>

Abstract:

In this paper, we introduce the subject of the special issue *Trends in Argumentation Logic*. Here we mainly describe two approaches to argumentation logic with explicating monotonic and non-monotonic, or defeasible, reasoning and explain the role of artificial intelligence in applying argumentation logic. Then we give a short overview of the papers contributed to the special issue.

Keywords: Aristotle, artificial intelligence, argumentation logic, monotonic reasoning, defeasible reasoning.

Argumentation logic is a formalized description of the methods in which humans reason and argue about their claims with the help of arguments for justifying and persuading [5]. In recent years, there has been an increasing interest in applying logic to study and address real-life decision-making procedures in the area of argumentation. With the development of next generation argumentation, the interplay between argumentation approaches and logic is gaining momentum.

Argumentation as an especial theory has its roots in the time of the ancient Greek philosophers, Aristotle and classical rhetoric, and has come a long way all these years with the models and techniques that have been developed so far and still are in a process of rapid evolution. Now, there are two main logical approaches to reasoning in argumentation, presenting its monotonicity and non-monotonicity. Let us remember that monotonicity holds true in any standard symbolic logic and means that if $\Sigma \vdash \phi$, then also $\Sigma \cup \Sigma' \vdash \phi$, where the sign \vdash denotes a deduction from premises. But we can introduce a new deduction relation \sim between premises and conclusions which is not monotonic: if $\Sigma \sim \phi$, then $\Sigma \cup \Sigma' \sim \phi$ does not hold true [6]. Hence, adding new premises does not expand, but restricts the set of our conclusions. Such reasoning is called non-monotonic. One of the cases of non-monotonicity appears, when we have a statement ψ that is inferred from Σ but in turn it cannot be a premise for other conclusions. Then we have $\Sigma \sim \psi$. But if $\Sigma \sim \phi$ holds true for some ϕ , from this it follows that $\Sigma \cup \psi \sim \phi$ does not hold true. This is exactly the case when we cannot build chains of inference. Aristotle demonstrated two logical techniques in argumentation at once: both monotonic (his *Prior Analytics* and *Posterior Analytics*) and non-monotonic (his *Topics* and *On Sophistical Refutations*).

Nevertheless, it has been recently learnt that some foundations of argumentation theory as a practice were laid down at the time of Ur III (about 1.5 thousand years earlier than Aristotle lived). The Sumerians and then the Akkadians were the first who proposed a monotony technique in argumentation [12], [13]. Mainly, they applied two inference rules: *modus ponens* and *modus tollens*, and then accepted inference chains.

Practical applications of argumentation seem to be suitable for dealing with problems that require expert reasoning with strict specifications and mostly with accepting the monotonicity of reasoning. First of all, it is presented by rule-based models in which we can obtain some forward and backward inference chains in accordance with some argumentation standards provided by domain experts. This is very applicable now, e.g., in medicine [2]. Another very significant area of studying argumentation is presented by legal norms and an applicability of logic and argumentation to them [3], [7], [8], [9], [11].

At the end of 20th century, a new trend in argumentation, called argumentation logic, emerged, driven by three notable and independent developments: updating and amending knowledge data bases [1], defeasible reasoning as an application of the non-monotonic logic [10] and artificial intelligence. Argumentation logic develops the idea that rational agents accept arguments as convincing, not just because their conclusions are justified by inferring them from their premises, but rather because those arguments are able to support their conclusions against counterarguments, supporting the opposite conclusion. Argumentation logic views disputes as sets of arguments that are taken as its atoms abstractly of their premise-conclusion logical form and

ordered on graphs by a binary attack relation symbolizing critical argumentation. It employs the non-monotonic skeptical or credible semantic algorithms and the notion of fixed point for interpreting different kinds of inferential relations within these abstract argumentative frameworks [4]. Contemporary developments in argumentation logic suggest using labelling or preference-based semantics, as well as considering deductive formalisms based on abstract argumentation frameworks.

Recently, logic-based systems for examining and assessing arguments have been broadly applied, generating various formal methods for argumentation-based reasoning which is not only monotonic. Moreover, argumentation logic has become a key research topic within Artificial Intelligence for formalizing both monotonic and non-monotonic human reasoning. It involves the examination of those procedures for the development and exchange of arguments, where arguments are efforts to persuade someone by providing reasons for accepting a conclusion or claim as valid. Thus, theories and approaches implementing argumentation logic can be found over a wide range of cases in related disciplines such as linguistics, sociology, law, ethics, computer science and others. This trend prompts researchers to pay attention to potential new related areas, based on either their theoretical foundations or their effective applications.

This special issue collects newly developed works from logic and argumentation, to stimulate possible outcomes from their interplay. This volume includes the selection of 6 papers from 14 submissions accepted to Argumentation Logic Workshop of the 7th World Congress and School on Universal Logic (UNILOG 2022) at Orthodox Academy of Crete (Greece). Among the experts who presented their research at the workshop but do not become the authors of the papers of the special issue, there are Katie Atkinson (University of Liverpool, UK), covering the topic of *Explainable AI for Legal Applications using Computational Models of Argument*, and Ivan Mikirtumov (St Petersburg University, Russia), covering the topic of *Processing: Metaphor and Model for an Interpretation of Arguments*. The selected papers of the issue discuss theoretical foundations in argumentation logic as well as challenges and real-world cases. Each submission underwent a peer-review process by at least two independent expert reviewers. A short overview of the six papers accepted for publication is presented below.

The paper *Argumentation: Reasoning Universalis* contributed by Antonis Kakas (Department of Computer Science, University of Cyprus) is a theoretical work in the area of computational argumentation. It presents how argumentation can form a universal basis for reasoning, capturing both informal and formal logical reasoning. It highlights why argumentation reasoning is proper for the logical foundations of AI, drawing an analogy between Aristotle's study of argumentation and computational argumentation in AI.

Dimitra Serakioti (Democritus University of Thrace, School of Educational Sciences, Greece) and Petros Stefaneas (School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Greece), in their joint paper *Ambiguity in Argumentation: The Impact of Contextual Factors on Semantic Interpretation*, apply Halliday's Systemic Functional Grammar to present how interpretation can reconstruct the meaning of a dialogue and how we analyze ambiguities by bringing together two important strands of research: argumentation theory and text linguistics.

Vladimir A. Stepanov (Moscow, Russia) wrote the paper *Dynamic Approximation of Self-Referential Sentences* in which he proposes a new 6-value lattice of a non-classical logic via dynamic approximation for modeling of self-referential sentences. It handles those sentences as infinite iterations of self-predications and determines their truth-values with truth tables. The obtained new dual truth functions elegantly obey De Morgan laws.

The submission *Determining Argumentative Dispute Resolution Reveals Deep Disagreement over Harassment Issue (a Case-study of a Discussion in the Russian Parliament)* by Elena Lisanyuk (St Petersburg University, National Research University Higher School of Economics in Moscow, Russia) presents a methodology that combines concepts from argumentation logic, new dialectics, and logical-cognitive approach to argumentation and

aggregated formal and informal tools of analysis to develop an algorithm for determining dispute resolution.

In the research *Argumentation-based Logic for Ethical Decision Making*, Sofia Almpani (School of Electrical and Computing Engineering, National Technical University of Athens, Greece), Petros Stefaneas (School of Applied Mathematics, National Technical University of Athens, Greece), and Panayiotis Frangos (National Technical University of Athens, Greece) propose to define context-based scenarios for formalizing ethical reasoning on how far something can be accepted or rejected according to appropriate ethical rules creating a tool for verifying whether agent's decisions are ethically justified.

The paper *Non-Monotonic Reasoning in Medieval Theology: Problems and Assumptions* by Marcin Trepczyński (University of Warsaw, Poland) presents and analyzes cases of non-monotonic reasoning in medieval theological texts and outlines problems connected with identification of non-monotonicity specific for theology.

Additionally, in this special issue we publish two interviews: the *Public Theology Facing a Planet in Turmoil* given by Ted Peters (emeritus professor at the Graduate Theological Union, where he co-edits the journal, *Theology and Science*, on behalf of the Center for Theology and the Natural Sciences, in Berkeley, California, USA) to Konrad Szocik and the *Intellectual and Ethical Virtues in the Situation of War* given by Vojko Strahovnik (Department Chair and Associate Professor at the Department of Philosophy and Research Fellow in Philosophy at the Faculty of Theology, University of Ljubljana) to Andrew Schumann.

The editors would like to thank all the authors that have contributed to this special issue.

References

1. Alchourrón, C., P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, *Journal of Symbolic Logic* 50, 1985, pp. 510-530.
2. Almpani, S., Y. Kiouvrekis, P. Stefaneas, P. Frangos, Computational Argumentation for Medical Device Regulatory Classification, *International Journal on Artificial Intelligence Tools* 31, 2022, DOI: 10.1142/S0218213022500051
3. Barbashina, E., E. Lisanyuk, Leon Petrażycki on Norms and Their Logical Study, *Studia Humana* 7: 4, 2018, pp. 30-38.
4. Dung, P.M., On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n -person games, *Artificial Intelligence* 77, 1995, pp. 321-357.
5. Eemeren van, F.H., R. Grootendorst, *A Systematic Theory of Argumentation. The Pragmadiialectical Approach*, Cambridge University Press, 2004.
6. Gabbay, D. M., K. Schlechta, *A New Perspective on Nonmonotonic Logics*, Springer International Publishing AG, 2016.
7. Kirillova, N., E. Lisanyuk, Truth and legal argumentation in Fyodor Dostoevsky's, The Karamazov Brothers, *Tomsk State University Journal of Philosophy, Sociology and Political Science* 48, 2019, pp. 193-204.
8. Lisanyuk, E., Argumentation, Pavilionis's meaning continuum and the Kitchen Debate, *Problemos* 88, 2015, pp. 95-113.
9. Lisanyuk E. Logical theories of rational behavior, law and argumentation [Логічні теорії раціональної поведінки, право і аргументація], *Philosophy of law and contemporary theory of law* [Філософія права і загальна теорія права] 1–2, 2014, pp. 148-159, in Ukrainian.
10. Pollock, J. L., Defeasible Reasoning and Degrees of Justification, *Argument and Computation* 1 (1), 2010, pp. 7-22.
11. Prakken, H., Formalising debates about law-making proposals as practical reasoning, [in:] M. Araszkievicz, K. Pleszka (eds.), *Logic in the Theory and Practice of Lawmaking*. Legisprudence Library Series, Springer, 2015.

12. Schumann, A. Did the Neo-Babylonians construct a Symbolic Logic for legal proceedings? *Journal of Applied Logics – IfCoLog Journal of Logics and Their Applications* 6: 1, 2019, pp. 31-82.
13. Schumann, A. *Archaeology of Logic*, Taylor & Francis Group, 2023.

Argumentation: Reasoning Universalis

Antonis Kakas

University of Cyprus
1 University Avenue
2109 Aglantzia, Cyprus

e-mail: antonis@cs.ucy.ac.cy
<https://orcid.org/0000-0001-6773-3944>

Abstract:

Can argumentation form the basis for any form of reasoning, informal or formal logical reasoning? We examine this question from the particular perspective of the recent developments in logic-based Artificial Intelligence (AI). We propose that argumentation provides the wider framework encompassing uniformly all reasoning, with strict or formal logical reasoning being a special boundary case. We also attempt to link this unifying role of argumentation with Aristotle's original investigation of methods and formalisms for the systematic study of human reasoning.

Keywords: Reasoning, argumentation, Aristotle.

1. Introduction

Logic is traditionally separated into two forms: **Formal Logic** at the foundations of Mathematics and Science and **Informal Logic** as the study of human reasoning at large. These two forms of logic are generally considered to be very different. Yet they are both concerned with understanding the nature of human thought and, in fact, they share the same roots in Aristotle's work.¹ In this work we are interested in the question of whether formal and informal logic can be placed under a single framework and, if so, to understand their distinguishing features. In other words, we are interested in finding a universal form of reasoning that would be able to capture both informal and formal reasoning. In answering this question we will also attempt to link our proposal to the origins of the study of reasoning in Aristotle and how Aristotle's study can help in forming a unified view of reasoning. In a sense, the distinction of the two forms of logic seems to have evolved with the development of these over the last few centuries, especially with the development of Classical Logic and its foundational role in Mathematics and Science, drawing them more and more apart.²

In order to be concrete we will consider that Formal Logic is represented by Classical Logic or simply Propositional Logic. For the case of Informal Logic it is more difficult to select a representative example. Informal Logic relates to reasoning by humans at large in everyday tasks but also to critical thinking, rhetoric and debate. Whatever form we consider it is important to realize that in the study of

Informal Logic, within the humanities and particularly in Philosophy, scholars have essentially been equating informal reasoning with **Argumentation**. The entry on Informal Logic in the Stanford Encyclopedia of Philosophy (<https://plato.stanford.edu/entries/logic-informal/>) presents the overriding goal of informal logic as the task of providing a general account of argument as the basis of systems of informal logic. It then continues to state:

In the pursuit of its goals, informal logic addresses topics which include, to take only a few examples, the nature and definition of argument, criteria for argument evaluation, argumentation schemes, fallacies, notions of validity, the rhetorical and dialectical aspects of arguing, argument diagramming (“mapping”), cognitive biases, the history of argument analysis, artificial intelligence (AI), and the varying norms and rules that govern argumentative practices in different kinds of contexts.

One field which studies Informal Logic, in the sense of human reasoning at large, is that of Artificial Intelligence (AI), where the aim to formalize and automate common sense reasoning was set as an early foundational problem. This resulted in the search for and development of a plethora of new logics for AI, called **non-monotonic logics**, starting with the logic of Circumstantiation for formalizing the Situation Calculus, a system for common sense reasoning about the effects of actions and the change they bring about [15]. These new logics aimed to capture the non-monotonicity feature of human inference recognizing that it should be possible to abandon, in contrast to the monotonic inference of formal classical logic, earlier inferences in the face of new relevant information. Non-monotonicity was needed to render the inference flexible, in the same manner as human do when drawing inferences, to missing or ambiguous information and tolerant to (apparently) contradictory information.

Nevertheless, these new non-monotonic logics were developed based on the same formal and strict underpinnings of Classical Logic making it difficult to deliver on their promise of “AI systems with common sense” and “human-like natural intelligence”. Then in the 1990s, it was shown (see e.g. [1]) that using argumentation it was possible to reformulate (and in some cases extend) most, if not all, known logical frameworks of non-monotonic reasoning in AI. This AI approach to argumentation, sometimes referred to as **Computational Argumentation**, was motivated and to some extent grounded on earlier foundational work [26], [20], [21] on human argumentation in Philosophy and Cognitive Science. The result of reconciling non-monotonic logics through argumentation resulted in a strong focus on Computational Argumentation as a way of capturing human reasoning in AI along the same frame of interest as that of Informal Logic. For example, argumentation can provide a principled approach to knowledge representation and reasoning about actions and change [18], [8]. Based on this we can build computational models of narrative comprehension akin to the way humans perform this task [2].

Similarly, following recent work in the Psychology of Reasoning that strongly supports the link of argumentation to human reasoning (e.g. [16], [17]) we can synthesize a framework of computational argumentation informed by cognitive principles to obtain a framework, called **Cognitive (Machine) Argumentation**, as a suitable framework to model human reasoning in its various forms. This framework has been shown to capture well the human empirical data from several different experiments that are traditionally used in Cognitive Science to evaluate cognitive models of human reasoning. These empirical evaluation domains include “Syllogistic Reasoning” with experiments on how humans reason on the original Aristotelian syllogisms, the “Selection Task” where humans are tested on the way they reason about conditionals and the “Suppression Task” where the non-monotonic nature of human reasoning is observed [23], [24], [25]. Cognitive argumentation accounts for the empirical data in these domains in a cognitively adequate way that also reflects well the variation of human reasoning across the population.

We will therefore accept that human or informal reasoning is a matter of argumentation and ask whether argumentation can also encompass formal logic. Hence we will be interested in whether

argumentation can be given some formal structure and how this might also cover formal classical deductive reasoning. We will argue that this is possible so that both informal but also formal logic can be captured uniformly within the same formal structure of argumentation. Argumentation will thus form the wider notion of reasoning, **Reasoning Universalis**, encompassing all forms of reasoning with strict or formal logical reasoning being a special boundary case.

The next section, Section 2, reviews the formalization of argumentation as a reasoning system and gives its basic AI computational model. Section 3 presents how the formal propositional logic of deduction is captured within the above formalization of argumentation. It then discusses how this Logic of Argumentation extends smoothly beyond classical formal reasoning with premises which can be inconsistent under classical logic but admissible in an informal reasoning setting. Section 4, attempts to show a connection between this modern formalization of argumentation and the work of Aristotle in the books of Topics. Section 5 concludes with a brief discussion of the possible relevance of Aristotle’s work to today’s logic-based AI.

2. Formal Argumentation

Argumentation is a process of considering the alternative positions that we can take on some matter with the aim to justify or refute a standpoint on the matter. It can take place socially within a group of entities in a debate where entities argue for different standpoints, or within a single entity where the entity contemplates or reasons internally about the various standpoints on the matter, in order to decide on and self justify its own stance on the matter.

Argumentation has the general form of a **dialectical** process of (i) starting with some argument(s) directly supporting some desired standpoint or conclusion, then (ii) considering various counter-arguments against the initial argument(s) and (iii) defending against these counter-arguments, typically with the help of other arguments as allies of the initial arguments. The process repeats by considering further counter-arguments against these new allied defending arguments, until we have formed a **coalition of arguments** that “stands well” as a **case** for the standpoint or conclusion of interest.

We therefore have an “argumentation arena” where arguments attack and defend against each other in order to support their claims. This arena of argumentation can be captured by a formal **argumentation framework** which in an abstract form can be simply given as a tuple, $\langle \mathcal{A}rgs, \mathcal{A}TT \rangle$, where $\mathcal{A}rgs$ is a set of arguments and $\mathcal{A}TT$ is an attack (typically non-symmetric) relation between the arguments in $\mathcal{A}rgs$. Note that in this abstract formulation of argumentation we have no information on how the arguments and the attacking relation between them arises. In other words, at the most abstract level the only essential elements for the dialectic process and the result of argumentation are these two notions of the existence of arguments and the attacks between them. In practice though, the consideration of constructing the arguments and the attacks between them cannot be avoided and it is an important element of the whole process of argumentation. For example, we note that in this minimal abstract formalization of argumentation frameworks, the attack relation, $\mathcal{A}TT$, serves both the purpose of identifying conflicts between arguments but also to specify the relative strength between them so that we can identify whether an argument is strong enough to defend against another one by attacking it back under $\mathcal{A}TT$.

Given an argumentation framework we can then formalize, through some normative condition, the notion that a subset of arguments “stands well” as a case of arguments. In fact, the dialectical process of argumentation indicates how to give such a suitable semantics to formal argumentation. The dialectic argumentation semantics is defined via a relation $ACC(\Delta, \Delta_0)$ between any two sets of arguments Δ, Δ_0 . This relation specifies the **acceptability** of the set of arguments Δ under the context where the set Δ_0 of arguments is considered as given and so a-priori acceptable. Informally, the relative acceptability between sets of arguments captured by “ $ACC(\Delta, \Delta_0)$ ” is defined recursively to hold when the argument

set Δ can render all its attacking (or counter-arguments) non-acceptable in the context of accepting Δ_0 together with Δ . This natural and intuitive description of the acceptability relation between sets of arguments can be formally defined as the least-fixed point of an associated formal operator satisfying the following (see [14] for the technical details):

$ACC(\Delta, \Delta_0)$ holds, iff $\Delta \subseteq \Delta_0$, or,
 for any A such that $(A, \Delta) \in \mathcal{ATT}$ (i.e. A attacks Δ),
 $A \subseteq \Delta_0 \cup \Delta$, and there exists D such that:
 $(D, A) \in \mathcal{ATT}$ (i.e. D counter-attacks or defends against A), and
 $ACC(D, \Delta_0 \cup \Delta \cup A)$ holds.

Then the **acceptable or case** subsets of arguments are defined as those that are acceptable in the context of the empty set, i.e. the subsets Δ for which $ACC(\Delta, \emptyset)$ holds in the least fixed point of this relation.

The technical details of the definition of this semantics of argumentation are not important for this paper. The technical form of the above definition is included here only for the reader to have some appreciation that at the abstract level, argumentation can be given a strict and well defined meaning and that it is this meaning that permeates through all forms of reasoning.

Informally, the above least-fixed point definition says that for any counter-argument to Δ we need to have a defending argument against this counter-argument which is acceptable within the context of Δ . Indeed, acceptable subsets of arguments can be computed following the fixed point definition of acceptability as illustrated here by Figure 1.

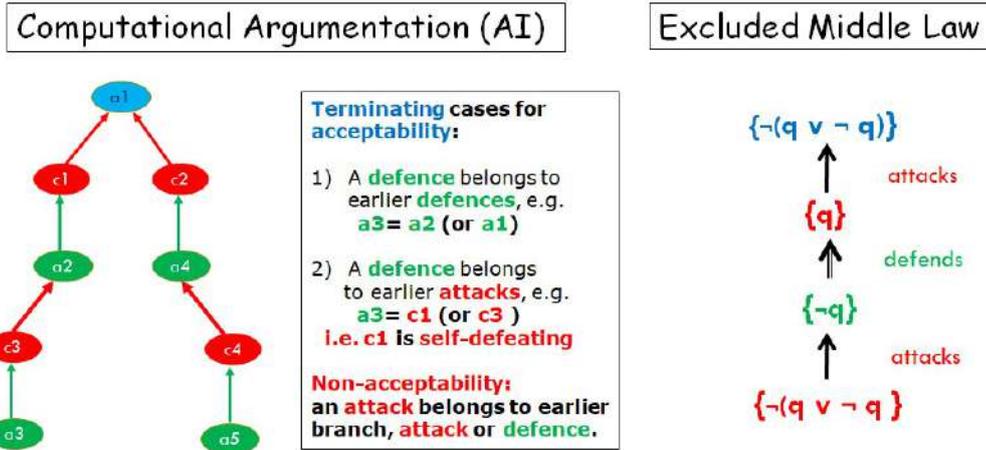


Figure 1: Dialectic Acceptability/Non-Acceptability of Arguments. The left part shows the general abstract case. The right part shows a concrete case in propositional logic.

This also helps us understand the formal semantics of argumentation as we can connect these trees to the dialectic process of arguing for and against a position. The left hand part of this figure, under the heading of Computational Argumentation in AI, shows this dialectic process to construct acceptable arguments in terms of labelled trees. One starts with an argument at the root of these computational trees (such as a_1 in the figure) supporting a position of interest. We then consider counter-arguments to the root argument, as indicated by the red arrows from arguments c_1 and c_2 in the figure, and then for each one of these a corresponding defending argument. Defending arguments are shown in green attacking with green arrows the attacking arguments, e.g. in the figure a_2 defends against c_1 and a_4 defends against c_2 . The process then repeats by considering new attacking arguments against each one of the newly

introduced defending arguments and in turn considering defending arguments against each one of the new attacking arguments.

The purpose of these trees is to help us construct an acceptable subset of arguments containing the argument at the root of the tree (which would typically support a desired conclusion). Red nodes in the tree indicate attacking counter-arguments whereas green nodes indicate defending arguments. The collection of defending arguments together with the root argument form the “pro case” while the collection of the attacking arguments form the “con case”.

The termination conditions for the acceptability (respectively the non-acceptability) of the root argument are shown in the box of the left part of Figure 1. They show that the process terminates when a defense (respectively an attack) node belongs to the union of arguments in the branch of the tree above the defense (respectively the attack) argument³. When all branches of the tree terminate with a defense argument we have that the root argument is acceptable, whereas if one branch terminates with an attack argument the root argument is non-acceptable. This simple process can then be automated and used in real-life applications of AI [11].

3. Formal Logic as a Case of Argumentation

We can now present how this semantics of argumentation can be used to reformulate the formal logical reasoning of classical Propositional Logic in terms of argumentation. Indeed, it is possible to define **Argumentation Logic (AL)** [12] as a realization of the above abstract argumentation framework and show that this captures classical deductive reasoning. The arguments in AL are made up of sets of propositional formulae and the attack and defense relations are defined through the incompatibility between formulae and their negation. Arguments that are formed solely from formulae within a given theory T under which we are reasoning are stronger than arguments which are not of this form, i.e. arguments that contain at least one formula outside the premises T .

It is then possible to show that Argumentation Logic is logically equivalent to classical deductive reasoning whenever the given theory T that we are reasoning from is classically consistent [10]. This means that when the (propositional) premises T are classically consistent then formal logical entailment, in the classical sense of truth in all models, coincides with **sceptical entailment** in Argumentation Logic, defined as follows:

A formula ϕ is a **sceptical conclusion** in AL if and only if the argument $a = \{\phi\}$ is acceptable and the opposite argument $a' = \{\neg\phi\}$ is non-acceptable.

The non-acceptability of the argument a' of the negation of the formulae means that the negation is not possible under any circumstance and hence the positive conclusion is an absolute winner, i.e. necessarily follows.

Non-surprisingly, as in most works that aim to bring formal logic closer to human reasoning, e.g. the early example of Intuitionistic Logic [19], the central element for this result of reformulating formal logical reasoning in terms of argumentation lies in the way that Reductio ad Absurdum is captured within the framework of argumentation. This is done by identifying structurally **self-defeating** (or fallacious) arguments and relating these to indirect logical proofs, i.e. proofs requiring *Reductio ad Absurdum*, within Propositional Logic.

Informally, a self-defeating argument, S , is one that “turns on itself” by rendering one of its attacking arguments acceptable in its own context of S . This means that the self-defeating argument renders the arguments that it needs for its defence, against some attacking counter-argument, non-acceptable. More formally, we can define a self-defeating argument S as one for which there exists a counter-argument A such that $\neg ACC(A, \emptyset)$ and $ACC(A, S)$ hold. So, although the attack A is in

general (i.e. when we do not take any argument to be as given) non-acceptable, under S , this attack is rendered acceptable. Hence S brings about its own defeat and non-acceptability. The simplest example of a self-defeating argument is one that attacks itself, since in its own context its self-attack is acceptable.

