

A Model for the Burden of Persuasion in Argumentation

Roberta CALEGARI ^{a,1} and Giovanni SARTOR ^{a,b}

^a CIRSIFID - Alma AI, University of Bologna, Italy

^b European University Institute, Florence, Italy

Abstract. This work provides a formal model for the burden of persuasion in legal proceedings. The model shows how the allocation of the burden of persuasion may induce a satisfactory outcome in contexts in which the assessment of conflicting arguments would, without such an allocation, remain undecided. The proposed model is based on an argumentation setting in which arguments may be accepted or rejected according to whether the burden of persuasion falls on the conclusion of such arguments or on its complements. Our model merges two ideas that have emerged in the debate on the burden of persuasion: the idea that allocation of the burden of persuasion makes it possible to resolve conflicts between arguments, and the idea that its satisfaction depends on the dialectical statuses of the arguments involved. Our model also addresses cases in which the burden of persuasion is inverted, and cases in which burdens of persuasion are inferred through arguments.

Keywords. burden of persuasion, argumentation, legal reasoning

1. Introduction

The burden of proof is a central feature in legal decision-making and yet no agreed theory of it exists [1,2]. Generally speaking, we can say that burdens of proof distribute dialectical responsibilities to the parties: when a party has a burden of proof of type b relative to a claim ϕ , then, unless the party provides the kinds of arguments or evidence required by b , the party will lose on claim ϕ , i.e., that party will fail to establish ϕ . Burdens of proof can complement the analysis of dialectical frameworks that are provided by argumentation systems. In particular, they are important in adversarial contexts: they are meant to facilitate the process of reaching a single outcome in contexts of doubt and lack of information. In the legal domain, two types of burdens are distinguished: the *burden of production* (also called burden of providing evidence, or ‘evidential’ burden), and the *burden of persuasion*. The focus of this paper is on the burden of persuasion, and its purpose is to show how an allocation of the burden of persuasion may induce single outcomes in contexts in which the assessment of conflicting arguments would, without such an allocation, remain undecided. Our approach is based on providing specific criteria for accepting and rejecting propositions upon which there is a burden of persuasion.

¹R. Calegari and G. Sartor have been supported by the H2020 ERC Project “CompuLaw” (G.A. 833647).

2. Burdens of production and burdens of persuasion

Following the account in [3], we distinguish the burden of production from the burden of persuasion. A party burdened with production needs to provide some support for the claim he or she is advancing. More exactly, we can say that the party has the burden of production for ϕ if the following is the case: unless relevant support for ϕ is provided – i.e., unless an argument for ϕ is presented that deserves to be taken into consideration – then ϕ will not be established (even in the absence of arguments against ϕ). When knowledge is represented through a set of rules and exceptions, the party interested in establishing the conclusion of a rule has the burden of production relative to the elements in the rule’s antecedent condition, while the other party (who is interested in preventing the conclusion from being derived from the rule) has the burden of production relative to the exceptions to the rule (as provided in a separate exception clause or in an unless-exception within the rule). Note that meeting the burden of production for a claim ϕ is only a *necessary* condition, and not a sufficient one, for establishing ϕ , since the produced arguments may be defeated by counterarguments. This aspect is addressed by the burden of persuasion, under which the burdened party looking to establish a claim needs to provide a ‘convincing’ argument for it—that is, an argument that prevails over arguments to the contrary to an extent that is determined by the applicable standard of proof. If there is a burden of persuasion on a proposition ϕ , and no prevailing argument for ϕ is provided, then the party concerned will lose on ϕ . In this paper, we focus on the burden of persuasion. We shall discuss it by way of three running examples: one from criminal law, one from civil law, and one from antidiscrimination law.

In criminal law, the burden of production is distributed between prosecution and defence, while the burden of persuasion (in most legal systems) is always on prosecution. More exactly, in criminal law, the burden of production falls on the prosecution relative to the two constitutive elements of crime, namely, the criminal act (*actus reus*) and the required mental state (*mens rea*, be it intention/recklessness or negligence), while it falls to the defendant relative to justifications or exculpatory defences (e.g., self-defence, state of necessity, etc.). In other words, if both *actus reus* and *mens rea* are established, but no exculpatory evidence is provided, the decision should be a criminal conviction. On the other hand, the burden of persuasion falls on the prosecution for all determinants of criminal responsibility, including not only for the constitutive elements of a crime but also for the absence of an exculpatory defence.