For a more elaborate example of a self-defeating argument let us consider an example from the argumentation-based reformulation of formal logic, related to how we can derive the excluded middle law in Argumentation Logic. This is illustrated in the right part of Figure 2 where we see that the negation of the law, i.e. $\neg(q \vee \neg q)$, is shown to be non-acceptable, as the computational tree has a branch that terminates with an attacking argument. The tree shows that the root argument is attacked by the formula q , as from q we can directly derive $q \vee \neg q$. This attack by q can only be defended by taking on the opposing position of $\neg q$. But this defense is attacked by the root formula of $\neg(q \vee \neg q)$ since we can directly derive $q \vee \neg q$ from $\neg q$. We therefore have that an attack belongs to the branch above it and so the argument $\neg(q \vee \neg q)$ renders its required defense non-acceptable and thus, indirectly, it also renders itself non-acceptable.

Posing a hypothesis as a premise in a **Reductio ad Absurdum** proof corresponds to considering a context in which the hypothesis as an argument is accepted. Then the hypothesis leading to an inconsistency corresponds to the dialectic argumentation process leading to the non-acceptability of a (necessary) defending argument in the context of the posited argument. This correspondence is exact when the propositional theory of given premises is classically consistent in which case the non-acceptability of a formula argument also means the acceptability of the complement of the formulae, in the same way that Reductio ad Absurdum is used to derive the complement of the posited hypothesis. For the general case where the given theory under which we are reasoning is inconsistent then this latter step does not hold and we can have that both a formulae and its complement are non-acceptable. This signifies that we cannot have a position on such formulae. Nevertheless, this does not mean that the whole reasoning of Argumentation Logic trivializes but only that for some isolated formulae we are completely agnostic.

In summary, classical formal reasoning is captured as a special case of argumentation where a logical conclusion emerges as the result of contemplating arguments for and against the conclusion. Argumentation Logic is constructed by adopting a set of direct proof rules as basic argument schemes together with the recognition of self-defeating arguments to cover the indirect proofs through Reductio ad Absurdum. Then the acceptability semantics of argumentation and the sceptical form of entailment under this semantics realized in this concrete framework is equivalent to classical deductive reasoning.

3.1. Beyond Classical Reasoning: Back to Informal logic

The above correspondence between classical Propositional Logic and Argumentation Logic shows that classical truth models correspond to cases of acceptable subset of formulae, but this breaks down when the given premises are classically inconsistent. Acceptable cases of Argumentation Logic continue to exist and the logic does not trivialize. Hence, Argumentation Logic with its paraconsistent⁴ form of argumentative reasoning can be understood as a smooth conservative extension of strict classical logical reasoning, in cases where indeed the given premise information is contradictory [13]. One interesting consequence of this is that (some of) logical paradoxes are dissolved. For example, for the Barber Paradox we have that in AL the two complementary sentences that “the Barber shaves himself” and that “the Barber does not shave himself” are both non-acceptable, showing that the logic is agnostic about who shaves the barber. But, in contrast to classical logic this does not affect the logical reasoning for any other person in the universe of discourse.

Generally, in such cases of inconsistent premises we can build arguments from subsets of the given premises that attack each other and therefore form alternatives. In addition, we may have information on the relative strength between these premises that would then feed into the definition of

attack (and defense) between such arguments that are grounded on the given premises. Such information comes from the content of the premises with which we are reasoning, i.e. it is specific to the domain about which we are reasoning. This is typical in the realm of informal reasoning e.g. within common sense reasoning, where general or individual human biases give preference to some statements over others. Furthermore, the whole such reasoning is context sensitive as these preferences may change from one context to another. Hence, with informal reasoning, although this is captured under the same framework of argumentation as formal reasoning, the various constructs of argument schemes, attacks and defenses depend on the content of arguments and the dynamically changing environment in which the reasoning takes place.

Let us illustrate this by a simple example of text comprehension. Consider the following start of a piece of text that we reading:

- Mary was very busy at the office.
- Her phone rang.

Did Mary want to answer the phone? An argumentation process to answer this question in favour of not wanting to answer would require to construct an argument supporting this conclusion and acceptably defending it against its counter-arguments. One such argument might use as its premises the common sense knowledge that “normally, when people are busy they do not want to answer their phone”. Grounding these with the information that Mary was busy we get an argument supporting the conclusion that Mary does not want to answer the phone. Let us call this the “Busy Argument (BA)”. Another argument that would be present in our mind is a general one based on the common sense knowledge that “normally, people want to answer their phones”. Let us call this the “General Argument (GA)”. These two arguments are in conflict and would therefore form counter-arguments for each other. In general we may consider them to be of equal strength and therefore formally they attack each other as shown on the left of Figure 2. This figure shows pictorially how this argumentation based interpretation of the text is captured within a formal argumentation framework $\langle Args, ATT \rangle$ (Nodes in the picture show the arguments in $Args$ and the red arrows show the ATT relation between these arguments).

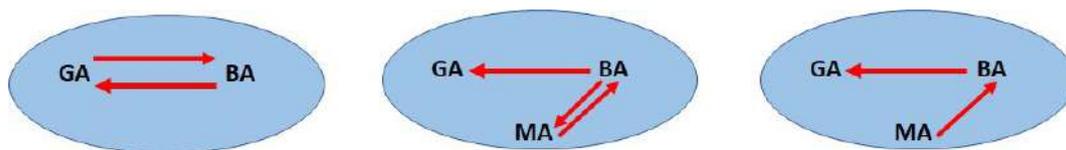


Figure 2: Argumentation Frameworks corresponding to the “phone story” text

But the author by including the qualification of “very” awakens, perhaps in the minds of only some human readers, a preference of the “busy argument” over the “general argument”. Then the “general argument” as a weaker argument would not attack, within the formal argumentation framework, the “busy argument” and so the corresponding argumentation framework would now not include the arrow from GA to PA . This means that the “busy argument” can defend against the counter-argument but not vice versa and hence we can draw the skeptical conclusion of “not wanting to answer the phone”.

Let us assume that the story continues as follows:

- Mary was very busy at the office.
- Her phone rang.
- It was her mother phoning.

Now a new argument, let us call it the “Mother argument (MA)”, enters the argumentation arena, based on the common sense knowledge that “normally, people want to communicate with their mothers”. In general, the bias giving preference to this argument may not be strong enough to overcome the earlier “busy argument” supporting the contrary position of not wanting to answer the phone. The corresponding formal argumentation framework is shown in the middle of Figure 2 where we have assumed that the two arguments of *PA* and *MA* are of equal strength and thus they attack each other. In such a case even if *PA* is considered stronger than *GA* (as shown in the middle of the figure) the absolute sceptical conclusion of “not wanting to answer the phone” is lost and this is now only a plausible or credulous conclusion as we have acceptable subsets of arguments for both this conclusion and its converse of “wanting to answer the phone”, namely the set $\{PA\}$ for not wanting to answer and the set $\{GA, MA\}$ for wanting to answer. The argument *MA* comes to defend *GA* against the strong attack against this by *PA*.

Suppose that the text continues with the following two sentences:

- Mary’s mother fell ill last week.
- She was still (very) ill in the hospital.

The author has now revealed more details of the context that render the “Mother argument” clearly stronger than the “busy argument”. In the corresponding argumentation framework, as shown on the right side of Figure 2, we now only have an attack from *MA* to *PA*. In this, the argument *PA* is not acceptable any more and hence “wanting to answer the phone” becomes a sceptical conclusion.

Using this argumentation-based text interpretation it is possible to construct, along the lines described above, what in Cognitive Psychology is called the **mental or comprehension model** [9] and understand the process of revision of this model as the text unfolds through the non-monotonicity of reasoning through argumentation [4], [3].⁵ It is interesting to note that once we lock ourselves into a specific comprehension model (at any particular point in the text) the conclusions that this model contains can be seen as formal logical conclusions of the explicit information in the story and the world knowledge that we have used to construct the arguments and the model. In other words, the informal logic dialectic argumentation that we use to construct the comprehension model is one and the same as the formal dialectic argumentation process when projected down on the specifics used for constructing the model. Hence, the informal logic of comprehension has a formal logical interpretation under the strict form of argumentation for capturing formal logic, as we have described in the previous sections. Argumentation thus glues together informal and formal logic in both directions.

4. Aristotle: The origins of Systems of Reasoning

We will now briefly look into Aristotle’s work on dialectic argument and study this from a contemporary argumentation perspective. Specifically, we will examine the resemblance between the basic acceptability semantics, that as we have argued above unifies informal and formal reasoning, with the method of Aristotle for dialectic argumentation found in the books of *Topics*.

In these books Aristotle considers the wider context of what today we associate with informal reasoning and laid argumentation as the foundational element of his investigation. Aristotle states from the very start that the purpose of *Topics* (100a18-20) is:

To discover a method by which we shall be able to reason from generally accepted opinions about any problem set before us and shall ourselves, when sustaining an argument, avoid saying anything **self-contradictory** (copied from Rigotti and Greco [7]).

His study of dialectic argument is extensive and quite thorough in an attempt to provide a pragmatically

effective method of applying argumentation to support a position or a claim. He categorizes the different possible positions in terms of four types of “predicables” and goes into great length to give, for each different type of predicable, elaborate prescriptions (topoi) or strategies of how to go about supporting, attacking and defending each particular type of position. From a contemporary point of view these topoi can be linked to the notion of *argument schemes* [27], [28] that associate premises to a position or to the contrary of a position, together with the pragmatics or heuristics to follow when carrying out the process of argumentation, as for example in the pragma-dialectal approach to argumentation in [6].

It is prudent to note that irrespective of the particular details of each topos the overall general condition on the process of dialectic argumentation is to **avoid a contradiction** on “our side” or to arrive at a contradiction on the “opposing side”. This is the only *normative condition* on the process as we can clearly see in the above statement of Aristotle. It oversees the process as a requirement with which the process needs to be compliant. But its role is not merely that of a passive checker on the process. In some strong sense it actively drives the process of dialectic argumentation.

At the very general level the strategy of dialectic argumentation in Aristotle is to bring the opposite view into a situation which is unacceptable because it is self-contradictory. Aristotle describes how this strategy can be executed through a process between a *Questioner* and an *Answerer*. This process can be understood as a semi-formal computational structure consisting of three stages:

(a) **Opening:** The Questioner presents a statement to which the Answerer can reply either yes or no. The overall aim of the Questioner is to force the Answerer to accept that his answer is self-contradictory and thus not reasonable. (b) **Interrogation:** The Questioner introduces questions to the Answerer to establish beliefs that the Answerer holds. The aim of the Questioner in this stage is to gather such beliefs from the Answerer that would allow him to build a strong argument against the Answerer’s claim. (c) **Conclusion:** Once the Questioner has all the information s/he needs s/he reveals to the Answerer the counter argument, which s/he builds through a *sylogism* based on premises that the Answerer has accepted. The fact that this is build through a syllogism means that this is quite a strong argument and cannot be dismissed. Hence the Answerer has no option than to accept that his initial position is in contradiction with his other beliefs, i.e. his case is self-contradictory and therefore defeated by the dialectic process of argumentation.

As mentioned above, in this adversarial process, the goal for the Answerer is to prevent the Questioner from succeeding by reasonably rejecting the premises that would lead to self-contradiction. The difficulty for the Answerer lies in realizing the counter-argument that the Questioner has in mind to build so that s/he can be careful on the beliefs he accepts during the second interrogation stage.

We can then observe a resemblance between this method of Aristotle for dialectic argumentation and the notion of acceptability and non-acceptability of arguments that we have presented above as the unifying foundation of contemporary informal and formal reasoning. The central task in Aristotle to bring the Answerer into a *self-contradiction* is analogous to the identification of self-defeating arguments under the formal notion of acceptability of arguments. Just like the dialectic method of Aristotle concludes with the exposition of a contradiction in the beliefs held by the Answerer, in the same way the computational trees of acceptability (see Figure 2 and termination conditions for non-acceptability) closes with an attacking argument playing also the role of a needed defense argument in the same dialectic branch of the tree, thus rendering the defending argument as self-defeating and non acceptable. Let us illustrate this correspondence through an example, shown in Figure 3.

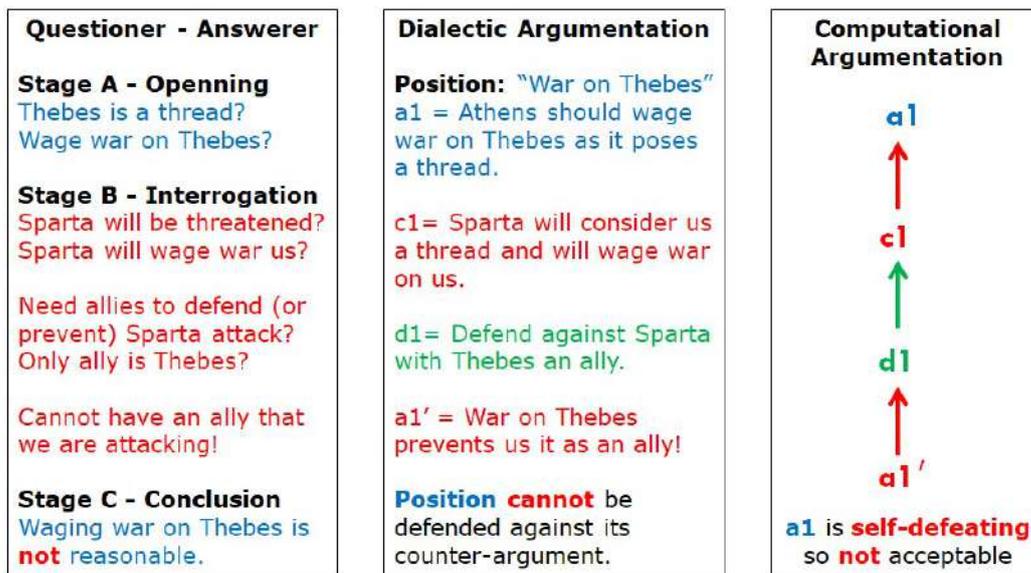


Figure 3: Example of Aristotle’s Dialectic Argument

In the leftmost box of the figure we see the questions asked by the Questioner. We assume that the Answerer has answered “yes” to all these questions. The Questioner can then re-construct an explicit dialectic argumentation process (seen in the middle box of the figure) where the attacking counter-argument of $c1$ is revealed together with the fact that the proposed defense $d1$ against this, i.e. to use “Thebes as an ally”, is in conflict with the original position of the answerer of “waging war on Thebes” and therefore could not form a coalition with the initial argument of $a1$. The rightmost part of the figure shows the abstract computational structure of this argumentation process and how it ends up with the non-acceptability of the initial argument supporting the original position of the answerer⁶. We can thus see the direct correspondence between Aristotle’s dialectic method of argumentation with the computational model of argument trees in modern argumentation.

5. Conclusions: Reasoning in AI

We have presented how argumentation can form a universal basis for reasoning. The single notion of an acceptable (set of) argument(s) as one that can defend against its counter-arguments can uniformly capture informal and formal logical reasoning. Different forms of reasoning are thus attributed to the “intensity” of the argumentation process to consider to a varying degree a complete set of arguments and counter-arguments.

One of the main tasks of today’s AI is to understand, formalize and effectively compute human reasoning. Hence, if we accept the universality of argumentation for reasoning, i.e. that *Reasoning is Argumentation*, then argumentation presents itself as a suitable candidate for the logical foundations of AI. We are then naturally led to *re-enact Aristotle’s study of argumentation* in the *Organon* and particularly in the books of *Topics*. Just like Aristotle studied how to conduct argumentation in an effective way and proposed different *topoi* as guidelines for achieving this we can carry out an analogous study for the effective realization of computational argumentation in AI. To do so we need to consider, as Aristotle did, the dynamic and uncertain nature of the environment in which argumentation takes place where the computational process of argumentation should adapt to new information and in many cases be guided to actively seek new relevant information. There is of course one major difference: Aristotle’s argumentative reasoning was to be carried out by the “machine of the human mind/brain” whereas in AI the machine is a poor artifact of the human mind/brain. Nevertheless, we can draw from the study of

argumentation over the centuries in philosophy, the psychology of reasoning and other disciplines to help us in this task of finding an effective process of reasoning through argumentation. Combining this with the study from a modern perspective of Aristotle's extensive work on the good practice of argumentation, as for example in the recent work of [7], can provide us with valuable insights for the development of AI.

References

1. Bondarenko, A., Dung, P.M., Kowalski, R.A. and Toni, F. An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence* 93, 1997, pp. 63–101.
2. Diakidoy, I.A., Kakas, A.C., Michael, L. and Miller, R. Story comprehension through argumentation. In Simon Parsons, Nir Oren, Chris Reed, and Federico Cerutti, editors, *Computational Models of Argument - Proceedings of COMMA 2014, Scottish Highlands, UK*, volume 266 of *Frontiers in Artificial Intelligence and Applications*, 2014, pp. 31–42, IOS Press.
3. Diakidoy I.A., Michael, L. and Kakas, A.C. Knowledge activation in story comprehension. *The Journal of Cognitive Science*, 18, 2017, pp. 439–471.
4. Diakidoy I. A., Michael, L., Kakas, A.C. and Miller, R. Story Comprehension through Argumentation. In *Proceedings of the 5th International Conference on Computational Models of Argument (COMMA)*, 2014, pp. 31–42, IOS Press.
5. Novaes, C.T. *The Dialogical Roots of Deduction: Historical, Cognitive, and Philosophical Perspectives on Reasoning*, 2020, Cambridge University Press.
6. van Eemeren, F.H. and Grootendorst, R. *A Systematic Theory of Argumentation: The pragma-dialectical approach*, 2003, Cambridge University Press.
7. Greco, S. and Rigotti, E. *Inference in Argumentation: A Topics-Based Approach to Argument Schemes*, 2019, Springer Verlag.
8. Hadjisoteriou, E. and Kakas, A. C. Reasoning about actions and change in argumentation. *Argument Computation*, 6 (3), 2015, pp. 265–291.
9. Johnson-Laird, P. *Mental Models*, 1983, Cambridge University Press.
10. Kakas, A.C., Mancarella, P. and Toni, F. On argumentation logic and propositional logic. *Studia Logica*, 106 (2), 2018, pp. 237–279.
11. Kakas, A.C., Moraitis, P. and Spanoudakis, N.I. *GORGias*: Applying argumentation. *Argument & Computation*, 10 (1), 2019, pp.55–81.
12. Kakas, A.C., Toni, F. and Mancarella, P. Argumentation logic. In Simon Parsons, Nir Oren, Chris Reed, and Federico Cerutti, editors, *Computational Models of Argument - COMMA 2014*, volume 266 of *Frontiers in Artificial Intelligence and Applications*, 2014, pp. 345–356. IOS Press.
13. Kakas, A.C. Informalizing formal logic. *Informal Logic*, 39 (2), 2019, pp.169–204.
14. Kakas, A. C. and Mancarella, P. On the Semantics of Abstract Argumentation. *Journal of Logic and Computation*, 23 (5), 2013, pp.991–1015.
15. McCarthy, J. Programs with common sense. In *Proc. of Teddington Conf. on Mechanization of Thought Processes*, 1959, pp. 75–91.
16. Mercier, H. and Sperber, D. Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34 (2), 2011, pp.57–74.
17. Mercier, H. The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences*, 20 (9), 2016, pp.689 – 700.
18. Michael, L. and Kakas, A.C. Knowledge qualification through argumentation. In Esra Erdem, Fangzhen Lin, and Torsten Schaub, editors, *Logic Programming and Nonmonotonic Reasoning, 10th International Conference, LPNMR 2009, Potsdam, Germany*. Volume 5753 of *Lecture Notes in Computer Science*, 2009, pp. 209–222. Springer.
19. Moschovakis, J. Intuitionistic Logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2021 edition, 2021.

20. Perelman C. and Olbrechts-Tyteca, L. *The new rhetoric. A treatise on argumentation*, 1969, Notre Dame/ London: University of Notre Dame Press.
21. Pollock, J. L. Defeasible reasoning. *Cognitive Science*, 11(4), 1987, pp.81–518.
22. Priest, G., Tanaka K. and Zalta, E.N. Paraconsistent logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy (Spring 2022 Edition)*, 2022.
23. Dietz, E. and Kakas, A.C. Cognitive argumentation for human syllogistic reasoning. *Künstliche Intell.*, 33(3), 2019, pp.229–242.
24. Dietz, E. and Kakas, A.C. Cognitive argumentation and the suppression task. *CoRR*, abs/2002.10149, 2020.
25. Dietz, E. and Kakas, A.C. Cognitive argumentation and the selection task. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43, 2021, pp. 1588-1594, <https://escholarship.org/uc/item/4n0867c7>, 2021.
26. Toulmin, S.E. *The Uses of Argument*, 1958, Cambridge University Press.
27. Walton, D.N. *Argumentation Schemes for Presumptive Reasoning*, 1995, Routledge.
28. Walton, D.N., Reed C. and Macagno, F. *Argumentation Schemes*, 2008, Cambridge University Press.

Notes

1. All statements in this paper relating to Aristotle are to be understood as hypotheses posed by the author in the context of his extremely limited knowledge of Aristotle’s work. They are therefore subject to disproof by any Aristotelian scholar. They are made in an attempt to understand how Aristotle, as the first logician and his general study of systematizing human reasoning, relates to current attempts in AI to formalize and automate human reasoning.
2. The recent book, entitled “The Dialogical roots of Deduction” [5] provides a unique exposition of the evolution of logical reasoning and how its genealogical connection with the process of dialectics survives into today’s various forms of logical reasoning.
3. These conditions complement the base termination conditions of the non-existence of an attacking (respectively defending) argument.
4. It is evident that Argumentation Logic is related to Paraconsistent Logics [22] which similarly consider how we can define forms of reasoning that do not trivialize under inconsistent premises.
5. The construction of a comprehension model depends also on other factors, e.g. that of coherence where only conclusions in the main thread of the story are considered, but these are extra-logical processes outside the scope of this paper.
6. Strictly speaking the attacking argument $a1'$ is not the same as $a1$ but has the same effect of terminating the branch at an attack level. The only way to defend against $a1'$ is either by an argument against its premise of waging war on Thebes or an argument against Thebes being an ally. In either case this new defense will be attacked either by $a1$ or by $a3$ resulting in the non-acceptability of the branch.

Ambiguity in Argumentation: The Impact of Contextual Factors on Semantic Interpretation

Dimitra Serakioti

Democritus University of Thrace,
68100 Nea Chili,
Alexandroupolis, Greece

e-mail: dimser21@hotmail.com
<https://orcid.org/0000-0002-6306-5106>

Petros Stefaneas

National Technical University of Athens,
Zografou Campus 9, Iroon Polytechniou str.,
15772 Zografou, Greece

e-mail: petros@math.ntua.gr
<https://orcid.org/0000-0002-2096-9914>

Abstract:

This article is concerned with the concept of ambiguity in argumentation. Ambiguity in linguistics lies on the coexistence of two possibly interpretations of an utterance, while the role of contextual factors and background/encyclopedic knowledge within a specific society seems to be crucial. From a systemic point of view, Halliday has proposed three main language functions (meta-functions): a) ideational function, b) interpersonal function, c) textual function. Language could reflect speaker's experience of his external and internal world, interpersonal relationships and organization of text, respectively. Lexico-grammatical choices under a micro-level perspective and context (the environment of language) may lead to inconsistent interpretations through semantic or syntactic ambiguities. In philosophy and argumentation logic, strategies of ambiguity have been investigated by Aristotle, since the first sophistic movement. In his *Topics*, *Metaphysics* and *Rhetoric*, has pointed out the notion of “τὸ διπλῶς / διχῶς λεγόμενον”, meaning that a term can have different senses and double interpretation. In this paper we discuss how we reconstruct the meaning of an utterance in dialogue through the mechanism of interpretation and how we analyze and construe ambiguities, combining the insights of argumentation theory and text linguistics. Research results show that in case of misunderstanding, the “best interpretation” is the less defeasible one according to contextual presumptions.

Keywords: argumentation logic, ambiguity, context, semantic interpretation, Systemic Functional Grammar.

1. Introduction

The notion of ambiguity has been investigated since the first sophistic theories in Aristotelian Topics, Rhetoric and Sophistical Refutations [14], [5], [25], [7].¹ In the field of his dialectics, Aristotle supports that a term could have two different meanings (*παρὰ τὸ διπτόν, τὸ διχῶς λεγόμενον / τὸ διπτῶς λεγόμενον*). In *De Sophisticis Elenchis* an Aristotle's classification of linguistic fallacies is included, which is the first one in the Ancient Greek world. More specifically, Aristotle supports that there are 13 types of ambiguity. Six of these are called *linguistic ambiguities*, such as *syntactic ambiguity* and *lexical ambiguity* (*ὁμωνυμία*)², and they depend on the use of language [22]. Answerer (*interlocutor*) is allowed to ask for clarification from the questioner (*dialectician*) when he does not understand a term [24]. In this framework, problems of argumentation and communication may arise, resulting in different interpretations of an utterance and misunderstanding [1, p. 112b], [22]. It is remarkable the fact that Aristotle illustrates how expressions and definitions that involve temporal qualifications (i.e. *νῦν* = now) must be rejected in dialectic because of the ambiguity (*ἀμφιβολία*) which occurs in them [1, p. 142b21-33].³

In a pragma-dialectical approach, context seems to be crucial considering cultural factors, the purpose of dialogue and interlocutors' attitude [11], [26], [15], [18], [17]. In the last decades, context has become a significant concern for text linguistics and discourse analysis, taking into account interpersonal relationships in a dialogue, co-text (what precedes of an utterance and what follows), encyclopedic knowledge and social and cultural environment [20], [23]. Systemic Functional Grammar (SFG) could be a useful tool in order to analyze ambiguity in interpersonal relationships through argumentation [9], [8].

In this paper we examine a) the way we interpret ambiguities in argumentation applying Halliday's Systemic Functional Grammar (SFG), b) the role of contextual factors in argumentation analysis.⁴ The paper is structured as follows: Section 2 concludes the research methodology; Section 3 presents Halliday's Systemic Functional Grammar while Section 4 discusses the notion of ambiguity in argumentation and the role of contextual factors. Finally, Section 5 describes the main research conclusions providing perspectives for future work.

2. Methodology

For argumentation analysis, Systemic Functional Grammar (SFG) is applied, in order to reveal and analyze ambiguity/uncertainty, taking into account contextual factors. Systemic Functional Grammar (SFG) or Systemic Functional Linguistics (SFL) as proposed by Halliday in the 1960's is a model related to social semiotic approach to language, concerned with lexico-grammatical choices according to interlocutors' intention [8], [9]. For the analysis of ambiguities in argumentation, the specific model was chosen, due to the fact that offers a useful tool in a micro- and macro-level perspective. In this way, SFG could be a basic framework in order to solve ambiguity, viewed as a strategy of persuasion, and reveal the "best interpretation" taking into account contextual factors.