Example 1 (Criminal law example) *Let us consider a case in which a woman has shot and killed an intruder in her own home. The applicable law consists of the rule according to which intentional killing constitutes murder, and in the exception according to which there is no murder if the victim was killed in self-defence. Assume that it has been established with certainty that the woman shot the intruder and that she did so intentionally. However, it remains uncertain whether the intruder was threatening the woman with a gun, as claimed by the defence, or had turned back and was running away on having been discovered, as claimed by the prosecution. The burden of persuasion is on prosecution, who needs to provide a convincing argument for murder. Since it remains uncertain whether there was self-defence, prosecution has failed to provide such an argument. Therefore the legally correct solution is that there should be no conviction: the woman needs to be acquitted.*

In civil law, both the burden of production and the burden of persuasion may be allocated in different ways in the law, depending on various factors, such as the ability of a party to provide evidence in favour of his or her claim. In matters of civil liability, for example, it is usually the case that the plaintiff, who asks for compensation, has to prove both that the defendant caused him harm, and that this was done intentionally or negligently. However, in certain cases, there is an inversion of the burden of proof for negligence (both the burden of production and the burden of persuasion). This means that in order to obtain compensation, the plaintiff only has to prove that he was harmed by the defendant. This will be sufficient to win the case unless the defendant provides a convincing argument that she was not negligent.

Example 2 (Civil law example) *Let us consider a case in which a doctor caused harm to a patient by misdiagnosing his case. There is no doubt that the doctor harmed the patient: she failed to diagnose a cancer, which consequently spread and became incurable. However, it is uncertain whether or not the doctor followed the guidelines governing this case: it is unclear whether she prescribed all the tests that were required by the guidelines in such a case, or whether she failed to prescribe some tests that would have enabled the cancer to be detected. Assume that, under the applicable law, doctors are liable for any harm suffered by their patients, but they can avoid liability if they show that they were diligent (not negligent) in treating the patient, i.e., that they exercised due care. Thus, doctors have both a burden of production and a burden of persuasion concerning their diligence. Let us assume that law also says that doctors are considered to be diligent if they followed the medical guidelines that govern the case. In this case, given that the doctor has the burden of persuasion on her diligence, and that she failed to provide a convincing argument for it, the legally correct solution is that she should be ordered to compensate the patient.*

These two examples share a common feature. In both, uncertainty remains concerning a decisive issue, namely, the existence of self-defence in the first example and the doctor's diligence in the second. However, this uncertainty does not preclude the law from prescribing a single legal outcome in each case. This outcome can be achieved by discarding the arguments that fail to meet the required burden of persuasion, i.e., the prosecution's argument for murder and the doctor's argument for her diligence, respectively. Our third example addresses anti-discrimination law. According to the European law against discrimination – or at least according to an interpretation of some of its controversial provisions – where there is evidence for discrimination in employment, it is on the employer to prove that there was no discrimination.

Example 3 (Anti-discrimination law example) *Let us consider a case in which a woman claims to have been discriminated against in her career on the basis of her sex, as she was passed over by male colleagues when promotions came available, and brings evidence showing that in her company all managerial positions are held by men, even though the company's personnel includes many equally qualified women, having worked for a long time in the company, and with equal or better performance. Assume that this practice is deemed to indicate the existence of gender-based discrimination, and that the employer fails to provide prevailing evidence that the woman was not discriminated against. It seems that it may be concluded that the woman was indeed discriminated against on the basis of her sex.*

In this paper, we put forward a formal model for the burden of persuasion which captures the patterns of reasoning that are exemplified above. Our model originates from legal considerations and is applied to legal examples. However, the issue of the burden of proof carries a significance that goes beyond the legal domain and involves other domains – public discourse, risk management, etc. – in which evidence and arguments are needed and corresponding responsibilities are allocated according to dialectical or organisational roles.

3. Argumentation Framework

We introduce a structured argumentation framework relying on a lightweight ASPIC⁺-like argumentation system [4]. In a nutshell, arguments are produced from a set of defeasible rules, and attack relationships between arguments are drawn into argumentation graphs. Then arguments from the graph are labelled by following an acceptance labelling semantics that takes burdens of persuasion into account.

3.1. Defeasible theories, argumentation graphs and burden of persuasion

Let a literal be an atomic proposition or the negation of one.