3. Systemic Functional Grammar (SFG)

Halliday's Systemic Functional Grammar emphasizes to the use of language. Language choices are related to speaker's and hearer's intention and what is important to SFG is the meaning and not the structure, as suggested by other linguistic models. In this framework, language is a system of meanings and socio-cultural context plays a crucial role (see Figure 1). Lexico-grammatical choices, such as adjectives, active or passive voice, epistemic modality, indefinite pronouns, present perfect tense, "construct" different aspects of social reality. Halliday proposes three main (meta)functions: a) ideational function, b) interpersonal function, c) textual function. The *ideational function* refers to the way the world is represented through language and it is related to encyclopedic knowledge. The *interpersonal function* deals with the way that language reflects identities or relationships within communicative discourse, for instance the relationship between

interlocutors. Finally, the *textual function* describes the use of language, structural relationships through lexico-grammatical choices [8].

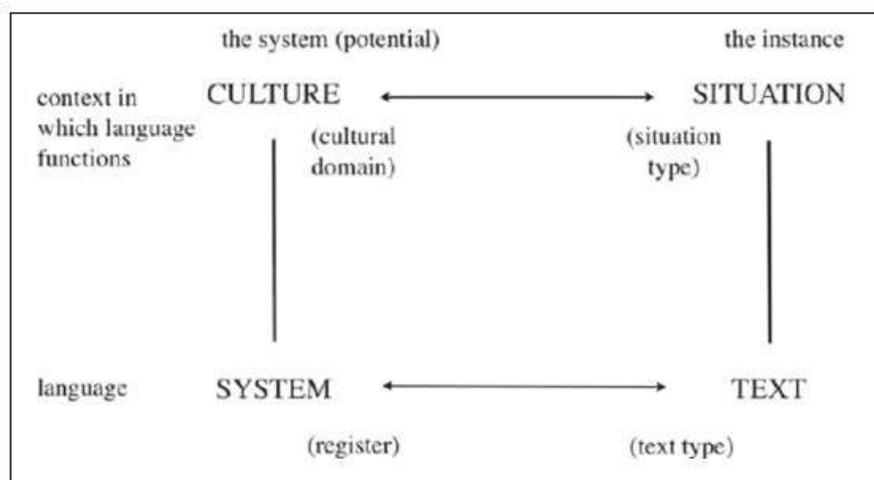


Figure 1: Text in context according to Systemic Functional Grammar [10].

4. Ambiguity in Argumentation and Context

Ambiguity, the property of a lexical item having more than one meaning, is a basic notion in semantics, syntax and pragmatics [2], [16], [19]. Lexical/semantic ambiguity refers to the presence of two or more possible meanings for a single word:

(1) I've brought the *seal* [4, p. 3]

In the above example, the multiple meaning of the word *seal* results to several interpretations. More particularly, the word *seal*, as a noun, could refer to “a sea animal that eats fish” or to “an official design or mark, stamped on a document to show that it is genuine and carries the authority of a particular person or organization” [27]. In this instance, in the word *seal* the phenomenon of *polysemy* is observed (a single lexeme has multiple meanings).

(2) a. *Paris* is a bustling metropolis.

b. *Paris* begins with the sixteenth letter of the English alphabet [4, p. 251].

In (2a) *Paris* refers to the city of France, while in (2b) it is mentioned to the word itself. The context in which the word is used plays a crucial role for the interlocutors. Speakers and hearers in a conversation rely on their background/encyclopedic knowledge (ideational function) and the co-text (in this case, the co-text is what follows the word *Paris*). The utterance is interpreted according to interlocutors' intention, taking into account lexico-grammatical choices, i.e. in (2b) *letter* and *alphabet* are nouns related to the same domain and they refer to the entry *Paris* in a dictionary (textual function).

In addition, ambiguity, paradox and vagueness are observed in arguments. Some examples from Modern Greek are:

(3) (premise 1) Ο Γιώργος είναι ευχάριστο άτομο (George is a pleasant person)

(premise 2) Το άτομο διασπάται (the atom disintegrates)

(conclusion) Ο Γιώργος διασπάται (George disintegrates)

(4) (premise 1) Ο αστυνόμος είναι όργανο (The police officer is an instrument)

(premise 2) Το μπουζούκι είναι όργανο (bouzouki⁵ is an instrument)

(conclusion) Ο αστυνόμος είναι μπουζούκι (The police is bouzouki)

In (3) a syllogistic argument is observed. The single lexeme *átomo* in the two premises is a polysemous word and it has multiple meanings. In the first premise, the word *átomo* refers to George as a person, i.e. the human being as a unit with its particular and unique characteristics as opposed to the species or the society. In premise 2, *átomo* is a term of physics or chemistry and refers to the least and invisible part of the matter. In this argument the premises are true⁶, but the conclusion is false.

In (4) the premises are true, but the conclusion is false. In this case, we observe a paradox that lies on the phenomenon of polysemy. In the first premise the word *όργανο* means *a person charged with a certain employment, especially within the framework of the state*, while in the second premise the notion *όργανο* refers to the Greek stringed-traditional instrument.

The phenomenon of polysemy is obvious in verbs, such as the verb *make* [28] in English, which has different meanings (see Figure 2). For example, the single lexeme *make* could mean *prepare* or *create* (she makes the table / she makes bread), *represent* (he made him a truly tragic figure), *appoint* (he made her his personal assistant), *equal* (5 and 8 makes 13), *force* (they made me follow the rules), etc. Ambiguity also occurs in cognitive verbs, such as *think*, *believe*, *assume*, *guess* and *suppose* [21], [13].

Polysemous word	Meaning
Make	1. prepare 2. represent 3. force 4. appoint 5. equal 6. calculate 7. reach 8. be suitable

Figure 2: Polysemous word *make* in English.

Although the verb *make* has different meanings, interlocutors consider all the contextual presumptions and confront ambiguities, choosing the less defeasible interpretation and taking into account the co-text of the utterance. In this way, they avoid miscommunication and misunderstanding. The “best interpretation” is related to contextual presumptions, such as the interaction, the background knowledge, interlocutor’s interests/values and the communicative purposes of utterances.

In another example, the preferred interpretation could be (5a)⁷:

- (5) a. The view could be improved by the addition of a *plant* out there.
b. The view would be destroyed by the addition of a *plant* out there [19 p.174].

In the above instance (5a), the best interpretation is *plant = living organism such as flower, tree or vegetable*, while in (5b) *plant = factory*.

In addition to lexical ambiguities often appear in speech and syntactic ones, as the following example⁸:

- (6) Flying planes can be dangerous.
 - a. The act of flying planes can be dangerous.
 - b. Planes that are flying can be dangerous.

In the above cases, interpretation of these utterances is mainly related to the macro-level perspective according to Systemic Functional Grammar and not to micro-level, that is the lexico-grammatical elements. The choice of “best interpretation” depends on factors, such as the knowledge of native speaker, his background and his communicative intention. For instance, for the first interpretation (6a) interlocutor may have personal experience with the planes, perhaps as a pilot, and may be able to evaluate possible imminent dangers. On the other hand, one could argue that the second interpretation (6b) is more possible not to be chosen by the speaker/hearer because is more diffuse and hard to follow, as we all know the fact that planes are a safe means of transport. In addition in the second example, the verb *fly* determines the noun *plane* and there is a distinction between flying planes and non-flying planes. In a macro-level perspective, solving this misunderstanding presupposes the encyclopedic interlocutors’ knowledge, their internal and external knowledge, according to ideational function, as proposed by Halliday in Systemic Functional Grammar. Relationship between the interlocutors seems to be important in order to choose the “best interpretation” and specifically in 6a, in case that one of them (or both) is pilot (interpersonal function). Taking into account the lexico-grammatical choices (textual function), i.e. the adjective *flying* that determines *planes*, a corpus analysis through concordances could reveal lexical collocations and the co-text information [3], [6], [12].

5. Conclusions

Halliday’s Systemic Functional Grammar could be a useful tool in order to construe ambiguities in argumentation. Through the three functions (ideational function, interpersonal function and textual function) is possible to construct the “best interpretation” of an utterance. Relationship between the speaker and the hearer, their intentions and their knowledge about the world according to the social and cultural environment, contribute to the analysis of ambiguity as a persuasion strategy. In addition, lexico-grammatical choices (textual function), such as passive voice, epistemic modality and verbs i.e. *think*, *suppose*, *believe*, could imply the phenomenon of semantic or syntactic ambiguity. In this framework, co-text (what precedes of an utterance and what follows) and lexical collocations may solve misunderstanding problems, while the role of lexicalized verbs (i.e. *make*) seems to be crucial. Finally, the use of corpus analysis through concordances may enlighten cases of vagueness observed in argumentation.

References

1. Aristotle. *Topics*. In *the complete works of Aristotle*, vol. I, ed. Jonathan Barnes. Princeton: Princeton University Press, 1991.
2. Atlas, J. *Philosophy Without Ambiguity*. Oxford: Oxford University Press, 1989.
3. Chang, C.-H., & Krulee, G. Resolution of ambiguity in Chinese and its application to machine translation. *Machine Translation*, 6(4), 1991, pp. 279–315.
4. Chapman, S., & Routledge, C. (Eds.). *Key Ideas in Linguistics and the Philosophy of Language*. Edinburgh University Press, 2009.
5. Crivelli, P. *Aristotle on Truth*, Cambridge, 2004.
6. Degano, C. Corpus linguistics and argumentation: Retrieving argumentative patterns in UK prime ministerial debates. *Journal of Argumentation in Context* 5(2), 2016, pp.113-138.

7. Golitsis, P. "Aristotle on Ambiguity". *Strategies of Ambiguity in Ancient Literature*, edited by Martin Vöhler, Therese Fuhrer and Stavros Frangoulidis, Berlin, Boston: De Gruyter, 2021, pp. 11-28.
8. Halliday, M.A.K. & R. Hasan. *Language, context and text: a social semiotic perspective*. Deakin University Press, 1985.
9. Halliday, M.A.K. *An introduction to functional grammar*. 2nd edition. London: Edward Arnold, 1994.
10. Halliday, M. A. K. The notion of context in language education. Jonathan J. Webster (ed.), *Language and education*. Volume 7 in the collected works of M. A. K. Halliday, 250–264. London & New York: Continuum, 2007.
11. Hamblin, C. L. *Fallacies*. London: Methuen, 1970.
12. Hinton, M. Corpus Linguistics Methods in the Study of (Meta)Argumentation. *Argumentation* 35, 2021, pp. 435–455.
13. Horn, L.R. A pragmatic approach to certain ambiguities. *Linguist Philos* 4, 1981, pp. 321–358.
14. Hutchby, I. Oh', Irony and Sequential Ambiguity in Arguments. *Discourse & Society* 12 (2), 2001, pp. 123–41.
15. Jackson, S. Frans H. van Eemeren and Wu Peng (ed.). Contextualizing Pragma-Dialectics: John Benjamins, Amsterdam/Philadelphia, Volume 12 of the Argumentation in Context Series. *Argumentation* 32 (2), 2017, pp. 293-299.
16. Jaszczolt, K. *Discourse, beliefs and intentions*. Oxford: Elsevier, 1999.
17. Koniak, P. and Cwalina, W. Does it pay to avoid speaking straight about controversial issues? Impact of argumentative ambiguity on the perception of the speaker, *Journal of Communication Management*, Vol. 26 No. 1, 2022, pp. 84-97
18. Larina, T., Ozyumenko V. and Ponton D. Persuasion strategies in media discourse about Russia: Linguistic ambiguity and uncertainty. Lodz Papers in Pragmatics Levinson, S. *Presumptive meanings: The theory of generalized Conversational 22 implicature*. Cambridge, Mass.: MIT Press, 2019.
19. Levinson, S. C. *Presumptive meanings: The theory of generalized conversational implicature*. The MIT Press, 2000.
20. Malinowski, B. The problem of meaning in primitive languages. Charles K. Ogden & Ian A. Richards (eds.), *The meaning of meaning*, London: Routledge, 1935, pp. 146–152.
21. Postal, P. On Certain Ambiguities. In *Linguistic Inquiry*, 5, 3, 1974.
22. Schiapparelli, A. Aristotle on the fallacies of combination and division in *Sophistici Elenchi* 4. *History and Philosophy of Logic*, 2003.
23. Smith, L. W. John-Steiner, V., Panofsky, C. P., Smith, L. W (eds.). "Context." *Sociocultural Approaches to Language and Literacy: An Interactionist Perspective*. Cambridge University Press, 1994.
24. van Eemeren, F. H. *Strategic maneuvering in argumentative discourse: Extending the pragma-dialectical theory of argumentation*. Amsterdam and Philadelphia: John Benjamins, 2010.
25. van Eemeren, F. H. Strategic maneuvering in argumentative discourse in political deliberation. *Journal of Argumentation in Context* 2 (1), 2013, pp. 10-31.
26. Walton, D. N. *Informal logic: A pragmatic approach*. Cambridge: Cambridge University Press, 2008.
27. https://www.oxfordlearnersdictionaries.com/definition/american_english/seal_2
28. https://www.oxfordlearnersdictionaries.com/definition/english/make_1?q=make

Notes

1. Also, in the Euthydemus Plato mentions some fallacies, but he does not attempt to classify them.
2. Homonymy (ὁμωνυμία) is a lexical ambiguity in which a single word has two or more different meanings.
3. See among others Schiapparelli [22].
4. It is worth mentioned the fact that the paper does not examine the case of lexicalized metaphors (for instance, he is a "legend"/"star").

-
5. Greek stringed instrument.
 6. The first premise is a personal view about George's character, that is may be true or false.
 7. It is worth noting the fact that the "best interpretation" in argumentation may be not identical for the interlocutors. For example, the opponent may consider (5b) as the best, while the proponent has (5a) in mind.
 8. Structural/syntactic ambiguity refers to the structure of a sentence that has multiple interpretations.

Dynamic Approximation of Self-Referential Sentences

Vladimir A. Stepanov

Independent researcher, Moscow, Russia

e-mail: vastvast@yandex.ru

Abstract:

Non-classical logic via approximation of self-referential sentences by dynamical systems are consistently presented. The new 6-valued truth values $\langle T, va, A, V, av, F \rangle$ (here $A=Liar$, $V=TruthTeller$) are presented as a function of the classical truth values $x_i \in \{0,1\}$, which resulted in a philosophical standpoint known as Suszko's Thesis. Three-valued truth tables were created corresponding to Priest's tables of the same name. In the process of constructing 4-valued truth tables, two more new truth values (va, av) were revealed that do not coincide with the four original ones. Therefore, the closed tables turned out to be 6-valued. Prof Dunn's 4-valued truth tables are compared with our 4-valued truth tables. De Morgan's laws are confirmed by six-valued truth tables. Constructed 3-, 4- and 6-valued lattices obeying De Morgan's laws.

Keywords: self-reference, dynamic, Liar, TruthTeller.

1. Introduction

Sentences that refer to themselves are called self-referential. The most popular of these is the 'Liar' sentence. It can be noted that the study of self-referencing admits two possible approaches:

- external – which describes the reaction of self-referential sentences to the system under study. These include the popular studies of Priest in 1978 (LP), see [9] and Dunn [1];
- internal – when the emphasis is on the study of the structure of self-referential sentences, which began with Peirce in 1855, [8], [4]. We will devote our article to this last approach.

The constructive analysis of the Liar sentence was carried out by Charles Peirce, [8], who, as far as we know, was the first to notice in his lectures in 1864 – 1865, that self-referential sentences generate an infinite sequence of substitutions into themselves. That was the first application of the principle, which in the second half of the 20th century was called “turning a vicious circle into a generating circle”.

We are talking about the **S** icon, which first appeared in the article by Albert Johnstone, 1981, [3]: $Q =_{df} \mathbf{S}_Q P$. (Formulas are given in A. Johnstone's notation; we do not decipher them). Our understanding of the icon **S** is somewhat different from A. Johnstone.

2. Basic Definitions

We define a dynamic approximation of self-referential sentences, which for the Liar and the TruthTeller, generates three-valued Kleene logic, and allows us to obtain new 4- and 6-valued truth tables [10]. We use a special self-referencing icon $\mathbf{S}x$ as a symbol for the self-referential sentences and place it front of the predicate $P(x)$. We call the predicate $P(x)$ the core of a self-referential sentence. A self-referential sentence looks like this:

$$\mathbf{S}xP(x). \quad (1)$$

The expression $\mathbf{S}xP(x)$ reads as follows: “self-referential by x P of x ”. The symbol $\mathbf{S}x$ in the formula $\mathbf{S}xP(x)$ connects the free variable x of the predicate $P(x)$. That is why we will call $\mathbf{S}x$ as a quantifier, a self-referential quantifier.

Expression (1) obeys the axiom of self-reference by Feferman, [2]:

$$\mathbf{S}xP(x) \leftrightarrow P(\mathbf{S}xP(x)). \quad (2)$$

Peirce [8] applied (2) to generate an infinite Liar sentence:

$$\mathbf{S}xP(x) \leftrightarrow P\left(P\left(P\left(\dots \mathbf{S}xP(x)\dots\right)\right)\right). \quad (3)$$

Consider the iterative steps that bring Peirce to the infinite formula:

$$\mathbf{S}xP(x) \leftrightarrow P(\mathbf{S}xP(x)) \leftrightarrow P\left(P(\mathbf{S}xP(x))\right) \leftrightarrow P\left(P\left(P(\mathbf{S}xP(x))\right)\right) \leftrightarrow \dots \quad (3.1)$$

Let us arrange formulas (3.1) in the natural order of increasing their lengths:

$$\langle \mathbf{S}xP(x), P(\mathbf{S}xP(x)), P\left(P(\mathbf{S}xP(x))\right), P\left(P\left(P(\mathbf{S}xP(x))\right)\right), \dots \rangle. \quad (3.2)$$

In the formulas of the sequence (3.2), we replace the formulas $\mathbf{S}xP(x)$ by the variable x . The resulting sequence (3.3) will be denoted as

$$\mathbf{S}xP(x) = \langle x, P(x), P(P(x)), P(P(P(x))), \dots \rangle. \quad (3.3)$$

Definition 0: The expression $\mathbf{S}xP(x)$ will be called an approximation of the expression $\mathbf{S}xP(x)$:

$$\mathbf{S}xP(x) \approx \mathbf{S}xP(x). \quad (4)$$

Expression (4) is the definition of the trajectory of a dynamical system of the form $(\{0,1\}, P(x))$ with orbits $\langle P^n(x), n \in \mathbb{Z}^+ \rangle$, where $P^n(x) = P(P^{n-1}(x))$, by [6]. Consider the case when the kernels of self-referential sentences $P(x)$ are composed of $Tr(x)$ using the propositional connectives of equivalence and negation:

$$P(x) \in \{Tr(x), \neg Tr(x), Tr(x) \leftrightarrow Tr(x), Tr(x) \leftrightarrow \neg Tr(x)\}. \quad (5)$$

It is easy to see that expression (4) is periodic, with a maximum period of 2. This means that the second and third terms of the sequence (4) determine the rest of the infinite sequence. Therefore, in our case, we rightfully shorten the definition of the self-referencing quantifier as follows:

$$SxP(x) = \langle x, P(x), P(P(x)) \rangle. \quad (6)$$

The variable x and the predicates $P(x)$ from (5) in our case take values from $\{0,1\}$.

Definition 1: For $SxP(x) = \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \}$:

$$\begin{aligned} \neg SxP(x) &= \neg \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \} \\ \neg SxP(x) &= \{ \neg \langle 1, P(1), P(P(1)) \rangle, \neg \langle 0, P(0), P(P(0)) \rangle \} \\ \neg SxP(x) &= \{ \langle \neg 1, P(\neg 1), P(P(\neg 1)) \rangle, \langle \neg 0, P(\neg 0), P(P(\neg 0)) \rangle \} \end{aligned} \quad (7)$$

This is the table for the negation:

$SxP(x)$	$\neg SxP(x)$
$\{ \langle 1, 1, 1 \rangle; \langle 0, 1, 1 \rangle \} = T$	$F = \{ \langle 0, 0, 0 \rangle; \langle 1, 0, 0 \rangle \}$ (False)
$\{ \langle 1, 0, 1 \rangle; \langle 0, 1, 0 \rangle \} = A$	$A = \{ \langle 0, 1, 0 \rangle; \langle 1, 0, 1 \rangle \}$ (Antinomy, Liar)
$\{ \langle 1, 1, 1 \rangle; \langle 0, 0, 0 \rangle \} = V$	$V = \{ \langle 0, 0, 0 \rangle; \langle 1, 1, 1 \rangle \}$ (Void, TruthTeller)
$\{ \langle 1, 0, 0 \rangle; \langle 0, 0, 0 \rangle \} = F$	$T = \{ \langle 0, 1, 1 \rangle; \langle 1, 1, 1 \rangle \}$ (True)

Definition 2: We define two-place connectives $o \in \{\wedge, \vee, \rightarrow, \leftarrow\}$ for two S-formulas $SxP(x)$ and $SxQ(x)$. We study such a variant of two-place connectives, when the trajectories of estimates of the formula $SxP(x)$ of the one branch ($x = 1$ or $x = 0$) interact with the trajectories of the formula $SxQ(x)$ of the same branch ($x = 1$ or $x = 0$, respectively):

$$\begin{aligned} SxP(x) \circ SxQ(x) &:= \\ \{ \langle 1, P(1), P(P(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \} \circ \{ \langle 1, Q(1), Q(Q(1)) \rangle, \langle 0, Q(0), Q(Q(0)) \rangle \} &= \\ \{ \langle 1, P(1), P(P(1)) \rangle \circ \langle 1, Q(1), Q(Q(1)) \rangle, \langle 0, P(0), P(P(0)) \rangle \circ \langle 0, Q(0), Q(Q(0)) \rangle \} &= \\ \{ \langle 1 \circ 1, P(1) \circ Q(1), P(P(1)) \circ Q(Q(1)) \rangle, \langle 0 \circ 0, P(0) \circ Q(0), P(P(0)) \circ Q(Q(0)) \rangle \}. \end{aligned}$$

$$\begin{aligned} \text{Example.: } F \wedge V &= \{ \langle 1, 0, 0 \rangle, \langle 0, 0, 0 \rangle \} \wedge \{ \langle 1, 1, 1 \rangle, \langle 0, 0, 0 \rangle \} = \\ &= \{ \langle 1, 0, 0 \rangle \wedge \langle 1, 1, 1 \rangle, \langle 0, 0, 0 \rangle \wedge \langle 0, 0, 0 \rangle \} = \\ &= \{ \langle 1, 0, 0 \rangle, \langle 0, 0, 0 \rangle \} = F. \end{aligned}$$

3. Main Results

Let's compare Kleene-Priest tables for \wedge of the Liar sentences with the tables obtained for values A and V:

Kleene-Priest p				Hypothesis: p = A				Hypothesis: p = V			
\wedge	t	p	f	\wedge	T	A	F	\wedge	T	V	F
t	t	p	f	T	T	A	F	T	T	V	F
p	p	p	f	A	A	A	F	V	V	V	F
f	f	f	f	F	F	F	F	F	F	F	F

Lemma 1: 1. The sentences *Liar* (=A) have a tabular model isomorphic to Priest's tabular model

for *Liar* (= p) [9].

2. The sentences *TruthTeller* (=V) have a tabular model isomorphic to Priest's tabular model for *Liar*(p).

\wedge	T	A	V	F
T	T	A	V	F
A	A	A	<i>av</i>	F
V	V	<i>av</i>	V	F
F	F	F	F	F

\vee	T	A	V	F
T	T	T	T	T
A	T	A	<i>va</i>	A
V	T	<i>va</i>	V	V
F	T	A	V	F

Lemma 2: When constructing the interaction of V and A, new truth values were obtained: $A \wedge V = \{ \langle 1,0,1 \rangle, \langle 0,0,0 \rangle \} = av = \neg(va)$, $A \vee V = \{ \langle 1,1,1 \rangle, \langle 0,1,0 \rangle \} = va = \neg(av)$.

The author has not come across any statement in the literature that the sentences $A \wedge V$ and $A \vee V$ have truth values similar to *av* and *va*, respectively.

For comparison, here are the Dunn [1] tables : Dunn [1] compiled 4-value tables for TBNF truth values. They are intended for reasoning to the computer on inconsistent data B or their absence N. Dunn used the truth values of T and F to close the tables when the scores N and B interacted. They are labeled inside the tables.

\wedge	T	B	N	F
T	T	B	N	F
B	B	B	<i>F</i>	F
N	N	<i>F</i>	N	F
F	F	F	F	F

\vee	T	B	N	F
T	T	T	T	T
B	T	B	<i>T</i>	B
N	T	<i>T</i>	N	N
F	T	B	N	F

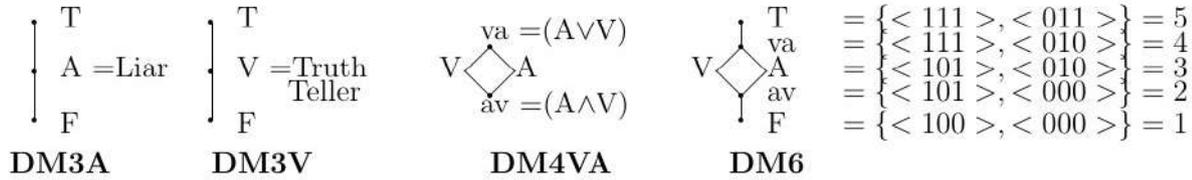
However, many researchers use these tables to analyze self-referential sentences, assuming $N=V$ and $B=A$. In our case, the tables are not closed: $A \vee V = va$ and $A \wedge V = av$, which encourages the construction of new, already six-valued ones. Fortunately, they are already closed. These are the complete 6-valued tables:

\neg		\wedge	T	<i>va</i>	A	V	<i>av</i>	F
T	F	T	T	<i>va</i>	A	V	<i>av</i>	F
<i>va</i>	<i>av</i>	<i>va</i>	<i>va</i>	<i>va</i>	A	V	<i>av</i>	F
A	A	A	A	A	A	<i>av</i>	<i>av</i>	F
V	V	V	V	V	<i>av</i>	V	<i>av</i>	F
<i>av</i>	<i>va</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	<i>av</i>	F
F	T	F	F	F	F	F	F	F

\vee	T	<i>va</i>	A	V	<i>av</i>	F
T	T	T	T	T	T	T
<i>va</i>	T	<i>va</i>	<i>va</i>	<i>va</i>	<i>va</i>	<i>va</i>
A	T	<i>va</i>	A	<i>va</i>	A	A
V	T	<i>va</i>	<i>va</i>	V	V	V
<i>av</i>	T	<i>va</i>	A	V	<i>av</i>	<i>av</i>
F	T	<i>va</i>	A	V	<i>av</i>	F

Lemma 3: The next four lattices are DeMorgan lattices, á la Leitgeb, [7]:

$$\{ F \leq av \leq A \leq V \leq va \leq T \}; (1 \leq 2 \leq 3 \leq 3 \leq 4 \leq 5):$$



4. Conclusion

The proposed truth-values are finite estimates of infinite periodic classical sequences of kernels of the self-referential statements. This result is consistent with R. Suszko's Thesis of transforming of the sets of non-classical truth-values into the sets of classical truth-values.

Acknowledgements

Let me thank I. B. Mikirtumov and L. D. Beklemishev for their support at a difficult moment, and anonymous reviewers for good advice.

References

1. Dunn J.M., Two, three, four, infinity: The path to the four-valued logic and beyond, *New Essays on Belnap-Dunn Logic*, ed. by H. Omori and H. Wansing, Cham: Springer, 2019, 77-97.
2. Feferman S. Toward Useful Type-Free Theories I, *The Journal of Symbolic Logic*, 1984, 49, 75-111.
3. Johnstone A. Self-reference, the Double Life and Godel, *Logique et Analyse*, 1981, 24, 35-47.
4. Kauffman L.H., Knot Logic. In: *Knots and Applications*, ed. by L. Kauffman, World Scientific Pub., 1994, 1-110.
5. Kleene S. C., *Introduction to Metamathematics*. D. Van Nostrand Co., Princeton, 1950.
6. Konev B. Kontchakov R., Wolter F., Zakharyashev M. On Dynamic Topological and Metric Logics, *Studia Logica*, 2006, 84, 129-160.
7. Leitgeb H., Truth and the Liar in De Morgan-Valued Models, *Notre Dame Journal of Formal Logic*, 1999, 40(4), 496-514.
8. Michael E., Peirce's Paradoxical Solution to the Liar's Paradox, *Notre Dame Journal of Formal Logic*, 1975. XII, No. 3, 369-374.
9. Priest G., The Logic Paradox, *Journal of Philosophical Logic*, 1979, 8, 219-241.
10. Stepanov V., In defense of the self-referencing quantifier Sx . Approximation of self-referential sentences by dynamic systems, *Logico-Philosophical Studies*, 2021, 19(2), 145-150, DOI: 10.52119/LPHS.2021.49.50.014.

**Determining Argumentative Dispute Resolution Reveals
Deep Disagreement Over Harassment Issue (A Case-Study
of a Discussion in the Russian Parliament)**

Elena Lisanyuk

St Petersburg State University
7-9, Universitetskaya nab.,
199034 St Petersburg, Russian Federation;
National Research University Higher School of Economics
21/4 Staraya Basmannaya St.,
105066 Moscow, Russian Federation

e-mail: e.lisanuk@spbu.ru
<https://orcid.org/000-0003-0135-4583>

Abstract:

In 2018, three journalists accused one of the Members of the Russian Parliament of harassment at workplace. Many influential persons of the Russian elite engaged themselves in the public discussion of the conflict. We studied that high-profiled discussion using a hybrid method merging human- and logic-oriented approaches in argumentation studies. The method develops ideas of the new dialectics, the argumentation logic and the logical-cognitive approach to argumentation, on which is based the algorithm for determining of dispute resolution by aggregating formal and informal tools of analysis. We have reconstructed the discussion as two disputes about questions A and B. A: Did the MP violate the code of conduct by making statements or actions against the journalists? B: Are actions like the behavior of the MP harassment? The opinions of the discussion participants were grouped into the four points of view: A1 – the MP did not violate the code of conduct, A2 – the MP violated the code of conduct, B3 – the actions are not harassment, B4 – the actions are harassment. We mapped arguments in support or against each of them using OVA software, evaluated the arguments with the help of the critical questions, a tool proposed in the new dialectics, and determined the ultimate A + B resolution by applying of the algorithm that combines elements of gradual and labelling semantics from the argumentation logic and the classification of disputes from the dialectical approaches. The resolution was a subset of four arguments that ensured the victory of A1+B4. However, the substantial incompatibility of those arguments highlighted a deep disagreement, an unresolvable difference of opinion, between the parties about the permissibility of courtship. The deep disagreement, a bonus result yielded by the application of the hybrid method, excluded the interpretation of the determined resolution

as convincing for the parties, but pointed out a way to smooth the difference of opinions by elaborating of legal, social and moral aspects of the problem of harassment at workplace.

Keywords: argumentation logic, new dialectic, logical-cognitive approach to argumentation, computing of dispute outcomes, evaluation of arguments, critical questions.

1. Introduction

In February 2018, a scandalous conflict over harassment erupted in the State Duma of the Russian Federation – the Lower Chamber of the Russian Parliament. Three journalists, Ekaterina Kotrikadze (RTVI Channel <https://rtvi.com>), Farida Rustamova (BBC Russian <https://www.bbc.com/russian>) and Darya Zhuk (TV Rain <https://tvrain.ru>), hereinafter referred to as the Journalists, complained to the Ethics Commission of the State Duma about the indecent behavior of Leonid Slutsky, one of the deputies, hereinafter referred to as the MP. The MP denied all charges. The Ethics Commission found no “violations of the code of conduct” in the MP’s actions and pointed out signs of bias in the accusations of the Journalists [35]. Later, the MP apologized and said that he did not want to offend the Journalists and “did not cross borders” [24].

The Journalists’ complaints triggered a discussion in the media, which exposed legal, social and moral aspects of the problem of harassment. Russian legislation provides no definition of bullying or harassment at workplace, not necessarily for sexual purposes, as inadmissible forms of behavior that “should be separated from other forms of reprehensible behavior <...> and prohibited by law as harassment” [13, p. 57]. The social boundary of courtship permissibility is blurred, and its establishment is hindered by “elements of archaism, the inseparability of family and sexual relations in labor relations” [25, p. 49]. Women are morally humiliated by “obscene offers <...> in exchange for a job or other benefits, < ... > emphasizing their unequal position with men” [2, p. 8].

The purpose of this study is to demonstrate the potential of a hybrid method of argumentation analysis aimed at determining dispute resolution. The method elaborates the concept of new dialectics [27] and implements an algorithm for search and selection of the dispute resolutions, based the logical-cognitive approach to argumentation [14] which employs the idea of defeasible argumentation, as developed in the argumentation logic, see [21] for an overview. The new dialectics and other dialectical conceptions of argumentation exhibit human-oriented approach in their studies of argumentation while the argumentation logic is an influential branch of logic-oriented approaches. The algorithm for determining of the dispute resolutions proposed in the vein of the logical-cognitive approach combines the advances in the human- and logic-oriented approaches for providing of an effective tool for the argumentation analysis. The algorithm includes reconstruction of argumentative discussion as a dispute of definite kind, evaluation of the arguments in it, computing the outcomes and determination of the resolution of the dispute. The human-oriented and logic-oriented approaches both view argument as a piece of reasoning consisting of premises and conclusion, but the latter pursues the normative ways of how the conclusion follows out the premises, while the former proceeds both in the normative and descriptive directions and focuses on how its premises serve as reasons one party offers to the other party in a dialogue in order to get her to agree to its conclusion. Deductive inference and formal entailment are the cornerstones of logic-oriented approaches. S. Toulmin’s model of argument is an influential sample of a human-oriented conception.

In the new dialectics, arguments are evaluated with the help of the critical questions formulated in relation to the scheme of argumentation each of the arguments instantiate, and in the argumentation logic, they are evaluated with respect to their relations to each other such as attack or

support. By merging those evaluations of the arguments, the algorithm allows defining the outcomes of the dispute and selecting the resolutions of the dispute out of them. The outcomes are the stronger arguments in the dispute, they consist of the sets of arguments which are proposed in favor of the one viewpoint that are not rejected by counterarguments supporting the opposite viewpoint. The selection of the resolutions of the dispute out of its outcomes amounts to determining the subsets of the outcomes that either belong to the position of one or another party, or make up an intersection of them, depending on the type of dispute, as proposed in the formal dialectics [4] and adopted in many dialectical approaches including the pragma-dialectics [10]. In this study, the search of the outcomes and resolutions is outlined with the help of argumentation mapping of the dispute visualized with OVA software <http://ova.arg-tech.org/>, one of the remarkable achievements of the collaboration between the human- and logic- oriented approaches, see [28] for its evolution.

The four substantial results obtained with the hybrid method manifest its efficiency and efficacy. They are a fruitful merging of human- and logic- oriented approaches to argumentation; the determining of the dispute resolution; revealing of the deep disagreement and pointing to a way of smoothing it. By the deep disagreement philosophers [29] and argumentation theorists [12] call agents' discord over issues regarded both so fundamental and uncompromisable that, whenever those issues arise in discussions, they lead to unresolvable deadlocks. The two former results are expected outcomes of creating and applying of the algorithm, but the two latter ones are unexpected happy bonuses of that, and they provide a convenient tool for handling high-profile conflicts and assessing public discussions over sensitive issues. The relevance of the bonus results of the algorithm application amounts to its ability to discover hidden controversies in the views of the parties in conflicts where the obvious fact that its parties share some views belonging to the explicit agenda of discussion conceal their opposition over some other issues, which may appear irrelevant to the agenda but constitute parties' deeply implicit standpoints employed to support their explicitly put views. Further discussion of those implicit standpoints may lead to a resolution, but may end up in a deadlock, too, which in the sensitive issues risk polarization in public opinions and endanger public support for important decisions regardless of whatever reasonable and well prepared they are. Discovering of the deep disagreement at a certain point of public conflict helps, on the one hand, to stop chancy search of a resolution before the disagreement gets radicalized, and, on the other hand, to develop the issues over which the parties agree, if needed.

2. Reconstruction of the Discussion

We reconstructed the harassment discussion as two disputes over issues A and B, which affected all the three aspects of the problem of harassment – legal, social, and moral:

- A. Did the MP violate the code of conduct by making statements or actions against the Journalists?
- B. Are statements or actions like the behavior of the MP harassment?

The opinions of the discussion participants amounted to the following four points of view:

- A1 – the MP did not violate the code of conduct,
- A2 – the MP violated the code of conduct,
- B3 – the actions are not harassment,
- B4 – the actions are harassment.

See Diagrams 1-3 for the visualization of the reconstruction.

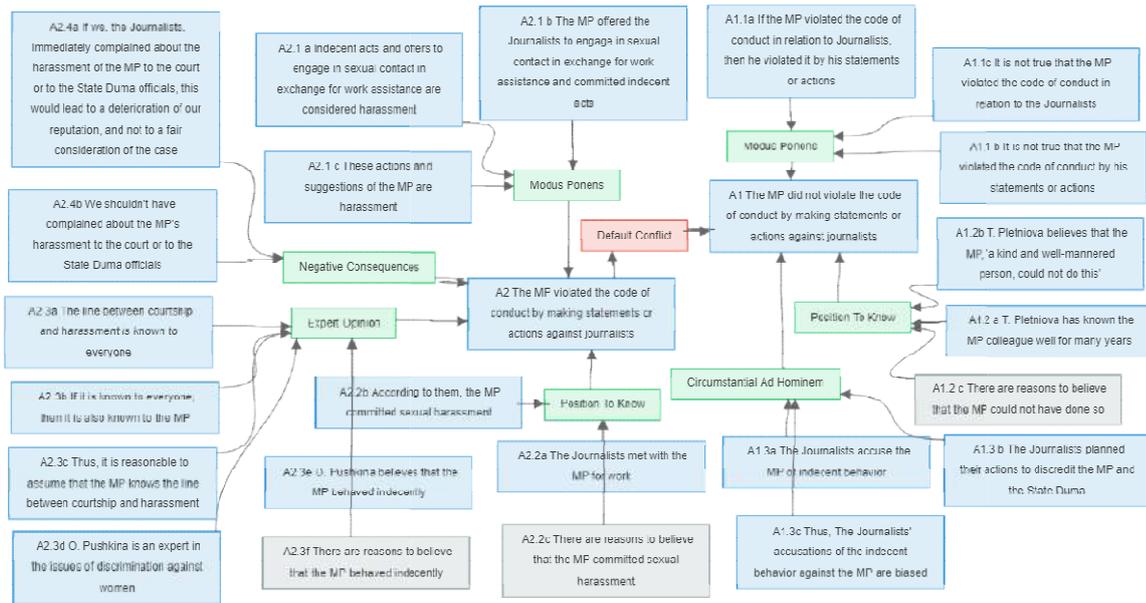


Figure 1: Diagram 1. Reconstruction of Dispute A. Did the MP violate the code of conduct by making statements or actions against the Journalists?

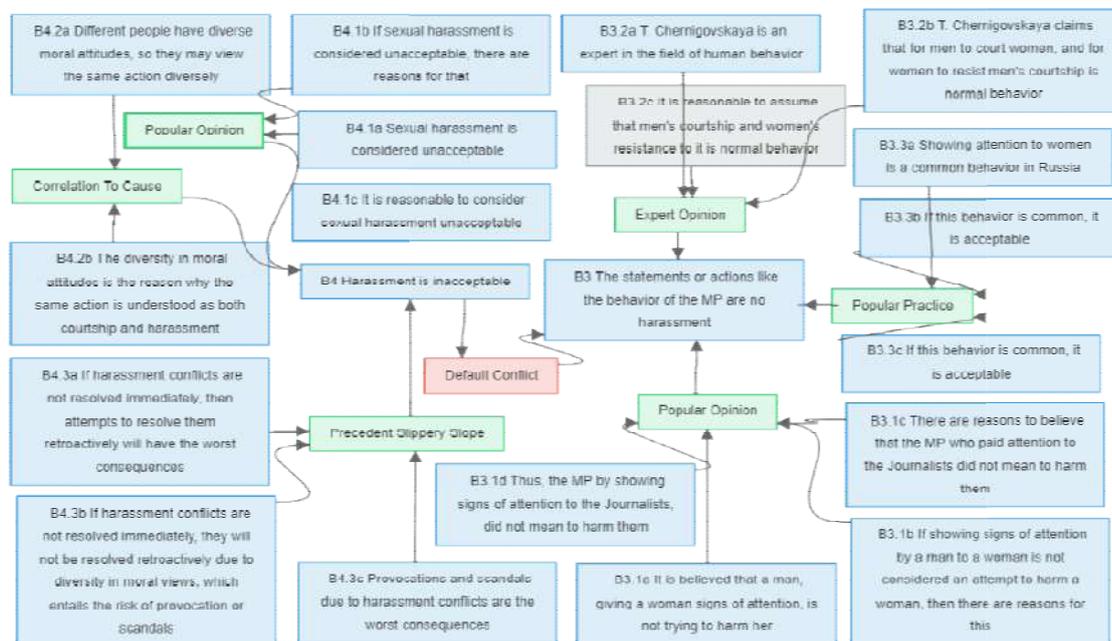


Figure 2: Diagram 2. Reconstruction of Dispute B. Are actions or statements like the behavior of the MP harassment?

There were three arguments put in favor of each of A1, B3 and B4, and four arguments were put forward in support of A2, see Fig.1 and Fig. 2. In dispute A, A2 prevailed and was then rejected in dispute B, for the arguments in favor of B4, aimed at rejecting B3, indirectly attacked A2 and overruled it (Fig. 3).

The ultimate resolution amounted to the subset of the four arguments that ensured the victory of A1+B4:

A1.1 The MP did not violate the code of conduct for the accusations are unproven, and
 A1.2 this is not in his nature.

B4.1 the harassment like the actions of the MP is unacceptable, and

B4.2 and 4.3 as courtships can be regarded in different ways, charges of indecent behavior must be brought immediately in order to be investigated without delays.

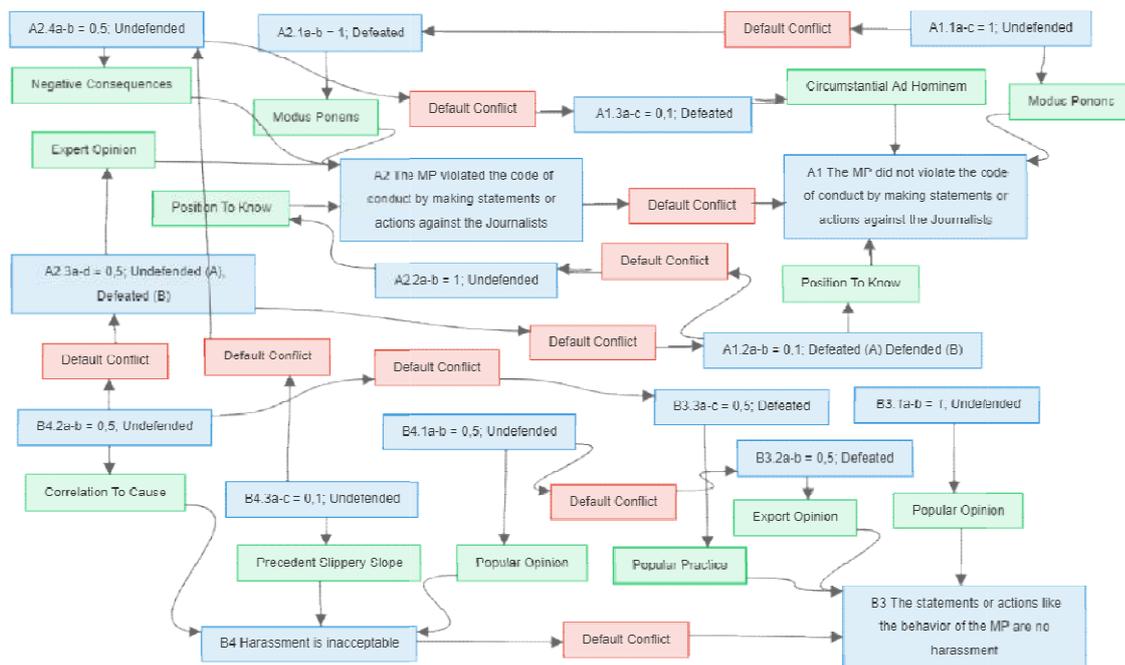


Figure 3: Diagram 3. The outcomes of Disputes A and B with the evaluation of arguments.

3. Visualization and Mapping the Disputes

Mapping the disputes and their digital visualizations are an integral part of the reconstruction of the discussion, it contributes to establishing of the content of the points of view, identifying of the arguments, including implicit ones, clarifying of their argumentation schemes, and determining of the type of dispute. It results in the argumentative maps of the disputes (Diagrams 1 and 2 on, respectively, Fig. 1 and Fig. 2), which make it possible to determine the resolution of the whole dispute by evaluating the arguments and computing of their outcomes (Diagram 3 on Fig. 3). Argumentation mapping is widely used in argumentation analysis in both human- and logic-oriented approaches, see [5], [20] for good examples. The advantage of our approach is that the mapping is done with the help of the specialized software designed to visualize aspects of

argumentation that traditional flowcharts may fail to discover or distinguish, since accurate mapping is essential for a more precise assessment of arguments in determining the resolution of a dispute. This applies to the enthymematic reasoning, characterizing the majority of the arguments in the dispute, when one or more premises or conclusions are presupposed by the author without being explicitly stated in the dialog, and to such element of arguments' structure as their demonstration, the connection of its premises to conclusions, which hardly ever gets explicit. In Diagrams 1, 2 and 3, the types of the demonstration of the arguments are marked in the green cells.

The visualization and argumentation mapping opened a perspective of examining the public discussion in its entirety and in detail in the absence of a unified text or protocol that reflected it. Due to the technical limitations of argumentation mapping by traditional flowcharts or formulaic notation, researchers normally have to deal with single arguments, excerpts from discussion transcripts, pre-prepared or abbreviated texts, any of which risk distorting the result. Visualization through OVA allowed us to reduce the impact of such technical limitations on the result and enabled to collect and map the opinions and arguments of the parties published in various media during several months of 2018, when the discussion was going on.

This made it possible to abstract from the secondary branches of the discussion, to establish the key points of view and arguments of the parties and to group the opinions of dozens of participants in the discussion around them, reducing the number of characters in our study to either the direct participants in the conflict over the harassment or the most influential people in the aftermath public discussion. In addition, the visualization made it possible to establish relations of support and criticism between the arguments of the parties, making explicit the premises or conclusions initially left implicit.

4. Evaluation of Arguments

In argumentation logic, there are several kinds of semantics designed to evaluate which of the arguments proposed in a dialogue are acceptable for a rational agent. In our algorithm, the primary evaluation of arguments employs ideas of labelling and gradual semantics. However, it executes that not on a graph which is a general way for it in the argumentation logic, as, for instance [1] remarkably does, but with the help of mapping instead. The gradual semantics qualify the strength of arguments by special functions assigning weighs to arguments for determining how strong or weak an argument of definite weight has to be in relation to other arguments in order for a rational agent to agree or disagree with its conclusion, respectively [6]. The labelling semantics evaluates arguments in relation to the set of arguments withstanding counter-argumentation as belonging to it – *in*, not belonging – *out* or *undecided* [3].

Answers to the critical questions divide the set of arguments in the dispute into *strongly sustainable* ones with a conventional weight = 1, which to all critical questions give answers compatible with their premises and conclusion, *averagely* = 0.5 or *weakly sustainable* = 0.1, if the answers reject more or less than half of the critical questions, respectively. In Diagram 3 (Fig. 3), the sustainability of arguments is indicated in the premise cell that combines the premises of each argument, which in their expanded form are mapped in Diagrams 1 and 2 (Fig. 1 and Fig. 2). Some interim conclusions in the dispute A and B are left out in the Diagram 3, with their evaluation provided in the relevant premises' cells.

5. Computing of the Outcomes and Determining the Resolution

Computing of the outcomes and determining the resolution of the disputes A and B are based on the extension semantics of the argumentation logic. According to it, the ordering of the set of arguments of the dispute is modelled on a directed graph by means of the binary abstract attack

relation *attack* [α , β] between arguments α and β , symbolizing how criticism of argument α rejects argument β ; and how counter-argumentative attack [γ , α] rejects argument α by counterattacking it and thereby returns argument β as defended [10]. A practical application of the extension semantics to the evaluation of arguments in a meaningful dispute is proposed in [15].

Instead of a graph and formalism, we relied on the mapping and visualization of disputes, which is more convenient for a meaningful analysis of the argumentation. Along with that we preserved the terminology of attacks, counterattacks and defenses inherent in the argumentation logic. The use of mapping restricts the interpretation of the outcomes and resolutions of the dispute to that given dispute but opens up the prospect of a formalized analysis of meaningful discussions. The visualized mapping enabled us to detect a deep disagreement concealed under the inconsistency of A + B dispute resolutions, which was hardly possible to do by means of formalization on a graph.

The outcomes of the disputes are established by determining of a stronger argument in each pair of arguments attacking one another (shown by red cells in Fig.3). An argument is *defeated* (shown in Figure 3 as “defeated”) if it is attacked by an equal or more sustainable argument; *undefended* (shown in Figure 3 as “undefended”) if it is attacked by a less sustainable argument or not attacked; and *defended* (shown in Figure 3 as “defended”), if it was attacked by an argument that was counterattacked and defeated. The position of the party with the biggest number of *not defeated* arguments, i.e. *defended* or *undefended*, is considered the most convincing. In Dispute A (Fig. 1), A2 is most convincing, supported by undefended A2.2, A2.3 and A2.4, against undefended A1.1 in support of A1. In Dispute B (Fig. 2), B4 is the most convincing argument with the three undefended arguments in its support against undefended B3.1 in support of B3.

In the contemporary dialectical approaches, disputes are divided into single and multiple ones according to the number of propositions that constitute the content of the parties 'points of view, and into unmixed and mixed ones, depending on the parties' intentions to defend their point of view or criticize other opinions, or do both, respectively. Dispute A is a single mixed dispute where A1 seeks to prove that A2 is unjustified. To resolve such a dispute in favor of A2, it is sufficient that there is at least one defended argument in the set of arguments in support of it, otherwise A1 prevails. Dispute B is a multiple mixed dispute, in which each party seeks to get defended its point of view and refuted the opposite. To resolve this dispute in favor of either party, it is sufficient that in its position the number of not defeated arguments exceeds the number of such arguments in the opposite position.

A uniform assessment of the validity of deductive, inductive, and plausible arguments is provided by the special algorithm for evaluating each of them using critical questions, which allows computing the outcome of the dispute with respect to the demonstrative quality of the arguments. A plausible argument is a reasoning that provides prima facie acceptable conclusion, based on the assumption that its premises are true in the absence of evidence of the contrary, and that conclusion will have to be discarded if there arrives novel evidence of the falsity of the assumption. Critical questions to plausible arguments test the assumptions in relation in their schemes of argumentation. Plausible arguments are often considered fallacies, because their acceptability is defeasible and depends on the context of the dialog, and a plausible argument that is acceptable in one dialog can be found not acceptable in another dialog. Unlike that, deductive or inductive arguments, the assessment of correctness or validity of which is indefeasible and independent of context of the dialog where they appear, because it is based on their logical form or probabilistic support their premises provide for their conclusions, respectively. Critical questions to deductive arguments check their correctness and verify the premises. No inductive arguments were found in arguments A and B.

6. Deep Disagreement

Unlike normal difference of opinions, a necessary condition for any discussion to start, deep disagreement is an abnormal difference of opinions. It points to the impotence of deductive or inductive arguments to convince amid “the assumption that earnest clear thinking can resolve fundamental issues.” The deep disagreement is the situation in a dialog when “the parties on both sides might agree on all historical and statistical matters, but still disagree. The dispute is, in fact, one concerning *moral standing*.” [12, pp. 10-11]. In the case of a normal difference of opinions, the parties share some views that form epistemic or procedural foundations for resolving it, so “341...that some propositions are exempt from doubt, are as it were like hinges on which those turn”... But when “611. two principles really do meet which cannot be reconciled with one another” parties face the deep disagreements pointing to the emptiness of the set of those hinges-like views they have to share in order to come to an agreement, and ‘then each man declares the other a fool and heretic.’ [29]. To resolve a deep disagreement there is a need for persuasion where the non-demonstrative plausible arguments play their key role. A non-empty resolution of the dispute where a deep disagreement occurs implies a compromise [16] or, otherwise, termination of the discussion as a way out [11].