Notation 3.1 For any literal ϕ , its complement is denoted by $\bar{\phi}$. That is, if ϕ is a proposition p , then $\bar{\phi} = \neg p$, while if ϕ is $\neg p$, then $\bar{\phi}$ is p .

Literals are brought into relation through defeasible rules.

Definition 3.1 A *defeasible rule* r has the form: $\rho : \phi_1, \dots, \phi_n, \sim \phi'_1, \dots, \sim \phi'_m \Rightarrow \psi$ with $0 \leq n$, and where

- ρ is the unique identifier for r , denoted by $N(r)$;
- each $\phi_1, \dots, \phi_n, \phi'_1, \dots, \phi'_m, \psi$ is a literal;
- $\phi_1, \dots, \phi_n, \sim \phi'_1, \dots, \sim \phi'_m$ are denoted by $\text{Antecedent}(r)$ and ψ by $\text{Consequent}(r)$;
- $\sim \phi$ denotes the weak negation (negation by failure) of ϕ : ϕ is an exception that would block the application of the rule whose antecedent includes $\sim \phi$.

The name of a rule can be used as a literal to specify that the named rule is applicable, and its negation correspondingly to specify that the rule is inapplicable [5].

A superiority relation \succ is defined over rules: $s \succ r$ states that rule s prevails over rule r .

Definition 3.2 A *superiority relation* \succ over a set of rules Rules is an antireflexive and antisymmetric binary relation over Rules , i.e., $\succ \subseteq \text{Rules} \times \text{Rules}$.

A defeasible theory consists of a set of rules and a superiority relation over the rules.

Definition 3.3 A *defeasible theory* is a tuple $\langle \text{Rules}, \succ \rangle$ where Rules is a set of rules, and \succ is a superiority relation over Rules .

Given a defeasible theory, by chaining rules from the theory we can construct arguments, as specified in the following definition; cf. [5,6,7].

Definition 3.4 An *argument* A constructed from a defeasible theory $\langle \text{Rules}, \succ \rangle$ is a finite construct of the form: $A : A_1, \dots, A_n \Rightarrow_r \phi$ with $0 \leq n$, where

- A is the argument's unique identifier;
- A_1, \dots, A_n are arguments constructed from the defeasible theory $\langle \text{Rules}, \succ \rangle$;
- ϕ is the conclusion of the argument, denoted by $\text{Conc}(A)$;
- $r : \text{Conc}(A_1), \dots, \text{Conc}(A_n) \Rightarrow \phi$ is the top rule of A , denoted by $\text{TopRule}(A)$.

Notation 3.2 Given an argument $A : A_1, \dots, A_n \Rightarrow_r \phi$ as in definition 3.4, $\text{Sub}(A)$ denotes the *set of subarguments* of A , i.e., $\text{Sub}(A) = \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_n) \cup \{A\}$. $\text{DirectSub}(A)$ denotes the *direct subarguments* of A , i.e., $\text{DirectSub}(A) = \{A_1, \dots, A_n\}$.

Preferences over arguments are defined via a last-link ordering: an argument A is preferred over another argument B if the top rule of A is stronger than the top rule of B .

Definition 3.5 A *preference relation* \succ is a binary relation over a set of arguments \mathcal{A} : an argument A is preferred to argument B , denoted by $A \succ B$, iff $\text{TopRule}(A) \succ \text{TopRule}(B)$.

We now provide definitions of possible collisions between arguments. Our definition focuses on cases in which an argument: (a) contradicts the conclusion of another argument (top-rebutting), or (b) denies the (applications of the) latter's top rule or contradicts a weak negation in the latter's body (top-undercutting).

Definition 3.6 A *top-rebuts* B iff $\text{Conc}(A) = \overline{\text{Conc}(B)}$, and $B \not\succeq A$; A *strictly top-rebuts* B iff $A \succ B$.

Definition 3.7 A *top-undercuts* B iff

- $\text{Conc}(A) = \neg N(r)$ and $\text{TopRule}(B) = r$; or
- $\text{Conc}(A) = \phi$ and $\sim \phi \in \text{Antecedent}(\text{TopRule}(B))$

Definition 3.8

- A *top-attacks* B iff A *top-rebuts* B or A *top-undercuts* B
- A *strictly top-attacks* B iff A *strictly-top-rebuts* B or A *top-undercuts* B

3.2. Labelling semantics

We use $\{\text{IN}, \text{OUT}, \text{UND}\}$ -labellings, where each argument is labelled IN, OUT, or UND, depending on whether it is accepted, rejected, or undecided, respectively.