In the discussion of harassment, the deep disagreement concerned the understanding of signals about unwanted attention: are those signals part of a flirting culture that implies recipients’ implicit consent to the courtship – A1 and B3, or do they express a clear rebuff that turns courtship into harassment – A2 and partly B4?

In favor of the former, Tatyana V. Chernigovskaya, professor at St. Petersburg State University and influential cognitivist, argued that men’s courtship of women is biologically justified by the need for procreation and therefore it is evolutionarily predetermined [7]. Her arguments B3.2b and B3.3a implied that there is no borderline between harassment and signs of courtship, on the existence of which insisted MPs Oksana V. Pushkina [32] and Alexey B. Veller [26] in A2.3a and B4.2 a-b who both regarded any forms of harassment intolerable.

Position B4 – harassment is inadmissible, and in order to stop it, in each case it is necessary to draw a borderline between it and signs of attention – highlighted the legal and social aspects of the problem of harassment and fell in between the polar positions of the parties of the deep disagreement A1+B3 and A2. The proponents of the drawing of the line of permissible courtship A2 and B4 came out in solidarity in support of the inadmissibility of harassment against the views of the supporters of A1, who considered harassment to be courtship. However, the supporters of B4 argued in favor of A1 in saying that it was not always possible to establish this boundary, and therefore that discussion about the harassment in the State Duma is just such a case. The deep disagreement over the issue of the borderline of courtship permissibility suggests that in further discussions which appear inevitable for many reasons, the polarization will increase. Due to the psychological phenomenon of group polarization and the cognitive confirmation bias, those who were initially supporting B4 would lean to one of the other poles [18] until after a conflict resolution procedure is established for the legal or moral aspects of sexual harassment, which would enable to eliminate the social aspect of this problem as well.

7. Evaluation of the Arguments in Tthe Discussion About Harassment

In this section, we evaluate the sustainability of the arguments proposed in the discussion using critical questions as they are formulated in [27].

Three arguments were put forward in defense of A1: a deductive argument marked as *modus ponens* A1.1 about the audio recording and the two plausible arguments *from position to know* A1.2 and an indirect argument *circumstantial ad hominem* A1.3.

A1.1 is a simple destructive dilemma:

- a) *If the MP violated the code of conduct in relation to the Journalists, then he violated it by his statements or actions.*
- b) *It is not true that the MP violated the code of conduct by his statements or actions.*
- c) *So, it is not true that the MP violated the code of conduct in relation to the Journalists.*

Premise *a* summarizes the questions the Journalists were asked by the Ethics Committee. The truth of *b* follows from the consensus of the parties that “the originality of the MP’s statements is known to everyone in our country” according to one of the members of the Ethics Committee. One of the Journalists agreed with this and added that for this reason, the Journalists tried not to pay attention to the peculiar manner of behavior of the MP [24]. “The Commission is not authorized to give expert assessments to audio recordings” [35], therefore, it was not possible to use audio recordings to confirm actions that violate the norms of behavior. The overall score of A1.1 is 1, strongly sustainable.

A1.2 was put forward by Tamara V. Pletniova, a colleague of the MP in the State Duma, who acted as an informed person – a person who, due to circumstances, happened to possess relevant information.

- a) *Tamara Pletniova has known the MP colleague well for many years.*
- b) *She believes that the MP ‘is a kind and well-mannered person, he could not do this’.*
- c) *There is a reason to accept that the MP could not have done so.*

1. Does she have reliable information that the MP could not have done this? Doubtful.

2. Is Tamara Pletniova trustworthy as a reliable source of information? No. She did not witness what was happening in the MP’s office and tried to victimize the Journalists. “These girls journalists should have behaved themselves, dressed better, ... not walked around with their bare navels [30].”

3. Did she claim that the MP could not have done this? Yes. “He treated women with warmth and never insulted them” [30].

The argument A1.3 score is 0.1 weakly sustainable, because the answers to the critical questions 1, 2 are not compatible with premises *a* and *b*.

A plausible *circumstantial ad hominem* argument A1.3 summarizes the conclusion supported by the Ethics Commission about a planned attack on the MP.

- a) *The Journalists accuse the MP of indecent behavior.*
- b) *The Journalists planned their actions to discredit the MP and the State Duma.*
- c) *The Journalists' accusations of the indecent behavior against the MP are biased.*

1. Are there any incompatible statements among the Journalists’ claims? Yes. The Commission found inconsistent the statements of the Journalists about the acts of the MP’s indecent behavior that had taken place several years ago, with the fact that they have continued to work in the State Duma and complained about those incidents much later, presumably in connection with certain political interests.

2. Did the Journalists manage to explain these incompatible statements and remove doubts about their reliability? Yes. Each of the Journalists complained about the MP independently, they could not plan their actions in advance, since they had never worked together before the filing their charges.

3. Were the Journalists subjected to personal discussion? No, they weren't.

The score of argument A1.3 is 0.1 weakly sustainable due to the incompatibility of the answers to questions 2 and 3 with premises *b* and *c*.

The correctness of A2.1 is ensured by the logical form of the syllogism:

- a) Indecent acts and offers to engage in sexual contact in exchange for work assistance are considered harassment.*
- b) The MP offered the Journalists to engage in sexual contact in exchange for work assistance and committed indecent acts.*
- c) These actions and suggestions of the MP are harassment.*

Premise *a* is based on the definition of the term “harassment” by legal theorists [13]. Premise *b* describes the Journalists' complaints about the MP's obscene suggestions in exchange for interviews and political comments the Journalists had to take according to their editorial assignments. Strongly sustainable A2.1 = 1 was attacked by A1.1 = 1 and defeated.

Argument A2.2 *from position to know* combines the testimonies of the Journalists who met with the MP and acted as informants about these events.

- a) The Journalists met with the MP for work.*
- b) According to them, the MP committed sexual harassment.*
- c) There are reasons to believe that the MP committed sexual harassment.*

1. Are the Journalists in a position to know of the MP's sexual harassment? Yes, they have met with the MP, had vertical relationships with him, and the success of their work in the State Duma depended on him to a large extent.

2. Are the Journalists trustworthy as reliable sources of information? Yes. Apart from the MP, they were the only participants in the incidents after which they privately complained about obscenities to their colleagues. One of the Journalists made an audio recording.

3. Did the Journalists claim that the MP harassed them? Yes. They described the details of his obscene suggestions and actions [8].

The score of argument A2.2 is strongly-sustainable 1, undefended, attacked by weakly-sustainable A1.2, and then – defended by the counterattack of A2.3 on A1.2, and at the end of the Dispute B – again not defended due to the attack of B4.2 on A2.3.

Argument A 2.3 *from expert opinion*, put forward by MP Oksana V. Pushkina, consists of two arguments: *a – c* refers to the common opinion expressing one of the opposing views in the deep disagreement between A2+B4 and A1+B3; *c – f* refers to the expert opinion of Pushkina herself.

- a) The line between courtship and harassment is known to everyone.*
- b) If it is known to everyone, then it is also known to the MP.*
- c) Thus, it is reasonable to assume that the MP knows the line between courtship and harassment.*
- d) Oksana Pushkina is an expert in the issues of discrimination against women.*
- e) She believes that the MP behaved indecently.*
- f) Thus, there are reasons to believe that the MP behaved indecently.*

1. What is the basis of the claim that everyone knows the line between courtship and harassment? It is based on the professional experience of Oksana Pushkina, as her public web-pages suggest.

2. Is there any reason to doubt that everyone knows the line between courtship and harassment? Yes. B3.1-3 indicates the deep disagreement on this issue.

1. How reliable is Oksana Pushkina as an expert? Very reliable. For many years she has worked with women's harassment petitions¹².

2. Is Oksana Pushkina an expert on harassment? Yes.

3. From which of her statements it follows that the MP behaved indecently? "People are held back by fear in 99% of the harassment cases. They stay muted. I < ... > realized how great the girls Journalists who decided to tell their stories were. After all, this is considered indecent in our country. And in general, this cannot happen in our society [17]."

4. Does Pushkina personally deserve trust as a source of opinion? Yes, she is a political activist experienced in overcoming discrimination against women [31].

5. Is Pushkina's statement *a* about the indecent behavior compatible with the statements of other experts? No, it is incompatible with the opinion of expert Chernigovskaya in B3.1.

6. On what evidence are Pushkina's statements based? They are based on her experience in evaluating the testimonies of victims of harassment.

The overall score of argument A2.3 is 0.5, averagely sustainable. A2.3 rejects weakly sustainable A1.2, but B4.2 rejects A2.3, returning A1.2 as defended.

A2.4 is an argument *from negative consequences* contained in the Journalists' objections to A1.3 [23].

a) If we, the Journalists, immediately complained about the harassment of the MP to the court or to the State Duma officials, this would lead to a deterioration of our reputation, and not to a fair consideration of the case.

b) We shouldn't have complained about the MP's harassment to the court or to the State Duma officials.

1. What is the basis of statement *a*? One of the Journalists consulted with experts and found out the legal and moral aspect of the problem of harassment: the lack of appropriate rule of law and the humiliating status of victims of harassment in Russian society.

2. How probable is that the negative consequences will happen? The comment of the head of the Moscow Union of journalists confirms this forecast [33].

3. Are there any positive consequences of an immediate complaint about the MP's behavior that should be taken into account? Yes. The Ethics Commission stated that the issue could have been considered without unnecessary publicity, without suspicion of the applicants' special motives, if it had been filed at a different time.

The overall score of the argument is 0.5, averagely sustainable, because it does not give a conclusion-compatible answer to question 3. A2.4 rejects A1.3, then is counterattacked by B4.3, which returns A1.3 as defended.

Three plausible arguments were put forward in defense of B3: B3.1 *from popular opinion*, B3.3 *from popular practice* and B3.2 *from expert opinion*. Let begin with B3.1.

a) It is believed that a man, giving a woman signs of attention, is not trying to harm her.

b) If showing signs of attention by a man to a woman is not considered an attempt to harm her, then there are reasons for this.

c) There are reasons to believe that the MP who paid attention to the Journalists did not mean to harm them.

d) Thus, the MP by showing signs of attention to the Journalists, did not mean to harm them.

1. What are the reasons to consider *a* as a popular opinion? Personal views of the authors of the argument, confirmed by the results of sociological surveys [34].

2. Is there any reason to doubt that statement *a* is a popular opinion? No. The fact that *a* is a popular opinion is confirmed by the positions of MPs O. Pushkina and A. Veller, who defended B4.

The overall score of B3.1 is strongly sustainable 1.

B3.2 together with B3.3. express the social aspect of deep disagreement. Consider B3.2 first.

a) T. Chernigovskaya is an expert in the field of human behavior.

b) T. Chernigovskaya claims that for men to court women, and for women to resist men's courtship, is normal behavior.

c) It is reasonable to assume that men's courting and women's resisting it is normal behavior.

1. How reliable is T. Chernigovskaya as an expert? Very reliable.

2. Is T. Chernigovskaya an expert in the field of human behavior? Yes.

3. Which Chernigovskaya's statement implies that such behavior is normal for men and women? "Everything I know about humans – anthropologically, physiologically, psychologically, linguistically, cognitively – tells me that this is [blurring the boundaries between male and female, including the rejection of "courtship"] it's a very bad road" [7].

4. What is the basis of T. Chernigovskaya's statements? They are based on her retelling and interpreting of research results.

5. Does T. Chernigovskaya personally deserve trust as a source of opinion? Not quite. In public lectures and talks, she often expresses her personal opinion and emphasizes her disagreement with other positions [24].

6. Are T. Chernigovskaya's statements compatible with the opinions of other experts? No, they are incompatible with the expert opinion of O. Pushkina in A2.3.

The overall score of B3.2 is averagely sustainable 0.5, due to the answers to critical questions 4-6 that are incompatible with *a*.

B3.3 refers to *popular practice* and summarizes the views of several MP in the public discussions.

a) Showing attention to women is a common behavior in Russia.

b) If this behavior is common, it is acceptable.

c) It is reasonable to believe that showing attention to women is acceptable behavior.

1. What is the basis of statement *a*? It is based on the personal views of the authors of the argument.

2. What are the grounds for considering the behavior described in *a* as acceptable? Unknown. We did not find any opinion polls confirming its acceptability.

The overall score of B3.3 is averagely sustainable 0.5.

In defense of B4, three plausible arguments were put forward: B4.1 *from popular opinion*, B4.2 *from correlation to cause*, and B4.3 *slippery slope* – a subspecies of the argument *to negative consequences*.

B4.1 attacks B3.2 on the social aspect of the deep disagreement:

a) Sexual harassment is considered unacceptable.

- b) *If sexual harassment is considered unacceptable, there are reasons for that.*
- c) *It is reasonable to consider sexual harassment unacceptable.*

1. What is the basis of statement *a*? It is based on the personal views and life experience of the author of the argument A. Veller: “any manifestations of sexual harassment are bad... the majority in our society shares this simple moral attitude,”[25] which is confirmed by the results of opinion polls [15].

2. Is there any reason to doubt *a*? Yes. The problem of harassment at workplace is considered important by 20% of Russians, and 63% believe that by their appearance women can "invite for" harassment, 36% support discussing harassment cases, 40% are against [34].

The overall score of B4.1 is averagely sustainable 0.5.

B4.2, the key argument supporting B4, places the B4 in between of the poles in the deep disagreement. In B4.2, the claim about the causal connection of the two phenomena is inferred out of the statement of a correlation noticed between them. B4.2 attacks B3.3 and indirectly A2.3:

- a) *Different people have diverse moral attitudes, so different people may view the same action diversely.*
- b) *The diversity in moral attitudes is the reason why the same action is understood as both courtship and harassment.*

1. Is there a correlation between the difference in people’s moral attitudes and their assessments of the actions of others? Yes, according to A. Veller.

2. Is there any reason to believe that this relationship is not a coincidence? Yes. The Journalists and many MPs recognized the manners of the MP as peculiar, but the Journalists considered his behavior in the situations under consideration as obscene, and the Ethics Commission did not.

3. Is there a third phenomenon that causes differences in the moral attitudes of people and in their assessments of the actions of others? Yes, it may be the involvement of the Journalists in the situation under consideration.

The overall score of B4.2 is averagely sustainable 0.5.

B4.3 is a *slippery slope* argument which justifies its conclusion B4 by the alleged causal chain of negative consequences that adopting of an opposite view would trigger. B4.3 reinforces B4's intermediate position in the deep disagreement:

- a) *If harassment conflicts are not resolved immediately, then attempts to resolve them retroactively will have the worst consequences.*
- b) *If harassment conflicts are not resolved immediately, they will not be resolved retroactively due to diversity in moral views, which entails the risk of provocation or scandals.*
- c) *Provocations and scandals due to harassment conflicts are the worst consequences.*

1. Which intermediate correlations in the event chain leading to *c* are explicitly specified? Four correlations are clearly identified: *c0* – a conflict with harassment is not immediately resolved; *c1* – it is impossible to resolve such a conflict retroactively; *c2* – a difference in moral views prevents such a conflict from being resolved retroactively; *c3* – provocations and scandals.

2. What intermediate correlations are missing in the chain of events *c0*, ..., *c3* for the validity of the transition from *c0* to *c3*? At least three. There are not enough reasons why it is impossible to resolve harassment conflicts retroactively, and *c2* is not a sufficient reason for this, as it suggests no

facts to justify itself. The chain contains no reference to the cases when attempts to retroactively resolve harassment conflicts failed. Neither in the immediate nor retroactive perspective B4.3 proposes definitive symptoms of actions that would indicate an action at question is or is not a harassment in social or legal terms.

3. Are there weak correlations in the chain c_0, \dots, c_3 ? Yes. These are c_1 , where there are no examples of any attempts to resolve such conflicts, successful or not; and c_2 , where differences in moral views are assumed to affect the resolution of harassment conflicts but it is unclear whether they hinder or contribute to that amid no facts of those resolution are described.

The overall score of argument B4.3 is weakly sustainable 0.1.

8. Conclusion

We analyzed the argumentation in the discussion on the socially sensitive issue of harassment using the hybrid method based on the new dialectic, the argumentation logic and the logical-cognitive approach to argumentation which merges the human-oriented stance in argumentation studies of the former with the logic-oriented stance of the latter. In this vein, the discussion was reconstructed with the help of digitally visualized dispute mapping, often employed in dialectical analysis of argumentation, instead of formalization on graphs used in the argumentation logic. The resolution of the discussion was determined by means of the algorithm specially designed in the logical-cognitive approach, which enabled us to achieve the following results. We identified the strongest conclusions in the discussion by the appraisal of the ability of the arguments proposed in support of them to both tolerate other opinions and withstand counter-argumentation and criticisms, which was a planned result of our case study, and revealed the deep disagreement about the limit of permissible courtship, which came up as a bonus result. These results demonstrate efficacy and efficiency of our hybrid method.

In the case-study, we reconstructed the discussion about harassment conflict in the State Duma between the MP and the Journalists who accused him of sexual misconduct. We divided the discussion into two Disputes A and B, the ultimate resolution of which amounted to the arguments in favor of A1+B4 that were not defeated in the Disputes. The resolution consisted in the two conclusions that the Journalists could not prove their accusations of the MP's indecent behavior which echoes the verdict of the Ethics Commission, but such actions are harassment and cannot be tolerated, which is inconsistent with both the resolution of Dispute B and the verdict. The inconsistency of the resolution A1+B4 is one of the symptoms of the deep disagreement revealed in the course of the study. It arose over the issue that initially did not belong to the agenda of the discussion, but the parties' contrary opinions about it came out as the core question of the conflict. Due to cognitive and psychological reasons, the polarization of opinions about harassment is likely to increase amid the absence of a legal definition of harassment and the blurring of the social boundary of the permissibility of courtship.

Acknowledgements

The support from the Russian Science Foundation, project No. 20-18-00158 'Formal philosophy of argumentation and a comprehensive methodology for the search and selection of the dispute resolutions' is kindly recognized.

The paper relies on some of the results by undergraduate students Ekaterina Balabanova, Anastasiya Vinogradova and Anastasiya Shevarenkova in reconstruction of the discussion about the harassment conflict in the State Duma in the fall of 2020, when that was their group assignment in the course "Theory of moral argumentation" at St. Petersburg State University. On top of that, I am

indebted to A. Shevarenkova for her kind assistance in collecting and analyzing the relevant media publications and to graduate student Nina Perova for her help with the English version of the paper.

References

1. Amgoud, L., Doder, D., Vesic, S. Evaluation of argument strength in attack graphs: Foundations and semantics, *Artificial Intelligence* 302, 2022, 103607.
2. Balabanov, S. S., Saraliev, Z. H. Seksual'nye domogatel'stva na rabote v Rossii [Sexual harassment at work in Russia], *Bulletin of the Nizhny Novgorod University. N.I. Lobachevsky. Series Social Sciences* 1 (17), 2010, pp. 7-12. (In Russian)
3. Baroni, P., M. Caminada, M. Giacomin. An introduction to argumentation semantics, *Knowledge Engineering Revue* 26 (4), 2011, pp 365–410.
4. Barth, E.M., Krabbe, E.C.W. *From Axiom to Dialogue*, Berlin - New York: Walter de Gruyter, 1982.
5. Bex F., Verheij B. Solving a Murder Case by Asking Critical Questions: An Approach to Fact-Finding in Terms of Argumentation and Story Schemes. *Argumentation* 26, 2012, pp. 325–353. DOI 10.1007/s10503-011-9257-0
6. Cayrol, C., Lagasquie-Schiex, M. C. Graduality in Argumentation, *Journal of Artificial Intelligence Research* 23, 2005 DOI: <https://doi.org/10.1613/jair.1411>
7. Chernigovskaya, T., Shevardnadze, S. *This is how the world works: a young lady must resist, and a man must persevere*. Spletnik. 08.03.2018. URL: <https://spletnik.ru/buzz/interview1/81128-tatyana-chernigovskaya-interview.html> (In Russian)
8. Chernyansky, P. *Daria Zhuk, a TV Rain producer, openly accused MP Slutsky of harassment. Daria Zhuk filed a complaint with the Ethics Commission of the State Duma*. The Village. 02.03.2018. URL: <https://www.the-village.ru/city/news/304107-zhuk-vs-slutsky> (RUS)
9. Dung, P.M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games, *Artificial Intelligence* 77, 1995, pp. 321-357.
10. Eemeren, F.H. van, Grootendorst, R. *A Systematic Theory of Argumentation*, Cambridge University Press, 2004.
11. Feldman, R. Deep Disagreement, Rational Resolutions, and Critical Thinking, *Informal Logic* 25 (1), 2005, pp. 13-23.
12. Fogelin, R.J. The Logic of Deep Disagreement, *Informal Logic* 25 (1), 2005, pp. 3–11.
13. Kharitonov, M. M. Ponjatie seksual'nogo domogatel'stva (harassmenta) i mehanizmy protivodejstvija emu v trudovom prave Rossii [The concept of sexual harassment and the mechanisms of countering it in the labor law of Russia], *Law. Journal of the Higher School of Economics* 3, 2019, pp. 52-75. (In Russian)
14. Lisanyuk, E. *Argumentation and conviction*. Saint-Petersburg: Nauka, 2015. (in Russian)
15. Lisanyuk, E. Argumentation, Pavilionis's meaning continuum and the Kitchen Debate, *Problemos* 88, 2015, pp. 95 -113.
16. Lisanyuk, E., Mazurova, M. Argumentation, peer-disagreement and birth of truth in dispute [Argumentaciya, raznoglasiye ravnyh I rozhdeniye istiny v spore]. *Epistemology & Philosophy of Science* 56 (1), 2019, pp. 81–100. DOI: 10.5840/eps20195619 (in Russian)
17. Nelyubin, N. MP Oksana Pushkina: 'Our men find it easier with 'fish''. Fontanka.ru. 02.03.2018. URL: <https://www.fontanka.ru/2018/03/02/035/> (In Russian)
18. Olsson, E. J. A Bayesian Simulation Model of Group Deliberation and Polarization, in F. Zenker (ed.), *Bayesian Argumentation*. Springer, 2013.
19. Ostrovskaya, Y., Miryasova, O. Problema harassmenta v vysshej shkole [The problem of harassment in higher education], *Friedrich Ebert Foundation. Russia, 2021*, URL: <https://www.fes-russia.org/meroprijatija/publikacija-problema-kharassmenta-v-vysshei->

- shkole/?fbclid=IwAR2h7eVFEed4KGIk89_LGjTSnh-gnYXWP4G6NLYzNltBWH-7Pfwyvy* (Accessed: 03.05.2021) (In Russian)
20. Prakken, H. Formalising debates about law-making proposals as practical reasoning, in Araszkievicz, M., Pleszka, K. (eds.). *Logic in the Theory and Practice of Lawmaking*, Legisprudence Library Series, Springer, 2015.
 21. Prakken, H., Vreeswijk, G. Logical systems for defeasible argumentation, in: Gabbay, D., Guenther, F. (Eds.), *Handbook of Philosophical Logic*, Kluwer, 2002, pp. 219–318.
 22. Samokhina, S. 'For lack of harassment' // Kommersant. 21.03.2018. URL: <https://www.kommersant.ru/doc/3579606> (In Russian)
 23. Smolyakova, D.T., Chernigovskaya, T. 'Humans share half of their genes with yeast'. SPb Sobaka.ru CИB 29.03.2018 URL: <https://www.sobaka.ru/city/society/70447> (In Russian)
 24. Sokolova, K. 'Harassment is like a virus. And my case is a vaccination against it'. *MP Slutsky on accusation of harassment*. SNOB. 13.06.2018. URL: <https://snob.ru/entry/162002> (In Russian)
 25. Stuchevskaya, O. Harassment i rossijskie zhenshhiny [Harassment and Russian women], *Vestnik obshhestvennogo mnenija* [Public Opinion Bulletin] 4 (96), 2008. (In Russian)
 26. Veller, A. *About the sexual scandal around a colleague in parliament*. SNOB. 26.03.2018. URL: <https://snob.ru/profile/31652/blog/135611> (In Russian)
 27. Walton, D., Reed, Ch., Macagno, F. *Argumentation schemes*, Cambridge UP, 2008.
 28. Wells, S., Reed, Ch. Testing formal dialectic, *Lecture Notes in Computer Science*, ed. Parsons S. et al., vol. 4049\2006, Berlin: Springer Verlag, 2006, pp. 74-87.
 29. Wittgenstein, L. *On Certainty*, Oxford: Basil Blackwell, 1969.
 30. Wittgenstein, L. 'He always treated women with warmth.' *The head of the State Duma committee stood up for Slutsky*. Esquire. 07.03.2018. URL: <https://esquire.ru/novosti/45002-warm-slutsky> (In Russian)
 31. Wittgenstein, L. 'MP Pushkina prepares amendments on sexual harassment amid scandal with MP Slutsky' Kommersant. 27.02.2018. URL: <https://www.kommersant.ru/doc/3559345> (In Russian)
 32. Wittgenstein, L. *Oksana Pushkina proposed to deprive politicians of their positions for harassment*. RBC. 22.03.2018. URL: <https://www.rbc.ru/politics/22/03/2018/5ab3eeee9a79477dd47ae462> ; (In Russian)
 33. Wittgenstein, L. *The head of the Union of Journalists of Moscow: 'The Journalists who complained about the harassment of the MP, have humiliated their profession'*. Fontanka.ru. 07.03.2018. URL: <https://www.fontanka.ru/2022/02/18/70454591/> (In Russian)
 34. Wittgenstein, L. We consider [*Project My schitaem*]. Mikhailov and partners. Analytica. 2019. URL: <https://m-p-a.ru/genderinequality.html> (Accessed: 16.03.2021) (In Russian)
 35. Wittgenstein, L. *The Ethics Commission of the State Duma found no violation of code of conduct in Slutsky's behavior*. Interfax 21.03.2018. URL: <https://www.interfax.ru/russia/604626> (In Russian)

Argumentation-Based Logic for Ethical Decision Making

Sofia Almpani

National Technical University of Athens,
9, Iroon Polytechniou Str.,
157 73 Zografou, Athens, Greece

e-mail: salmpani@mail.ntua.gr
<https://orcid.org/0000-0003-4823-479X>

Petros Stefaneas

National Technical University of Athens
9, Iroon Polytechniou Str.,
157 73 Zografou, Athens, Greece

e-mail: petros@math.ntua.gr
<https://orcid.org/0000-0002-2096-9914>

Panayiotis Frangos

National Technical University of Athens,
9, Iroon Polytechniou Str.,
157 73 Zografou, Athens, Greece

e-mail: pfrangos@central.ntua.gr
<https://orcid.org/0000-0002-8607-5737>

Abstract:

As automation in artificial intelligence is increasing, we will need to automate a growing amount of ethical decision making. However, ethical decision-making raises novel challenges for engineers, ethicists and policymakers, who will have to explore new ways to realize this task. The presented work focuses on the development and formalization of models that aim at ensuring a correct ethical behaviour of artificial intelligent agents, in a provable way, extending and implementing a logic-based proving calculus that is based on argumentation reasoning with support and attack arguments. This leads to a formal theoretical framework of ethical competence that could be implemented in artificial intelligent systems in order to best formalize certain parameters of ethical decision-making to ensure safety and justified trust.

Keywords: argumentation, argumentation logic, decision-making, ethics, artificial intelligence.

1. Introduction

As autonomous artificial intelligent (AI) systems take up a progressively prominent role in our daily lives, it is undoubtedly that they will sooner or later be called on to make significant, ethically charged decisions and actions [6]. Over the last years, the issue of ethics in artificial intelligence has gained great attention and many important theoretical and applied results were derived in the perspective of developing ethical systems [25]. But how could any AI agent be considered ethical? Some of the requirements needed are a broad capability to envisage the consequences of its own decisions as well as an ethical policy with rules to test each possible decision/consequence, so as to choose the most ethical scenario [25], [8]. The challenge is how we can guarantee that AI agents will always perform an ethically correct behavior as defined by the ethical code declared by their human supervisors. Argumentation reasoning can be used as a tool for the formal ethical development and justification of an AI system using the support and attack relationships of arguments and counter-arguments.

Moral reasoning is a key issue in AI ethics, and computational formal proofs are perhaps the single most effective tool for determining credible and trustful reasoning [9]. This work attempts to develop a *Moral* extension of the *Argumentation-based Proof Event Calculus* [3] (*MAPEC*) by integrating the ethical framework from [9] and the moral competence from [20] to develop a formal representation of ethical scenarios and integrate moral norms and concepts that are supported through argumentation (See Fig.1). A detailed description of the initial argumentation-based proof event calculus can be viewed in [3], [2].

For the realization of this effort, the objectives are:

- to formalize what it means for an AI agent's decision-making to be ethically correct;
- to provide a logical specification with which the system can be built and checked;
- to extend Argumentation-based Proof Event-Calculus to create an abstract Moral framework (MAPEC) with ethical logic-based argumentation.

The paper has four sections. Section 2 describes the theoretical background of AI ethics and formal reasoning systems. Section 3 outlines the formalization of ethical events in terms of argumentation theory. Section 4 concludes with an overview of this paper.

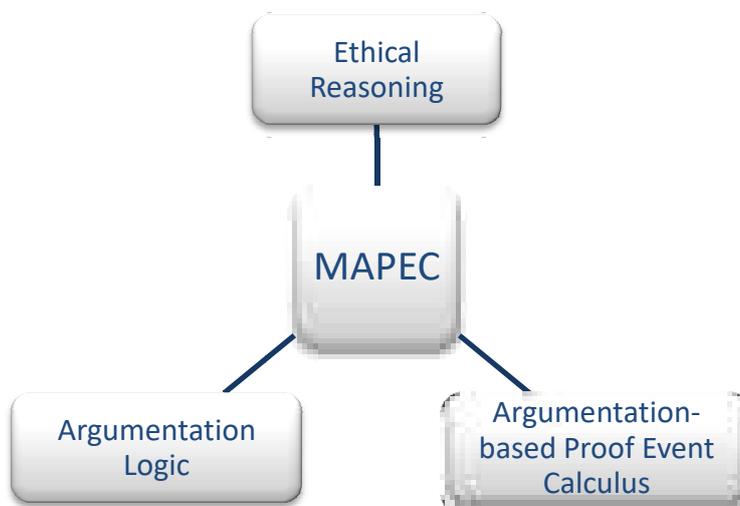


Figure 1: Research framework of MAPEC.

2. Theoretical Background

Academic research and real-life incidents of AI system failures and misuse have indicated the need for employing ethics in AI systems development [6]. Nevertheless, studies on methods and tools to address this need in practice are still lacking, resulting in a growing demand for AI ethics as a part of engineering [26]. But how can AI ethics be integrated in engineering projects when they are not formally considered? There has been some work on the formalization of ethical principles in AI [10]. Previous studies that attempt to integrate norms into AI agents and design formal reasoning systems has focused on: ethical engineering design [12], [27], [28] norms of implementation [15], [24], moral agency [13], [7], mathematical proofs for ethical reasoning [6], logical frameworks for rule-based ethical reasoning [1], [4], [16], reasoning in conflicts resolution [22], and inference to apply ethical judgments to scenarios [5].

One of the categories of AI ethics is Ethics by Design, which is the incorporation of ethical reasoning abilities as a part of system behavior, such as in ethical AI agents [26]. In this work, if we assume that an AI agent can be capable of ethical agency, the purpose is to enable AI agents to *reason ethically* [9] implementing argumentation reasoning. This includes taking into consideration societal and moral norms; hierarch the respective priorities of norms in various contexts; explain its reasoning with logical arguments; and secure transparency and safety [11]. These systems are often established with the purpose to assist ethical decision-making by people, identifying the ethical principles that a system should not violate [9].

In an autonomous system, it is not aimed to show that an agent always follows the moral thing, but that its actions are taken for the right reasons. In many real life scenarios, it is not easy to provide a complete set of decisions that will cover all situations [9]. Therefore, the system may have two modes of operation; either it uses its pre-existing set of arguments and actions in conditions which are within its anticipated parameters; or when new options appear it acts outside of these parameters based on various available resources that allow governing its actions using ethical reasoning [9].

3. A Formal Logic-Based Framework for Ethical Reasoning

To represent ethical codes and rules it requires an *ethical policy*, a hierarchy over the rules that are appropriate in different contexts (defining even which rule is more acceptable to violate when no ethical option is available). In order to demonstrate that a system has the property of making the right decisions (both operationally and ethically), it should be formally specified what the “right decisions” are.

Formal verification [21] includes proving or disproving that a system is compliant with a requirement determined in a mathematical language, i.e., a “formally specified property” expressed within a linear temporal logic, which in our case allows us to define what decisions should the rational agents made at some specific moment [9]. Thus, the ethical policy can be formalized in some computational logic L, whose well-defined formulas and proof theory specify the basic concepts required: the temporal structure, events, actions, sequences, agents, and so on [6]. The presented methodology proof-theoretically formalizes the ethical policy and implements it, meaning that this methodology encodes not the semantics of the logic L but its proof calculus [6].

Logic-based systems that are capable of dealing with increasing degrees of environmental uncertainty and variability are preferable [14] and argumentation constitutes a way to deal with an undefined and uncertain world, meaning not necessarily a chaotic one but just a complex one. Argumentation is a tool of cognition that can formalize the science of common sense reasoning on which new types of systems can be engineered [17].

Therefore, to address the challenge of ensuring ethically correct behavior, a logic-based argumentation approach such as MAPEC is proposed to guarantee that AI agents only execute events that can be proved ethically acceptable in a human-selected logic, by formalizing an ethical code [6].

3.1. Ethical Events Expressed Within an Argumentation-Based Framework

In an ethical framework, a moral vocabulary allows the agent to represent norms, ethically substantial behaviors, and their judgments (conceptually and linguistically) in order to fuel the moral communication. It contains: a *normative frame* referring to the features of norms and to the normatively-supported qualities of agents; a language of norm *violation* characterizing attributes of violations and of violators; and a language of *responses to violations* [20].

In our approach, the concept of norms is described with events, extending their context to **abstract ethical events**. The abstract ethical events present the arguments in a moral debate. The violations are analogous to the counterarguments. The role of ethical agents can be easily depicted as akin to the role of the supporter (or prover) and attacker in our argumentation framework [2], where the supporter plays the role of the ethical correct agent and the attacker the role of the violator. Their actions are the responses to moral violations with arguments or counterarguments. Moral communication expresses agent's efforts to recognize, clarify, or defend norm events, as well as interfere or rectify after a norm violation.

Definition 1: Abstract Ethical Events

An abstract ethical event is represented with argument e and its purpose is to defend an ethical principle c . The c can be interpreted also as "the supporter considers it immoral to permit or cause $\neg c$ (to happen)". The *Abstract Ethical Event* has the same structural components (data Φ , warrant w , ethical claim c) as a proof event in APEC [3]. Thus, an ethical argument e is in force when the event concludes to c , based on the data Φ and following the inference rules w and it has the following internal structure:

$$e \ c \ < \text{communicate} \ < \Phi, c \ >, w \ > ,$$

where $e \in E$, E the set of ethical events for the c . This means that an abstract ethical event refers to a fixed ethical *principle* specified by certain *data*, justified with a *warrant* that is based on ethical reasoning and a system of norms. Similarly, counter-argument e^* denotes the *violation event*.

A system of norms contains a society's principles for ethical behavior. They guide supporter's arguments and decisions to behave with specific (moral) actions and shape others' (moral) judgments of those behaviors [20]. Thus, they establish *an ethical policy with ethical rules*.

Definition 2: Ethical Policy

An ethical policy P is a tuple $P = \langle R, \geq \rangle$ where R is a finite set of ethical rules between the events e , with $e \in E$, and \geq is a complete (not necessarily strict) priority order on R . The expression $e_1 = e_2$ indicates that violating argument e_1 is equivalently unethical as violating argument e_2 , while $e_1 \geq e_2$ denotes that violating e_1 is equally or less unethical to violating e_2 . A special category of ethical event, symbolized as e_0 , is vacuously satisfied and encompassed in every policy so that $\forall e \in E: e > e_0$, indicating it is always strictly more unethical to do nothing and permit any of the unethical conditions to happen.

Moral action is an event, taking place in compliance with the norms and in specific time, which is accommodated to and harmonized with other social agents (violators or provers) who operate under the same context. The norm violations e^* of a violator are denoted as **attack**(e^*, t) events and the ethical proving action of a supporter are denoted as **support**(e, t), specified both by the time t to express the temporal sequence of the actions.

Definition 3: Ethical Actions

Given a certain context \mathbf{a} , an event \mathbf{e} , and an ethical principle \mathbf{c} , an ethical action can be the formulas:

$\text{support}(\mathbf{e}, \mathbf{t}) \stackrel{\mathbf{a}}{\Rightarrow} \mathbf{c}$, denoting the actions of a supporter to defend the ethical principle \mathbf{c} with ethical event (argument) \mathbf{e} in context \mathbf{a} and at time \mathbf{t} .

$\text{attack}(\mathbf{e}^*, \mathbf{t}) \stackrel{\mathbf{a}}{\Rightarrow} \neg\mathbf{c}$, denoting the actions of a violator to contravene the ethical principle \mathbf{c} with violation (counter-argument) \mathbf{e}^* in context \mathbf{a} and at time \mathbf{t} .

3.2. Prioritized Ethical Rules to Define Context-Based Scenarios

Context determines dynamic priorities on the decision policies of the agent [18]. To be able to reason about scenarios in terms of ethics we need a scenario selection process that uses the ethical policy, which can be represented within the argumentation theory. The agent can be in various contexts while deciding which scenario to choose, so the rules from all the contexts need to be considered when implement a plan. We advocate scenarios that are ethical or at least violate the fewest ethical principles, both in quantity and in severity.

The scenarios are ordered using $<$ which leads to a complete order over scenarios [9]. This can describe an agent's ethical policy based on the different contexts with **argumentation levels**. In the first level we have the rules that refer directly to the domain of the agent, the *object-level decision rules*. In the other *priority levels* the rules relate to the ethical policy under which the agent generates different possible scenarios that the agent can choose. In the *higher level priority* there are the rules representing the optimal course of action, the more ethical (or less unethical) scenario [18].

Definition 4: Levels of Ethical Rules

Given a policy $\mathbf{P} = \langle \mathbf{R}, \geq \rangle$ and a plan based on the ethical rules \mathbf{R} , \mathbf{V} is a set of abstract ethical events (including the events \mathbf{e} and the violations \mathbf{e}^* of the ethical principles \mathbf{c}) defined as:

$$\mathbf{V} = \langle \mathbf{e} \mid \mathbf{e}(\Phi, \mathbf{c}), \mathbf{e} \in \mathbf{E}, \text{support}(\mathbf{e}, \mathbf{t}) \stackrel{\mathbf{a}}{\Rightarrow} \mathbf{c} \rangle$$

In this set, we include all the ethical rules and ethical events \mathbf{e} that can be used to support an ethical principle \mathbf{c} . The aim is to create a priority between sets of ethical events, where a higher set means that includes more ethically important events in terms of moral values and norms. Thus, we define the operation *Higher* for the higher level of ethical scenarios \mathbf{L} based on the set of events \mathbf{V} , as follows:

$$\mathbf{L} = \text{Higher}(\mathbf{V}) = \{ \mathbf{e} \mid \mathbf{e} \in \mathbf{V}, \text{and } \forall \mathbf{e}_n \in \mathbf{V} : \mathbf{e} \geq \mathbf{e}_n \}$$

Consider a set of available, possibly ethical, scenarios \mathbf{L}_i for the different set of \mathbf{V}_i . The scenarios lead to different levels of ethical rules $\mathbf{L}_i \in \mathbf{L}$ that satisfies the following properties, in order to define using arguments $\mathbf{e}_n, \mathbf{e}_n \in \mathbf{E}$, which available scenario is more ethical (or less unethical). For every $\mathbf{i}, \mathbf{j} \in \mathbf{N}$, it holds that $\mathbf{L}_i \succ \mathbf{L}_j$ if at least one of the following holds:

1. $\mathbf{V}_i = \emptyset$ and $\mathbf{V}_j \neq \emptyset$.
2. $\mathbf{e}_1 \geq \mathbf{e}_2$ for every $\mathbf{e}_1 \in \text{Higher}(\mathbf{V}_j \setminus \mathbf{V}_i)$ and every $\mathbf{e}_2 \in \text{Higher}(\mathbf{V}_i \setminus \mathbf{V}_j)$
3. $\mathbf{e}_1 = \mathbf{e}_2$ for every $\mathbf{e}_1 \in \text{Higher}(\mathbf{V}_j \setminus \mathbf{V}_i)$, and every $\mathbf{e}_2 \in \text{Higher}(\mathbf{V}_i \setminus \mathbf{V}_j)$, while $|\text{Higher}(\mathbf{V}_j \setminus \mathbf{V}_i)| < |\text{Higher}(\mathbf{V}_i \setminus \mathbf{V}_j)|$.

If none of them holds, then \mathbf{L}_i and \mathbf{L}_j are equally (un)ethical, i.e., $\mathbf{L}_i \sim \mathbf{L}_j$.

The first relation makes sure that the ethical scenarios will always be favored by the unethical ones. The second one guarantees that when the principles that are the same in both scenarios are ignored, then the argument that defends the most valuable principle is considered “higher” ethical. The third states that when the arguments that in each scenario are violated are different, but equally valuable, the plan which violates less in number principles is “higher” ethical.

We can now define a logical property which specifies what it means that the reasoning and the decision-making of an agent are ethical. Informally, we have that whenever an agent selects a scenario, L_i , then all other applicable scenarios L_j should be ethically “lower”, *i.e.*, that $L_j < L_i$.

4. Conclusions

This work attempted to develop a proof-theoretical representation of norm scenarios and integrate ethical concepts into a system by developing a logic-based argumentation calculus. *Moral Argumentative Proof-Events Calculus* (MAPEC) is a framework to help stakeholders to various AI project build an ethics roadmap in a methodical way. This framework can present ethics foresight early in the deployment procedure, rather than implement it as an auditing or assessment tool. There are *three main stages* in this procedure which includes the interaction of three aspects (agents, ethical principles, and contexts):

1. identify the normative frame and the agents;
2. define the ethical events-arguments and rules for different scenarios; and
3. prioritize the ethical rules to define the order of scenarios.

The aim of this study is to establish that an ethical policy can be combined within an AI agent in such a way that the dedication to the policy can be *formally* verified and so it can be checked that the agent will always choose the most ethical decisions justified with arguments. The next step, in future research, is to build algorithms that can computationally capture ethical cognition and actions with formal decision-making that not only take ethics into consideration when reasoning but can be also proved with solid arguments.

References

1. Ågotnes T. and Wooldridge, M. Optimal social laws. in *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, 1, 2010, pp. 667–674.
2. Almpanti, S. and Stefaneas, P. On proving and argumentation. *AIC 2017, 5th International Workshop on Artificial Intelligence and Cognition*, Larnaka, 2017.
3. Almpanti, S. and Stefaneas, P. and Vandoulakis, I. On the Role of argumentation in discovery proof-events. *C3GI 2017, 6th International Workshop on Computational Creativity, Concept Invention, and General Intelligence*, Madrid, 2017.
4. Arkin, R. *Governing Lethal Behavior in Autonomous Robots*. CRC Press, 2009.
5. Blass, J. and Forbus, K. Moral Decision-Making by Analogy: Generalizations vs. Exemplars. *AAAI Conference on Artificial Intelligence; Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, Jan. 2015.
6. Bringsjord, S. and Arkoudas, K. and Bello, P. Toward a General Logician Methodology for Engineering Ethically Correct Robots. *IEEE Intelligent Systems*, 21, 2006, pp. 38–44.
7. Cunneen, M. and Mullins, M. and Murphy, F. and Gaines, S. Artificial Driving Intelligence and Moral Agency: Examining the Decision Ontology of Unavoidable Road Traffic Accidents through the Prism of the Trolley Dilemma. *Applied Artificial Intelligence*, 33 (3), 2019, pp. 267–293.
8. Danaher, J. Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism. *Sci Eng Ethics*, Jun. 2019.
9. Dennis, L. and Fisher, M. and Slavkovik, M. and Webster, M. Formal verification of ethical choices in autonomous systems. *Robotics and Autonomous Systems*, 77, Mar. 2016, pp. 1–14.

10. Dennis, L. A. and Fisher, M. and Winfield, A. F. T. Towards Verifiably Ethical Robot Behaviour. *arXiv:1504.03592 [cs]*, Apr. 2015, Accessed: Jul. 02, 2020. [Online]. Available: <http://arxiv.org/abs/1504.03592>.
11. Dignum, V. Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20, 2018, doi: 10.1007/s10676-018-9450-z.
12. Flanagan, M. and Howe, D. and Nissenbaum, H. Embodying values in technology: Theory and practice. *Information Technology and Moral Philosophy*, 2008, pp. 322–353.
13. Floridi L. and Sanders, J. W. On the Morality of Artificial Agents. *Minds and Machines*, 14, 2004, pp. 349–379.
14. Gomila, A. and Müller, V. Challenges for Artificial Cognitive Systems. *Journal of Cognitive Science*, 13, 2012, pp. 453–469.
15. Hofmann, B. Ethical Challenges with Welfare Technology: A Review of the Literature. *Science and engineering ethics*, 19, 2012.
16. Iba, W. and Langley, P. Exploring Moral Reasoning in a Cognitive Architecture. *Thirty-Third Annual Meeting of the Cognitive Science Society, Expanding the Space of Cognitive Science Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, Boston, 20-23 Jul 2011.
17. Kakas, A. and Michael, L. Cognitive Systems: Argument and Cognition. *IEEE Intelligent Informatics Bulletin*, 17, 2016, pp. 14–20.
18. Kakas, A. and Moraitis, P. Argumentation Based Decision Making for Autonomous Agents. in *Proceedings of the International Conference on Autonomous Agents*, 2, 2003, pp. 883–890.
19. Lutz, C. and Tamò Larrioux, A. RoboCode-Ethicists – Privacy-friendly robots, an ethical responsibility of engineers?. *ACM Web Science*, Oxford, 7, 2015.
20. Malle, B. and Scheutz, M. Learning How to Behave: Moral Competence for Social Robots. *Handbuch Maschinenethik*, 2019, pp. 1–24.
21. Fisher, M. and Dennis, L. and Webster, M. Verifying autonomous systems. *Commun. ACM*, 56 (9), Sep. 2013, pp. 84–93.
22. Pereira, L. and Saptawijaya, A. Modelling Morality with Prospective Logic. in *International Journal of Reasoning-based Intelligent Systems*, 1, 2007, pp. 99–111.
23. Robertson, L. J. and Abbas, R. and Alici, G. and Munoz, A. and K. Michael. Engineering-Based Design Methodology for Embedding Ethics in Autonomous Robots. *Proceedings of the IEEE*, 107 (3), 2019, pp. 582–599.
24. Sisk, B. A. and Mozersky, J. and Antes, A. L. and DuBois, J. M. The ‘Ought-Is’ Problem: An Implementation Science Framework for Translating Ethical Norms Into Practice. *The American Journal of Bioethics*, 20 (4), Apr. 2020, pp. 62–70.
25. Tzafestas, S. Ethics in Robotics and Automation: A General View. *International Robotics and Automation Journal (IRAT-J)*, 4, 2018.
26. Vakkuri, V. and Kemell, K.-K. and Abrahamsson, P. Implementing Ethics in AI: An industrial multiple case study. *ArXiv*, abs/1906.12307, 2019.
27. Wynsberghe, A. Designing Robots for Care: Care Centered Value-Sensitive Design. *Science and engineering ethics*, 19, 2012.
28. Winfield, A. F. and Michael, K. and Pitt, J. and Evers, V. Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems [Scanning the Issue]. *Proceedings of the IEEE*, 107 (3), Mar. 2019, pp. 509–517.

Non-Monotonic Reasoning in Medieval Theology: Problems and Assumptions

Marcin Trepczyński

University of Warsaw,
Krakowskie Przedmieście 3,
00-927 Warsaw, Poland

e-mail: m.trepczynski@uw.edu.pl
<https://orcid.org/0000-0003-0612-2597>

Abstract:

Some interesting cases of non-monotonic reasoning have already been identified in medieval theological texts. Jacob Archambault proved in 2015 that the argumentation presented by St Anselm of Canterbury in his *Proslogion* has non-monotonic “embeddings”. My own contribution from 2011 indicated that we can argue that a non-monotonic logic underlies some discussions provided by St Thomas Aquinas in his *Summa theologiae*, and showed that Boethius of Dacia used non-monotonic reasoning in his *De aeternitate mundi*. In this article, I would like to briefly present these examples and verify whether we can speak about similar cases in medieval Biblical exegesis. My aim is to outline particular problems connected with the identification of non-monotonicity which are specific to theology, as well as assumptions that should be adopted to successfully discuss this issue.

Keywords: non-monotonic reasoning, non-monotonic logics, medieval theology, Anselm of Canterbury, Bonaventure, Thomas Aquinas, Boethius of Dacia.

1. Introduction

In a nutshell, reasoning is monotonic when after adding new premises which are consistent with previous ones, we can still draw previous conclusions. Thus, as long as the set of premises grows, the set of conclusions grows too. It can be described by an increasing function which, according to mathematical terminology, is monotonic. However, it may happen that although the set of conclusions generally grows along with the growing set of premises, it suddenly fails to grow. That is, after adding some new premises which are consistent with the previous ones, some conclusions cannot be drawn anymore. Such a reasoning is called non-monotonic.

This label is applied to not only reasoning, but also to such concepts as: inference, theory, deductive system or logic. To quote Drew McDermott, “The word ‘nonmonotonic’ refers to first-order theories in which next axioms can wipe out old theorems” [10, p. 33], or as McDermott put it in an earlier work with Jon Doyle, “‘Non-monotonic’ logical systems are logics in which the introduction of new axioms can invalidate old theorems” [11, p. 41]. According to Dov Gabbay, “the non-monotonic deductive systems are built from rules which rely on \underline{P} and when $\underline{P} \cup \underline{P}'$ is

encountered the old deduction may fail” [5, p. 439]. In practice, the term “non-monotonic logics” may have been coined mainly thanks to Alfred Tarski, who introduced abstract logics as consequence operations. Such logics – as Philippe Besnard puts it – “are often identified with closure operators over sets of formulas of a logical language” [2, p. 77]. A classical consequence operator (classical logic) should satisfy at least the following three conditions, axioms or principles (for: X and Y – sets of sentences and C – the Tarskian consequence operator): 1) $X \subseteq C(X)$ (Reflexivity); 2) $C(C(X)) = C(X)$ (Idempotence); 3) $X \subseteq Y \Rightarrow C(X) \subseteq C(Y)$ (Monotony). If it does not satisfy the last one, such an operator or such a logic is non-monotonic, and thus it is also non-classical. Many systems can be generated in which that third condition is replaced by a weaker one (cf. [8], [2]), like, for example, the condition of cumulativity, as David Makinson showed [7]. Let us underline that it does not mean that such logics or systems necessarily exclude classical inference or consequence. To illustrate: in an “expert system,” Gabbay indicates, “one may take classical logic as the deductive component and some default system as the non-monotonic component” [5, p. 440]. Finally, we should note that we can speak about non-monotonicity in terms of an inference relation between sets of propositions or sentences. It is generally accepted that such a non-monotonic relation is represented by the symbol “ \sim ” instead of the symbol of the classical relation “ \vdash ”. Thus, a non-monotonic relation allows for a situation in which, for X – a set of premises, K – a new set of premises which do not contradict any premise from X , and α – any formula, we accept both: $X \sim \alpha$, and: $X \cup K \not\sim \alpha$.

The idea of non-monotonicity in reasoning, theories or logics has different applications. It can be used to describe everyday reasoning [9], advanced reasoning in difficult circumstances [19], including default logic [13], as well as to model effective decision making for robots, including artificial intelligence [14], [7]. We should also note that the idea of non-monotonicity in reasoning was embodied in the inference theory presented in the Talmud, in tractate *Zevachim* (49b–51a); according to this theory inferences which are generally accepted are not allowed if special configurations of hermeneutic rules on which premises and conclusions are based take place [4], [16].¹ I argue that the idea of non-monotonicity can also be applied to better understand processes of reasoning in theological writings. However, it is not easy to identify non-monotonic reasoning in such texts. It is connected with problems which can lead to serious controversies. I claim that in order to enable a successful debate on this issue, some key assumptions should be formulated.

In the first part of this article, I will briefly present publications which identify non-monotonic reasoning in important medieval theological treatises; however, in one case I will broaden the analysis. Next, I will extend this scope by adding an example from a medieval commentary on the Bible, which I have found recently. On this basis, it will be possible to achieve the main aim of this article – namely, to identify the above-mentioned problems connected with the identification of non-monotonicity which are specific for theology, as well as assumptions enabling farther debate.

2. Non-Monotonic Reasoning in Medieval Theology

2.1. Anselm’s Inconsistent *Proslogion*

In 2015, at the 1st World Congress on Logic and Religion, Jacob Archambault indicated, as he called it, “monotonic and non-monotonic embeddings” of St Anselm of Canterbury’s famous proof² for God’s existence, presented in the *Proslogion*. He elaborated this issue farther in an article published two years later. He demonstrated that the *Proslogion* has a special nature. To quote his conclusion:

The claims of the *Proslogion* should not be read as forming a systematic whole. Instead, in the movement of the work itself, the ascent of Anselm the protagonist sometimes involves a deepening of understanding that modifies or even jettisons claims advanced in earlier parts of the work [1, p. 134].

He gave many examples revealing that Anselm is, indeed, “inconsistent.” He showed that the *Proslogion* can be divided into two parts which are parallel. Chapters 1-13 represent *kataphatic* theology, so a discourse based on positive predication about God, and chapters 14-26 belong to *apophatic* (negative) theology, which reflects serious doubt about the possibility of knowing God and effectively saying anything positive about him, and which limits itself to say what God is not rather than what God is. Furthermore, each chapter has its counterpart: ch. 1 is a prologue, which corresponds to the recapitulation presented in ch. 14; ch. 2-4 contain the proof based on the argument of “that than which a greater cannot be thought,” whereas in ch. 15-17 Anselm tends to show that God is “greater than can be thought, emphasizing his distance even from the understanding of the believer,” Archambault shows [1, p. 131]; ch. 5 considers what God is generally, while ch. 18 “gives this specific content” – what God is actually, and so on. Finally, Archambault argues that this composition based on two parallel parts “exhibits these as two different stages in the spiritual life of Anselm the protagonist” [1, p. 131]. The earlier stage is represented by the *kataphatic* part, and the later by the *apophatic* one. In Archambault’s opinion, “each section of the latter half of the work revisits some theme from its earlier counterpart and modifies it in some way” [1, p. 131]. Thus, the whole work has a dynamic nature, reflecting the author’s development.

However, it does not mean that what was previously stated will be refuted in the second part. “The religious attitude of Anselm the protagonist requires the maintenance of both the affirmation and the denial at different moments” [1, p. 132]. Hence, his spiritual experience forces Anselm to pursue both discourse strategies, and to some extent to create an inconsistent environment which generates non-monotonic reasoning. It appears to be non-monotonic because the former premises are not negated, and the former reasoning is not challenged; however, the new approach of the second part is different and in some cases it is no longer possible to say about God what was said before. To demonstrate it, Archambault presents deductions, using formal notation. He clearly shows how a certain default rule which worked and made it possible to draw a conclusion in the context of ch. 2-4, no longer works in the context of ch. 15. He also points out a similar mechanism in other places [1, pp. 133-134].

To recapitulate, we can see that Anselm’s famous proof indeed has “non-monotonic embeddings.” I believe it is not an overstatement to say that in this way the discourse of the *Proslogion* is governed by a non-monotonic logic. However, in my opinion, the most important observation is that this special nature of Anselm’s work is strongly related to the situation of the theologian. First, because of the author’s spiritual experience and development, as Archambault emphasized. Second, due to the dialectical (in a Hegelian sense) nature of Christian theology, which absorbed two theological traditions which seem opposed to each other, but at the same time work perfectly together, both satisfying the need of talking about God and establishing a proper distance to what can be stated about God, namely: *via positiva* and *via negativa*. In this context, the non-monotonic approach seems natural to Christian theology, and – as we can see – it was adopted, though unconsciously – in the work of one of the most important representants of medieval theology.

2.2. Aquinas’s Non-Monotonic Theological Didactics

In an article published in 2011, I tried to show that some parts of Aquinas’s discourse presented in his *Summa theologiae* (hereinafter: *ST*) testify to the use of a non-monotonic logic [18]. The best example I found was question 3, article 1 from the first part of the treatise, where Thomas considers the problem of “whether God is a body.” In this subsection, I would like to discuss this case in detail. For this reason, it will be relatively longer than the other subsections.

According to a common strategy of 13th-century theological questions and the pattern Aquinas uses in the whole *ST*, at the beginning of q. 3, a. 1, he argues for the starting hypothesis: that God is a body; like always, he presents some of the most important arguments. In this case, he gives five of them, all based on passages from the Bible. Next, he formulates a strong

counterargument and solves the problem in the main part of the article – a section which often starts with the words “I answer that” (*Respondeo*) and which is usually labelled as *responsio* (response) or *corpus* (body – main part). This part is generally based on purely rational analysis and often includes “rational” proofs (that is, ones that do not refer directly to the Christian revelation). In q. 3, a. 1, we find as many as three such proofs. Finally, in the last part Thomas replies to each argument presented in the first part, in the light of his response.

Usually, the most interesting part of Aquinas’s questions, or generally medieval questions, are *corpora*, so the main responses to the problem. In our case, the situation is quite the opposite. The scope of this analysis consists of the arguments for the starting hypotheses and their revisions in the responses to those arguments. We will examine what happens within those revisions and how they refer to the previous arguments. In this situation, the *corpora* are not important. Of course, they refer to the main problem posed in each article, but they do not refer directly to the arguments for the starting hypothesis. They have different starting points, and thus, they do not follow the same thread. We can say the same about the counterarguments to those arguments which challenge them not by discussing their premises, but by offering other premises which give the opposite conclusion. Obviously, all those elements are interesting; however, for the sake of brevity, we will concentrate on the threads found in the arguments for the starting hypotheses and directly followed in the responses to them. Let us go through those five examples.

I mark the premises as “P” with indexes, even when they are both premises and conclusions drawn from other premises, and the final conclusions as “C.” Logical constants and reasoning indicators (like “but,” “for,” “therefore,” etc.) do not belong to the premises themselves; however, in order to make it more readable, I put the markers “P1,” “P2,” etc. at the beginning of each phrase which contains both those constants and a new premise. If there is a string of passages from the Bible, I mark them with one P, as one element, for the sake of brevity, although they could be treated as individual premises. The examples are quoted from the English edition of *ST* without modifications of wording.³ All bolds are mine. In cases when I think it is important to know the original technical term, I give the Latin text in brackets.

The first argument to prove that God is a body runs as follows:

P1. [A] body is that which has the three dimensions.

P2. But Holy Scripture attributes the three dimensions to God.

P3. [F]or it is written: *He is higher than Heaven, and what wilt thou do? He is deeper than Hell, and how wilt thou know? The measure of Him is longer than the earth and broader than the sea* (Job 11:8, 9).

C. Therefore God is a body.

As we can see, P1 and P2 create a syllogism (which can be recognized as the common syllogistic type Barbara) giving C. It is a deductive reasoning, called an inference, and as for its structure, it is formally correct and thus infallible. However, P2 is not a pure piece of information about God’s properties. P2 states that Holy Scripture attributes this property to God. Hence, first, it must be assumed that whatever Holy Scripture says is true (HP1). Here, it is a hidden premise, but in general it is the main assumption of Biblical hermeneutics and Christian theology. Second, P2 must be proven by giving evidence, and this is the role of P3. The set of basic premises consists of P1, P3 and perhaps HP1 as well.

Let us now examine Aquinas’s complete response to this argument:

As we have said above (q. 1, a. 9), Holy Writ puts before us spiritual and divine things **under the comparison** (*sub similitudinibus*) of corporeal things. Hence, when it attributes to God the three dimensions **under the comparison of corporeal quantity, it implies His virtual quantity**; thus, by depth, it signifies His power of knowing hidden things; by height, the transcendence of His excelling power; by length, the duration of His existence; by breadth, His act of love for all. Or, as says Dionysius

(*Div. Nom.* ix), by the depth of God is meant the incomprehensibility of His essence; by length, the procession of His all-pervading power; by breadth, His overspreading all things, inasmuch as all things lie under His protection.

How does Thomas revise the argument? He refers to the hermeneutic rule presented in the first *quaestio* of *ST*, which is a methodological introduction to the whole work. The rule is: “It is befitting Holy Writ to put forward divine and spiritual truths by means of comparisons with material things” (*ST*, I, q. 1, a. 9). Aquinas gives the following reason for such an approach: “For God provides for everything according to the capacity of its nature. Now it is natural to man to attain to intellectual truths through sensible objects, because all our knowledge originates from sense.” Next, he shows that according to this rule the passages quoted in P3 express the incorporeal properties of God, although by comparison to corporeal quantities. Thus, after adding this rule, either it is no longer possible to maintain P2 (if we understand three dimensions as corporal quantities), and thus: to derive C from the set of basic premises, or (if we are open to understand three dimensions in various senses) even if P2 can be upheld, we cannot derive C.

One could claim that Aquinas counters the previous argument. However, the text itself does not confirm such an approach. At this stage, let us note that Thomas adds the rule and shows its implications, but does not say that the premises were wrong or that the inference was invalid or incorrect.

Let us now examine arg. 2:

P1. [E]verything that has figure is a body,

P2. since figure is a quality of quantity.

P3. But God seems to have figure,

P4. for it is written: *Let us make man to our image and likeness* (Gen 1:26).

P5. Now a figure is called an image,

P6. according to the text: *Who being the brightness of His glory and the figure, i.e., the image, of His substance* (Heb 1:3).

C. Therefore God is a body.

This example is more complicated. P1 is proven by P2; P3 by P4 and P5 taken together; P5 by P6. P1 together with P3 create a syllogism (Barbara) giving C as a conclusion. The basic premises are: P2 (which is a definition), as well as P4 and P6 (which are the Biblical passages). It is debatable whether P1 can really be inferred from P2, as it seems it uses some hidden premises which are disputable. However, Aquinas does not refer to this premise and concentrates on P3, which is based on Biblical passages. His complete reply to arg. 2 is the following:

Man is said to be after the image of God, **not as regards his body, but as regards that whereby he excels other animals.** Hence, when it is said, *Let us make man to our image and likeness*, it is added, *And let him have dominion over the fishes of the sea* (Gen 1:26). Now man excels all animals by his reason and intelligence; hence it is according to his intelligence and reason, which are incorporeal, that man is said to be according to the image of God.

We can see that Aquinas adds a new piece of information, which is a new premise (P7) – namely, that the image in this context should be understood with respect to excellence in comparison to others. Together with another premise (P8), according to which “man excels all animals by his reason and intelligence,” it gives a basis for concluding that the image concerns the reason and intelligence, which are incorporeal. In this light, we cannot derive C anymore. However, Aquinas does not negate the conclusion or premises. He also does not claim that the reasoning is invalid.

Thus, the cases of arg. 1 and arg. 2 are similar. Thomas presents new premises, which are consistent with the sets of basic premises. He does not criticize previous reasonings. However, due

to the new premises, the previous conclusions cannot be inferred anymore. Thus, it seems that such a reasoning is non-monotonic, and the logic which is open for such reasoning is non-monotonic too. Let us review the next three arguments. Arg. 3 is much shorter, and very similar to arg. 1:

- P1. [W]hatever has corporeal parts is a body.
- P2. Now Scripture attributes corporeal parts to God.
- P3. *Hast thou an arm like God?* (Job 40:4); and *The eyes of the Lord are upon the just* (Ps 33:16); and *The right hand of the Lord hath wrought strength* (Ps 117:16).
- C. Therefore God is a body.

P1 and P2 (if we rephrase it to the following form: “[According to Scripture:] God has corporeal parts”) create a kind of Barbara syllogism which gives C, and P3 constitutes evidence for P2. P1 and P3 are the basic premises. P1 is a metaphysical statement, and P3 is a collection of Biblical passages.

In his reply to this argument, Thomas, again, only adds a hermeneutic rule, and an example showing how it should be applied. The whole answer to the argument is as follows:

Reply to 3. Corporeal parts are attributed to God in Scripture on account of His actions, and this is **owing to a certain parallel** (*secundum similitudinem*). For instance the act of the eye is to see; hence the eye attributed to God signifies His power of seeing intellectually, not sensibly; and so on with the other parts.

As can we see, this rule, albeit similar to the previous ones, refers to parallelism between God’s actions which are intellectual and natural actions in which corporal parts are involved, like the eye in the act of seeing. And again, the rule is consistent with the set of basic premises. The reasoning is not being assessed as invalid. However, we see that according to this rule one cannot infer P2 from P3, and for this reason we cannot infer C.

Arg. 4 has a similar structure:

- P1. [P]osture (*situs*) belongs only to bodies.
- P2. But something which supposes posture is said of God in the Scriptures:
- P3. *I saw the Lord sitting* (Isa 6:1), and *He standeth up to judge* (Isa 3:13).
- C. Therefore God is a body.

P1 and P2 create a syllogism giving C. P2 is proven by P3. P1 and P3 are the basic premises. P1 is a metaphysical statement and P3 are Biblical passages. Aquinas replies: “Whatever pertains to posture, also, is only attributed to God **by some sort of parallel** (*secundum quandam similitudinem*). He is spoken of as sitting, on account of His unchangeableness and dominion; and as standing, on account of His power of overcoming whatever withstands Him.”

We can see that Aquinas, again, indicates a rule according to which a kind of parallelism (*similitudo*) must be taken into account when talking about God with respect to such a category as *situs*. This rule is, again, consistent with the set of basic premises (here: P1 and P3). Aquinas does not criticize the argument. He just adds this rule, which either makes it impossible to draw P2 as a conclusion, if it uses the category of posture in a strict (natural) sense, or forces us to read P2 in a way which means that we are no longer allowed to infer C.

Arg. 5 also quotes the passages which refer to location, but this time within relationships of coming to and departing from, in which there is a local term, so an object to which or from which something goes:

- P1. [O]nly bodies or things corporeal can be a local term wherefrom or whereto.
- P2a. But in the Scriptures God is spoken of as a local term whereto.
- P3a. [A]ccording to the words, *Come ye to Him and be enlightened* (Ps 33:6),

P2b. and as a term wherefrom:

P3b. *All they that depart from Thee shall be written in the earth* (Jer 17:13).

C. Therefore God is a body.

P1 and P2a or P2b (they can be taken together or independently) create a syllogism giving C. P3a proves P2a and P3b proves P2b. P1, P3a and P3b are the basic premises. P1 is a metaphysical statement, P3a and P3b are Biblical passages. Aquinas replies:

We draw near to God by no corporeal steps, since He is everywhere, but by the affections of our soul, and by the actions of that same soul do we withdraw from Him; thus, to draw near to or to withdraw from **signifies merely spiritual actions based on the metaphor** (*sub similitudine*) **of local motion**.

This time he refers to the hermeneutic rule according to which we should understand local motion related to God as spiritual actions. It means that either we cannot infer P2a from P3a and P2b from P3b, if we understand local terms strictly, and thus we also cannot infer C, or we should read P2a and P2b according to the hermeneutic rule (that is, as they are proven by P3a and P3b), but then they cannot be combined with P1, which refers to the strict sense of local terms, so we cannot infer C. As we can see, the rule is consistent with other premises and Thomas, again, does not challenge the reasoning.

Let us now summarize the results of this presentation. All five arguments (which can be perversely labelled as “the five ways of proving God’s corporeity”) have a similar structure. They prove the starting hypothesis by leading to a single conclusion (that God is a body), which is drawn from the set of premises, in which we find the basic premises and the premises of – so to speak – a higher order, which are proven by these basic ones. The sets of the basic premises always consist of: 1) metaphysical general statements, which seem rather obvious, as elements of Aristotelian metaphysics, and 2) sentences from the Bible. The “higher-order” premises are almost always conclusions based on sentences from the Bible, except P1 in arg. 2, which is proven by P2, being a metaphysical statement. We can also divide the premises into those from which C is drawn directly and the others, “indirect” ones. The sentences from the Bible are always the indirect ones. Next, in each answer Aquinas gives a hermeneutic rule. When replying to arg. 2 the rule says that image of God should be interpreted according to the excellence of man in comparison to other animals. In the remaining answers, the rules refer to the concept of parallelism (*similitudo*) which can be observed between physical attributes, actions involving corporeal parts, staying in some place or being a term “whereto” or “wherefrom” (which we find in the Bible as related to God), and respectively: God’s spiritual attributes, God’s spiritual actions, God’s spiritual situation or relationship to others, being a goal or source of spiritual actions (which express the real sense of Biblical passages).

Now, we can either say that in each case, C is an element of the set of the consequences of (or can be drawn from) the direct premises, or that it is an element of the set of the consequences of the basic premises, or of all the premises. No matter which option we choose, let us label the set of the premises of each argument as X. In each case, C is inferred from X, and then Aquinas does not negate anything from X and does not claim that any reasoning is invalid, but adds a hermeneutic rule, R, which is consistent with X. However, then C cannot be inferred anymore. If we thus assume that Aquinas to some extent accepted the arguments as possible, but, as someone who knows more, added an important piece of information which changed the situation, we may claim that he violated the condition of monotonicity, and so his reasoning is governed by a non-monotonic logic. We can present the two steps of each thread (consisting of an argument and a reply) as follows:

(1) $X \vdash C$,

(2) $X \cup R \not\vdash C$.

One can claim that Aquinas himself would never have proposed any of the five arguments, so he in fact treats them as invalid ones. Thus, he does not uphold the first step in any way. However, if we

stick to the text, we do not find any critique of those arguments. In my opinion, we should follow the discourse, in order to keep its very real nature.

And this nature is special. *ST* was conceived as a textbook for young theologians. In the prologue, Thomas underlines that “we purpose in this book to treat of whatever belongs to the Christian religion, in such a way as may tend to the instruction of beginners.” Next, he points out what kinds of problems students of theology faced when reading books of other authors, to then promise that his work will present the sacred doctrine “as briefly and clearly as the matter itself may allow” and without frequent repetitions. Hence, *ST* contains Christian theology, but at the same time, it is an example of theological didactics. This context seems essential to me. It explains, among others, the role and the status of arguments for the starting hypotheses. Aquinas decided to keep the standard structure of medieval questions, and thus he included such arguments. It was important to teach students that each point can and should be discussed. However, in *ST*, he limited himself to presenting the most important ones, and that is why each question only features a few arguments (very often three), instead of thirty, like in many articles of Aquinas’s discussed questions (*quaestiones disputatae*). Next, I claim that he did not want to teach the students that if an argument leads to a conclusion which is later opposed in the counterargument and the main answer, it should be criticized as invalid or based on false premises. Otherwise, he would have criticized such a reasoning. In my opinion, he tried to show that the concerns expressed in such arguments are important. They could be formulated by a student who does not know as much as his master. Let us also recall that such discussions, in which students formulated arguments for and against a thesis, to finally listen to the master’s answer, were a common practice of 13th-century lessons. So in the didactical context, such arguments should be respected, especially since the five discussed in this article seem quite rational and well-constructed. At least in this particular case, according to this specific context, I claim that Aquinas agreed with them in the sense that it would be acceptable for an unexperienced beginner to conduct such a reasoning. However, as a master he must, first, ask for or give a counterargument, next, give a rational answer, and finally, return to the argument to provide the right hermeneutic tool which the beginner is being trained to use as an exegete. Such a didactical, “open” approach is to some extent bound to be non-monotonic.

2.3. Boethius Dacus’s Non-Monotonic *De aeternitate mundi*

A completely different example of non-monotonic reasoning can be found in the main work of Boethius of Dacia, also called Boethius Dacus, namely, *On the Eternity of the World (De aeternitate mundi)*. This late 13th-century philosopher discussed one of the hottest problems of his century: did the world have a beginning or not. He tried to show that there is no contradiction between the theses of philosophy and theology with respect to this issue. Christian theology argued that the first sentence of the Book of Genesis clearly proves that the world did have a beginning. However, the most influential ancient philosopher in the 13th century, Aristotle, suggested that the world had never begun. According to Boethius, from the philosophical perspective, it is not necessarily true that the world had a beginning. For this reason, he was accused of being a coryphaeus of the so-called double truth theory, according to which there were two different truths: the theological and the philosophical one. However, in the proemium of *De aeternitate mundi*, he claims that whoever believes that the world is eternal is a heretic. How is it possible? In a paper published in 2019, I argued that in order to understand Boethius’s position and to maintain together the inferences formulated in natural science, philosophy and theology, we can refer to non-monotonic logic, as a very helpful framework [17].

Boethius presented different situations of conducting reasoning concerning the eternity of the world, represented by a natural scientist, a Christian (or a theologian), a mathematician, and a philosopher. For the natural scientist the main and only principle is nature. And thus, his premises will be taken from the observation of nature. The natural scientist observes that every event (or movement) has its precedent causes, so the first movement could not be “new” (or produced as new by an immediate cause), unless it was not a first movement. On these grounds the natural scientist

concludes, as Aristotle did in his *Physics*, that there could not be any beginning. At the same time the Christian, on the basis of the Christian revelation, will conclude the opposite. Next is the mathematician, who cannot prove by natural reasons that the world and the first movement were “new,” as this problem is indifferent to every mathematical discipline. The metaphysician cannot do this either. However, his background is different, as his discipline is capable of making statements concerning God. He will say he cannot solve the problem because what he knows is that the world depends on divine will as its sufficient cause, and he (as a metaphysician) does not know God’s will, so he is unable to judge whether the world – as a result of divine will – was simultaneous with this cause or there was a time when this cause existed without this result, which came later. So for the metaphysician both options are possible [17, pp. 459-460].

To sum up, neither the natural scientist, nor the mathematician, nor the metaphysician can prove that the world had a beginning; however, their knowledge is different, which determines what they can and what they cannot infer. In turn the Christian (or the Christian theologian) has a richer set of premises, as he adds the statements taken from revelation. And on this basis he concludes that the world had a beginning. It is disputable, but we can argue that according to Boethius’s idea, the metaphysician knows and accepts the premises of the natural scientist, and the Christian (especially the Christian theologian) knows and accepts the premises of both.

It is very important that according to Boethius, from the absolute perspective, the natural scientist is wrong when talking about the eternity of the world. However, he correctly draws a conclusion from the principles of his science, and so his reasoning is correct and acceptable. Thus, Boethius seems to accept all inferences limited to the premises formulated within particular disciplines, as long as they are logically correct. Hence, according to him, it is acceptable to draw a conclusion about the eternity of the world from the premises available to the natural scientist, but with the metaphysician’s richer set of premises it is no longer possible. I claim that this is an example of the use of a non-monotonic logic. Next, the Christian theologian adds other new premises consistent with the premises of the natural scientist, and infers the negation of the natural scientist’s conclusion, which constitutes another strong sign of the use of a non-monotonic logic.

In order to show it more clearly, let us present the situations of the natural scientist, the metaphysician and the Christian theologian using symbols (for: N – the set of premises of a natural scientist, M – the set of additional premises of a metaphysician, R – the set of additional premises of a Christian, α – the statement “The world had no beginning”). We are excluding here the mathematician, as he has no premises on which he could base a conclusion concerning the eternity of the world. We can present it in the following way:

- (1) the natural scientist: $N \vdash \alpha$
- (2) the metaphysician: $N \cup M \not\vdash \alpha$
- (3) the Christian theologian: $N \cup M \cup R \vdash \neg\alpha$.

According to this approach, the situation of the theologian is very special. Although he believes he has an absolute perspective (based on divine revelation), he should accept the reasoning of the natural scientist (if it is logically correct), which is relative, as limited to the premises taken from natural science. Hence, the theologian’s thinking must be based on non-monotonic logic. The situation of the metaphysician is similar; however, he cannot claim to have the absolute perspective, as he should be aware that his knowledge is not complete.

3. Non-Monotonic Reasoning in Biblical Exegesis

In this section, I would like to present an example of non-monotonic reasoning found in the work of another prominent medieval theologian, namely, St Bonaventure of Bagnoregio’s *Commentary on the Gospel of John*. This example will show that the phenomenon of non-monotonicity can also appear in exegetical discourse.

The *Commentary* is composed in a special way. Bonaventure comments on the whole text passage by passage, but for each fragment he, first, gives a verse-by-verse explanation, and second, poses a few questions regarding the whole passage or its particular verses and then briefly discusses

them. It is likely that this second step reflects the later work of the author. It definitely makes it more “scholastic.” The example analyzed below is one such discussion.

Bonaventure presents the following reasoning related to the words of Christ “Receive the Holy Spirit” (John 20:22): “It seems that he should not yet be giving them the Holy Spirit, since it is said in John 16:7: “If I do not go, the Paraclete will not come to you.” Therefore, if he had not yet ascended, he should not yet be giving the Holy Spirit [3, pp. 973-974].”

We can see that the reasoning is intended to indicate an inconsistency of the Gospel or even of Christ’s words, as his utterance reported in John 20:22 seems to negate the one reported in John 16:7. Let us reformulate the reasoning, for the sake of clarity:

- P1. It is said in John 16:7: “If I [Jesus] do not go, the Paraclete will not come to you.”
- P2. He [Jesus] had not yet ascended.
- C. He [Jesus] should not yet be giving the Holy Spirit.

The reasoning seems correct. In fact, it is an example of *modus ponens*; however, the direct conclusion (“The Paraclete will not come to you”) is skipped, and the next conclusion is drawn, based on the assumption that Jesus should not do anything which contradicts his previous utterance (let’s include it as a hidden premise – HP). Jesus stated clearly that the Holy Spirit (called the Paraclete) will come after Jesus goes. If this condition is not met (as he has not ascended yet), according to Jesus’s words, the Holy Spirit should not come, and so Jesus should not yet be giving the Holy Spirit.

Bonaventure replies:

I respond that it has to be said that the Holy Spirit is said to be received or to be given, **not by reason of essence, but by reason of effect.**

So the disciples had the Holy Spirit before the passion, but for the work of their salvation which is by grace. They had the Holy Spirit after the passion, but before the ascension for the forgiveness of sins. They had the Holy Spirit after the ascension for the preaching of our faith. So they were confirmed when the Holy Spirit descended in tongues of fire [3, p. 974].

We can see that Bonaventure does not criticize the earlier reasoning. Neither does he undermine the premises. He just gives a new piece of information, according to which “the Holy Spirit is said to be received or to be given, not by reason of essence, but by reason of effect” (let us mark this new premise as “P3”). P3 makes it possible to claim that we can identify many different kinds of giving and receiving the Holy Spirit by reason of effect, which Bonaventure precisely enumerates (let us mark it as “P4”). One of such acts was performed by Christ when he resurrected and came to the disciples saying “Receive the Holy Spirit.” In this perspective, Christ’s conditional promise does not exclude the possibility of giving the Holy Spirit before the condition is met, as there are two or even three such actions. Thus, with this new knowledge, we cannot infer the previous conclusion anymore. Let us remark that although it seems that Bonaventure infers P4 from P3, it is a shortcut, as he also needs other premises to draw such a conclusion.

Thus, we can claim that Bonaventure accepts the reasoning, which reflects quite reasonable doubts, but only adds this new perspective to show that those doubts can be dispelled. Hence, his reasoning is non-monotonic. We can present these two steps as follows (for X – the set of the starting premises including P1, P2 and HP, K – the set of new premises, P3 and P4, C – the conclusion of the starting reasoning):

- (1) $X \sim C$
- (2) $X \cup K \not\sim C$

We can see that to some extent the examples from Aquinas’s *ST* were similar to this reasoning, as they were also related to the interpretation of Biblical passages. However, they come from texts of

different natures. *ST* belongs to a part of theological teaching called *disputatio*, as its aim is to discuss theological problems, and, moreover, it was a theological textbook, whereas Bonaventure's text is a commentary on a book from the Bible, and so it represents Biblical exegesis, as well representing a part of theological teaching called *lectio*, which preceded *disputatio*, both in didactic practice and in the order of theological work. We can also indicate a more important difference, though closely related to the first one. In Aquinas's *ST* the "disabled" conclusions were inconsistent not only with another Biblical passage presented in the counterargument (John 4:24: "God is a spirit"), but, before all, with a strong doctrinal statement developed in Christian theology, establishing that God is not corporeal. In Bonaventure's commentary, though, the inconsistency was identified only between Biblical passages. Of course, this seems natural if we take into account the purposes of the exegetical text. I have indicated these differences to underline that non-monotonic reasoning is also a phenomenon present in the part of theological work called *lectio*.

Finally, there is another difference between the examples from Aquinas's and Bonaventure's texts, which will be important in the discussion below. Let us note down that in the second step presented by Bonaventure, there is no hermeneutic rule which indicates how we should interpret the words of the Bible. Instead, Bonaventure adds information which makes the issue to which the starting premise refers more nuanced.

4. Problems and Assumptions

In his polemic published in 2012, Patryk Pogoda drew my attention to the most important problem connected with my interpretation of Aquinas's reasonings which I claimed to be non-monotonic [12]. He offered many valuable remarks. However, the essence of his critique can be found in the following sentence: "These reasonings are not examples of non-monotonic reasonings, because the increase of the set of premises is not limited to adding new premises, but is a modification of already-existing premises" [12, p. 131].

According to his view, this modification runs as follows: 1) X were read as literal, 2) X are read as metaphorical. So in fact, in the "second step" (when Aquinas gives a hermeneutic rule) the sentences X are replaced with new sentences X , with the same words, but different meaning. Thus, Pogoda's postulate was: first – translation, next – inference. Let us follow this interpretation. It means that Aquinas is not giving a hermeneutic rule as an additional premise, but, by showing such a rule, he is saying: the sentence taken from the Bible is a metaphor, so let us translate it to reveal the exact sentence expressing the spiritual reality of God. Then indeed we have no grounds to claim that Aquinas's reasoning is non-monotonic.

We can also point out another problem with my interpretation. One can claim that in the second step a master or a commentator (like in the examples from Aquinas and Bonaventure) in fact challenges the previous inference and indirectly claims it is invalid. Hence, if we again mark the set of the starting premises as " X " and the new ones as " K ", then K are not added to X to create a richer set together, but they express the critique of the inference, they indicate an error.

These two problems are serious. However, if we concentrate on the text, we will find no trace of either the interpretation according to which sentences from the Bible should be translated or the interpretation which presents the second step as an attack on the first step of the reasoning. Let us also recall that the first problem does not apply to the example from Bonaventure's commentary on John 20:22, because in the second step, there is no hermeneutic rule which could be read as a postulate to translate the Biblical metaphorical sentence into a non-metaphorical one. Bonaventure only adds information which nuances the issue. Finally, in my opinion, the postulate of translation stems from a general opposition to non-monotonic reasoning, based on the assumption that the first step of reasoning should not be conducted, as it is based on premises which are not sufficient to deductively infer the conclusion. Thus, such an approach assumes that only classical logic is allowed. Hence, it *a priori* excludes non-monotonic reasoning. It is then a matter of choice – whether we accept the non-classical approach or not.

However, if we rule out non-monotonic or even non-classical reasoning, we might overlook important features of the theological discourse. For instance, we can explain Anselm's *Proslogion* in terms of classical logic and argue that in the second part he was talking about slightly different concepts than in the first part, and so it is possible to defend the monotonicity of the discourse. But then we will lose a deep and true nerve of this text, which reflects a feeling of inconsistency which appears when an honest theologian talks about God and must somehow reconcile the *via positiva* and *via negativa*. We can also try to read Boethius of Dacia in a monotonic manner. But then we must say that according to Boethius only theological reasoning is allowed, and the reasoning of the natural scientist is invalid, as he did not take into account all possible knowledge. However, this would be contrary to what Boethius actually said.

I see at least a few special features of theological method and practice which are a basis to argue that especially in this field the non-monotonic approach is right. To some extent they relate to what has already been said. They are also intertwined. However, let us try to identify them:

1. Variety of senses. When working directly on Biblical passages, both within *lectio* (the example from Bonaventure) and *disputatio* (the example from Aquinas), the theologian must be open to this variety. Let us note that when Aquinas shows, in *ST I*, q. 3, a. 1, arg. 1, that the three dimensions attributed to God imply God's virtual quantity, he gives two possible ways of interpretation. Thus, he is not sure which of them is right, and perhaps both of them are, or there are also other possible ways to read this passage. The theologian must always be aware that his interpretation may not be the ultimate one and perhaps there is something deeper which can be coded by the passage he is trying to unpack. What is more, medieval theology developed the tradition of the four senses of Scripture (the historical or literal one, and the three spiritual: allegorical, moral/tropological, and anagogical). And in some cases, all four senses can be found in the same Biblical passage. Hence, the right way for the theologian is not to translate Biblical sentences into some unambiguous sentences, but to keep them as they are, and develop their theological reading by grasping new hermeneutic rules [cf. 15]. The examples from Bonaventure and Aquinas can be a good illustration of such a development, at least at the early stage of theological training.

2. Respect for Biblical sentences. This feature is very close to the previous one. It seems that theologians should always directly refer to the words from the Bible, which can always surprise them in later readings. For theologians, they are also considered to be the word of God. Thus, they should not be translated, in order to be ready as premises for inferences, but taken for granted. Hence, in my opinion the proper premises for theological reasoning based on the Bible should be the Biblical sentences, not their interpretation. We can call this approach "nominalist," to underline that the premises in such a reasoning are the exact words from the Bible and not the meaning assigned to them by an interpreter; in short: there are only words, as they are.

3. The humility of the theologian. As an interpreter, the theologian should never be sure if a particular reading of a Biblical passage is right or ultimate. However, this should not be an obstacle to conduct reasoning, although it may be the case that a new hermeneutic rule or perspective will change the situation. Thus, the theologian simply must use non-monotonic reasoning. Next, the example of Anselm's *Proslogion* shows that the humble theologians, after saying something positive about God's attributes, should follow the negative way of interpretation, thus to some extent distancing themselves from what has been said, to find the right balance in theological discourse. It also seems that this reflects the spiritual attitude of such a humble theologian. As Jacob Archambault has shown, in the case of Anselm, such an approach has a non-monotonic nature.

4. The processual or circular nature of theology. We can observe a kind of circulation in theological practice: first, reading the Bible and basic inferences; second, formulation of the doctrine; third, reading the Bible using the doctrine; and again, new inferences, and so on.⁴ It seems that the example from Bonaventure represents such a third step, in which the theory concerning receiving the Holy Spirit made it possible to read the Bible in a new light and to reconcile the two passages. This nature assumes a development and a process, in which we must conduct reasoning, but later we can acquire new information, which can make it impossible to infer the previous

conclusion, as in the example from Bonaventure. This special nature is then closely related to non-monotonic reasoning.

5. The compatibility with natural sciences. There is an approach in Christian theology (and in my view it is the most fruitful and today dominant one) according to which theologians respect the results of natural science and claim that their discipline should be compatible with them. The example of Boethius's *On the Eternity of the World* proves that it is feasible. And if theologians follow his path, they must use non-monotonic logic.

However, we can also learn from this presentation that in order to maintain the discussion on the non-monotonic nature of reasoning in theology on various levels, we need to accept some key assumptions. Without them, we will always be in danger of supporting arguments which can subvert the whole discussion. I propose the following assumptions:

1. Biblical exegesis has a special status different than in other disciplines, mainly due to the special status of Biblical sentences.
2. It is legitimate to infer from sentences with an undetermined sense, and so we reject the postulate to first translate the sentences in which the sense is not clear.
3. We should adopt the "nominalist" approach (identity of the sentence is determined by the words of the Biblical text).
4. Pro-classical interpretation of a reasoning will render it classical, and so we should refrain from arguments which *a priori* rule out non-classical approaches.

5. Conclusion

Different occurrences of non-monotonic reasoning in theological practice have been presented. They formed a basis for discussing the problems with interpreting the provided examples as non-monotonic, and for pointing out the specific features of theology which support such an interpretation. Next, the discussion served as a starting point for the formulation of a set of assumptions which should be accepted to continue the debate about the non-monotonic or monotonic nature of theological reasoning.

To conclude, in the light of the above-mentioned discussion and assumptions, I think I am allowed to claim that the non-monotonic approach really is inherent in theology. In my opinion, the examples from medieval theology found in the texts of Anselm of Canterbury, Bonaventure, Thomas Aquinas and Boethius of Dacia, are a good illustration of this claim.

Acknowledgements

This article was supported by the National Science Centre, Poland, under grant agreement No. 2020/39/D/HS1/03200.

References

1. Archambault, J. Monotonic and Non-Monotonic Embeddings of Anselm's Proof. *Logica Universalis* 11 (1), 2017, pp. 120-138. <https://doi.org/10.1007/s11787-017-0162-7>
2. Besnard, P. A Note on Directions for Cumulativity. *IfCoLog Journal of Logics and Their Applications* 1 (2), 2014, pp. 77-81.
3. Bonaventure of Bagnoregio and Karris R. J. (tr.). *Commentary on the Gospel of John*, Saint Bonaventure, NY: The Franciscan Institute and Saint Bonaventure University, 2007.
4. Franks C. The Realm of the Sacred, Wherein We May Not Draw an Inference From Something Which Itself Has Been Inferred: A Reading of Talmud Bavli Zevachim Folio 50. *History and Philosophy of Logic*, 33(1), 2012, pp. 69-86.
5. Gabbay, D. M. Theoretical Foundations for Non-Monotonic Reasoning in Expert Systems. In: K. R. Apt (ed.), *Logics and Models of Concurrent Systems* (pp. 439-457), Berlin: Springer, 1985.
6. Jamieson, R. B. and Wittman, T. R. *Biblical Reasoning: Christological and Trinitarian Rules for Exegesis*, Grand Rapids, MI: Baker Academic, 2022.

7. Makinson, D. General Theory of Cumulative Inferences. In: M. Reinfrank, J. Kleer, M. L. Ginsberg, E. Sandewall (eds.), *Non-Monotonic Reasoning: 2nd International Workshop, Grassau, FRG, June 1988. Proceedings* (pp. 1-18), Berlin: Springer, 1989.
8. Malinowski, J. Logiki niemonotoniczne. *Przegląd Filozoficzny – Nowa seria* 6 (1), 1997, pp. 31-53.
9. Malinowski, J. Rola oczekiwań w rozumowaniach. *Ruch Filozoficzny* 53 (4), 1996, pp. 757-583.
10. McDermott, D. Nonmonotonic Logic II: Nonmonotonic Modal Theories, *Journal of the ACM* 29 (1), 1982, pp. 33-57. <https://doi.org/10.1145/322290.322293>
11. McDermott, D. and Doyle, J. Non-Monotonic Logic I. *Artificial Intelligence* 13 (1-2), 1980, pp. 41-72. [https://doi.org/10.1016/0004-3702\(80\)90012-0](https://doi.org/10.1016/0004-3702(80)90012-0)
12. Pogoda, P. Tomasz z Akwinu monotonicznie. *Filozofia Nauki* 20 (2), 2012, pp. 129-133.
13. Reiter, R. A Logic for Default Reasoning. *Artificial Intelligence* 13 (1-2), 1980, pp. 81-132. [https://doi.org/10.1016/0004-3702\(80\)90014-4](https://doi.org/10.1016/0004-3702(80)90014-4)
14. Reiter, R. On Reasoning by Default. In: D. L. Waltz (ed.), *TINLAP '78: Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing* (pp. 210-218), Stroudsburg, PA: Association for Computational Linguistics, 1978. <https://doi.org/10.3115/980262.980297>
15. Roszak, P. Between Dialectics and Metaphor. Dynamics of Exegetical Practice of Thomas Aquinas. *Angelicum* 90 (3), 2013, pp. 507-534.
16. Schumann, A. Logical Cornerstones of Judaic Argumentation Theory. *Argumentation* 27, 2014, pp. 305-326. <https://doi.org/10.1007/s10503-012-9273-8>
17. Trepczyński, M. Non-Monotonic Logic and the Compatibility of Science and Religion. *Logica Universalis* 4 (13), 2019. <https://doi.org/10.1007/s11787-019-00230-4>
18. Trepczyński, M. Tomasz z Akwinu niemonotonicznie. *Filozofia Nauki* 19 (2), 2011, pp. 115-121.
19. Wójtowicz, A. and Trepczyński, M. Logiki niemonotoniczne jako sposób wnioskowania w niesprzyjających warunkach. *Filozofia Nauki* 19 (2), 2011, pp. 99-103.

Notes

1. I would like to thank the anonymous reviewer who has drawn my attention to this fascinating material and to the brilliant analyses. Tractate *Zevachim* is available in English at: <https://www.sefaria.org/Zevachim.49b.1?lang=bi>
2. Archambault rightly points out that argument and proof are not the same. He shows that, according to the ancient tradition, argument is the main idea and the basic point for a proof: “In the Boethian parlance, an ‘argument’ (argumentum) is not a set of premises which, when combined in the appropriate manner, lead to a conclusion; rather, it is an aspect of something whereby something further may be inferred about it or something else – typically, it is a concept signified by the middle term of a syllogism.” He also refers to Themistian typology to indicate that Anselm uses a topic *from description* and that in this case such an argument is the famous Anselmian phrase: “that than which a greater cannot be thought” [1, pp. 124-125].
3. I use the English translation of *ST* published in the New Advent repository: <https://www.newadvent.org/summa/1.htm>, and the original text available in Corpus Thomisticum: <https://www.corpusthomicum.org/sth1003.html>.
4. This is similar to what Robert B. Jamieson and Tyler R. Wittman identified in John Webster’s theory of Biblical reasoning: “Webster distinguishes within biblical reasoning two overlapping, mutually informing modes of reasoning: exegetical and dogmatic” [6, p. xvii].

Public Theology Facing a Planet in Turmoil



Abstract: The interview given by Ted Peters who pursues *Public Theology* at the intersection of science, religion, ethics, and public policy. Peters is an emeritus professor at the Graduate Theological Union, where he co-edits the journal, *Theology and Science*, on behalf of the Center for Theology and the Natural Sciences, in Berkeley, California, USA. His book, *God in Cosmic History*, traces the rise of the Axial religions 2500 years ago. He previously authored *Playing God? Genetic Determinism and Human Freedom?* (Routledge, 2nd ed., 2002) as well as *Science, Theology, and Ethics* (Ashgate 2003). He is editor of *AI and IA: Utopia or Extinction?* (ATF 2019). Along with Arvin Gouw and Brian Patrick Green, he co-edited the new book, *Religious Transhumanism and Its Critics* hot off the press (Roman and Littlefield/Lexington, 2022). Soon he will publish *The Voice of Christian Public Theology* (ATF 2022). See his websites: TedsTimelyTake.com and <https://www.patheos.com/blogs/publictheology/>.

Keywords: theology, war, Ukraine.

Konrad Szocik: Let's start with the situation in the US. How would you be to define the cultural and social situation in the U.S. from the point of view of theology and philosophy? One of the recently discussed topics is the tightening of abortion laws. How can this be reconciled with traditional American freedom, which is often a model for the world?

Ted Peters: We in America fear we might be inundated with an anti-intellectualism that would reject *a priori* any contributions offered by theologians to public discourse. Gone are the Reinhold Niebuhrs to whom the wider public listened for wisdom.

We're living in a postmodern nightmare (deconstructionist, not holistic, postmodern nightmare) where everything is adversarial, perspectival, and power-grabbing. There is no room for neutrality, fairness, or rational adjudication. In my judgment, social media has contributed significantly to the wind of ill will that is blowing across our society. See my recent blog post: "[The Chaos Machine, Social Media, and Public Theology.](#)"

The connection between the Republican Party and American evangelical Christians is most disturbing. It appears that Republican leaders have fooled the faithful in the churches. If the party supports the "pro-life" position on abortion, then the Christians should vote accordingly. No matter how heartlessly the Republicans treat penniless immigrants, victims of mass shootings, and the fragility of our environment, Christians are expected to fall in line. I find this as unfathomable as detestable.

Konrad Szocik: One of the most serious challenges facing the world is climate change. Radical critics accuse capitalism of being the root of all evil. But can we imagine any other economic, and social,

model than capitalism? And while remaining in capitalist structures, can we hope to stop environmental degradation at all?

Ted Peters: You're right in connecting economics and ecology. The problem is not merely global economic exchange. The problem includes an ideology that justifies it. My colleague at the University of California at Berkeley, economist Richard Norgaard, has labeled this ideology, the "church of economism." Functionally, economism is a religion.

The problem within economism which cries out for reform is found in both its form and substance. The amoral form of economism is found in the categories it imposes on our thinking, on our false consciousness. According to economism, everything becomes subjected to cost-benefit categories. Thereby, this marginalizes the tender values of intimacy, caring, sharing, and building. Similarly, the amoral substance of economism fosters just the opposite of what the Christian religion advocates: greed instead of charity, individual freedom without responsibility for the common good, anarchy without unity. If economism is in fact a religion, then it needs at least a reformation if not a supersession.

The [public theologian](#) needs to lift up a vision of a global common good in terms of a just, sustainable, participatory, and global society of moral deliberation. I lay out the details in a series of Patheos posts on [economism versus the common good](#).

Konrad Szocik: The war in Ukraine has shocked the world, especially the Western world, which seems to have "forgotten" the traditional wars waged on European territory. How can philosophy and public theology help us understand this war, and what role can they have to play?

Ted Peters: I'm glad you ask how the philosopher and theologian can "understand" the war. I'd rather put an end to it. But, "understand" may be the length to which I could go.

This is Putin's war. Pure and simple. His attempt at self-justification by appealing to "Mother Russia" and such is so transparently fake that everyone in the world recognizes the fraud. Everyone except Putin himself.

In Ukraine, the populace cowers in subway tunnels and under bridges to protect themselves from callous missile strikes on civilian targets. The Russian military seems heartless, ruthless, murderous.

So, what does the Patriarch of Moscow and All Russia do? Does Patriarch Kirill intervene with President Putin to plead for peace? To stop the war? To reconcile Russians with Ukrainians? No.

Rather, Kirill uses his bully pulpit to excoriate the people of Ukraine for holding gay pride parades. Kirill called the military action in Ukraine a conflict about matters "far more important than politics." And, he insinuated that the embrace of progressive western values would lead to the end of civilization.

Are you kidding? No, I'm not kidding. I'm not spreading false news. This is a dark day in global Christian leadership. See: "[Moscow: The Worst Public Theology](#)."

Konrad Szocik: Finally, it is worth asking whether the West can or should do more than it is doing so far in helping Ukraine. The discussion has been, as is usually the case, appropriated and dominated by politicians and the military. However, let's take the point of view of philosophy and public theology. Can we talk about the existence of special duties to those affected, and is there any universal morality to which we should adhere?

Ted Peters: As of this moment, the Ukrainian military is mounting a significant counter-offensive and reclaiming lost territory. The whole world is waving Ukrainian flags and cheering for the victims.

It appears to me that President Joe Biden and his NATO allies have been prudent. On the one hand, they wish Ukraine well. On the other hand, they wish to avoid raising the ire of Vladimir Putin any further. At all cost, NATO wants to avoid precipitating World War III. Because Ukraine is not yet a member of NATO, would it be illegal for NATO to send troops against Russia? Would it be legal to send weapons and advisors, but no troops? It seems the latter has been acceptable to Putin. NATO is walking a tight rope.

I'm grateful that we have thus far avoided World War III. Might it be possible to save Ukraine as well? As a philosopher and theologian and not a soldier, the best I can do is turn to prayer.

Intellectual and Ethical Virtues in the Situation of War



Abstract: The interview given by Vojko Strahovnik, Department Chair and Associate Professor at the Department of Philosophy and Research Fellow in Philosophy at the Faculty of Theology, University of Ljubljana. The impact of his work ranges from insights into the nature of normativity (the role of moral principles in the formation of moral judgments, the authority of the normative domain, epistemic agency, and epistemic virtuousness) to considerations related to practical dimensions of our lives (e.g., the role of guilt and moral shame in reconciliation processes, the importance of intellectual and ethical virtues in dialogue and education, global justice, animal ethics). His recent outreach activities include being a visiting lecturer (2017) and a Templeton and Fulbright research scholar (2016; 2022) at the University of Arizona, Department of Philosophy. The central question that incites him most is the structure and phenomenology of normativity. Webpage: <http://vojkostrahovnik.idh.si/papers/>.

Keywords: ethics, war, Ukraine, particularism, generalism.

Andrew Schumann: What is the role of moral principles in the formation of moral judgments? Is it possible to formalize moral judgments by means of some logic?

Vojko Strahovnik: I think moral judgment is more related to moral vision or moral sensitivity than ethical principles. This is not to say that the latter do not have some sort of regulative role or the role of enabling us to structure our moral outlook. In this sense, I lean towards moral particularism. At the heart of the debate between moral particularism and moral generalism lies the question of the acceptability of a principled approach to morality, including the question of whether universal, exceptionless moral principles govern morality. As Sean McKeever and Michael Ridge usefully put it recently, despite a diverse range of views under the label “particularism”, what they have in common is a negative attitude towards moral principles. Still, most forms of particularism nonetheless allow for a particular role that moral generalities nevertheless play in the structure of the proposed moral theory. There is a variety of views that occupy the region between strict moral generalism and radical particularism. All these accounts strive to incorporate some important aspects of morality that the debate on particularism pointed out, most noteworthy holistic nature of reasons and the seemingly ineliminable problem of exceptions to universal moral rules. In several of my papers, I have defended a version of moral pluralism that combines elements of a principled approach to morality and particularism.

Andrew Schumann: What is animal ethics? How much do people need animal ethics? What for?

Vojko Strahovnik: One way to delineate animal ethics is to define it as a domain of practical ethics that deals predominantly with the moral status of nonhuman animals and the ethics of our practices that include them. It harbours numerous topics as well as various approaches. The central question itself is, of course, very old. Perhaps most striking is what some (Cora Diamond, Stanley Cavell etc.), including me, have described as the difficulty of the animal question. Animal question is at the

heart of animal ethics and pertains primarily to the question of the moral status of nonhuman animals. For me, the moral vision that brings us close to an answer to this question is that what establishes the relationship between us, humans, and nonhuman animals is a sense of vulnerability and mortality, which we share with them as beings with living, vulnerable bodies. Yet still, the difficulty of animal questions related to the aspect that this question seems to defy attempts to articulate and pose it in its full significance. In answering the animal question, we should not reduce our answers to just a single morally important or decisive relationship. There is a plurality of morally relevant relationships, and each has its meaning inside a particular form of life. And the most tenacious aspect of the difficulty of the animal question is the perplexity and even perplexity and even anxiety that can arise due to the gap between philosophy (or our rational, detached judgment about the moral status of animals, the relevance of their pain and suffering, etc.) and our actual practices. In this sense, animal ethics is important, even for understanding ourselves, but it often lacks an apt bite in relation to changing our behaviour.

Andrew Schumann: What is your position on the war in Ukraine and how interesting is this topic for the Slovenian philosophers?

Vojko Strahovnik: I have no doubts that Russian aggression in Ukraine calls for the strongest moral condemnation, and the sanctions that were imposed are fully justified given the mentioned moral status of Putin's actions that prompted them. The whole situation, and in particular the people's suffering, saddens me the utmost. Many intellectuals in Slovenia have raised similar concerns. Some of them emphasized the importance of bringing the whole situation back behind the diplomatic table and prioritising the quest for peace. In this sense, they expressed important reservations regarding sending arms support to Ukraine and fuelling the war machine. But what is perhaps missing from this picture is a clear wrongness of Russia's acts in the first place. And given that a Ten-Day War or Slovenian War of Independence in 1991 bears some likenesses to what is going on in Ukraine, the whole context for evaluating the current war in Ukraine, together with the reminiscence of our past gets further complicated.

Andrew Schumann: Are intellectual and ethical virtues possible in dialogue in the context of information warfare, for example, in the situation of information warfare around Ukraine?

Vojko Strahovnik: In particular, intellectual or epistemic virtues are important. It is clear that both sides in the conflict extensively use propaganda and similar tools to put forward their preferred version of the story. On the other hand, in the context of the covid-pandemic many (fairly reliable and reputable) news outlets are being snubbed, and a plethora of conspiracy theories and delusive ideas are replacing them. And in such circumstances, epistemic virtues are even more important than ever. So, the question is, how can we establish (public) spaces and spheres that would enable us to develop and cultivate them?