Definition 3.9 Let G be an argumentation graph. An $\{\text{IN}, \text{OUT}, \text{UND}\}$ -*labelling* L of G is a total *function* $\mathcal{A}_G \rightarrow \{\text{IN}, \text{OUT}, \text{UND}\}$.

Notation 3.3 Given a labelling L , we write $\text{IN}(L)$ for $\{A \mid L(A) = \text{IN}\}$, $\text{OUT}(L)$ for $\{A \mid L(A) = \text{OUT}\}$ and $\text{UND}(L)$ for $\{A \mid L(A) = \text{UND}\}$.

Definition 3.10 An *argumentation graph* constructed from a defeasible theory T is a tuple $\langle \mathcal{A}, \rightsquigarrow \rangle$, where \mathcal{A} is the set of all arguments constructed from T , and \rightsquigarrow is an attack relation over \mathcal{A} .

Notation 3.4 Given an argumentation graph $G = \langle \mathcal{A}, \rightsquigarrow \rangle$, we write \mathcal{A}_G , and \rightsquigarrow_G to denote the graph's arguments, and attacks respectively.

Now, let us introduce the notion of a **BP-labelling**, namely a semantics which takes into account a set of burden of persuasion BurdPers , where BurdPers is a set of literals, in determining the status of arguments.

Definition 3.11 A **BP-labelling** of an argumentation graph G , relative to a set of burdens of persuasion BurdPers , is a $\{\text{IN}, \text{OUT}, \text{UND}\}$ -labelling s.t. $\forall A \in \mathcal{A}_G$ with $\text{Conc}(A) = \phi$

1. $A \in L(\text{IN})$ iff
 - (a) $\bar{\phi} \in \text{BurdPers}$ and
 - i. $\forall B \in \mathcal{A}_G$ such that B strictly top-attacks A : $B \in L(\text{OUT})$ and
 - ii. $\forall A' \in \text{DirectSub}(A)$: $A' \in L(\text{IN})$ or
 - (b) $\bar{\phi} \notin \text{BurdPers}$ and
 - i. $\forall B \in \mathcal{A}_G$ such that B top-attacks A : $B \in L(\text{OUT})$ and
 - ii. $\forall A' \in \text{DirectSub}(A)$: $A' \in L(\text{IN})$
2. $A \in L(\text{OUT})$ iff
 - (a) $\phi \in \text{BurdPers}$ and
 - i. $\exists B \in \mathcal{A}_G$ such that B top-attacks A and $B \notin L(\text{OUT})$ or
 - ii. $\exists A' \in \text{DirectSub}(A)$ such that $A' \notin L(\text{IN})$ or
 - (b) $\phi \notin \text{BurdPers}$ and
 - i. $\exists B \in \mathcal{A}_G$ such that B strictly top-attacks A and $B \in L(\text{IN})$ or
 - ii. $\exists A' \in \text{DirectSub}(A)$: $A' \in L(\text{OUT})$;
3. $A \in L(\text{UND})$ otherwise.

In Definition 3.11, items 1) and 2) concern conditions for acceptance and rejection, respectively, based on burdens of persuasion.

Condition for acceptance. Item 1.(a) concerns the case in which a burden of persuasion is on the complement $\bar{\phi}$ of the conclusion ϕ of argument A . A counterargument B for $\bar{\phi}$ is disfavoured by the burden of persuasion, while A is favoured. Thus, acceptance of A is not affected by a top-attacker B unless B is a strict top-attacker. Acceptance also require that all strict subarguments of A are IN . Item 1.(b) concerns the case in which the conclusion of argument A is contradicted by a counterargument B on which there is no burden of persuasion. Here, there is no favour for A . Thus, acceptance of A may also be affected whn B is a non-strict top-attacker. Acceptance also require that all direct subarguments of A are IN .

Condition for rejection. Item 2.(a) concerns the case in which the burden of persuasion is on the conclusion of argument A , so that A is disfavoured by the burden of persuasion. Here, the rejection of A may be determined by a counterargument B that is uncertain (UND), and also by any uncertainty on one of A 's direct subarguments. Item 2.(b) concerns the case in which there is no burden of persuasion on the conclusion of argument A .

Here, the rejection of A is only determined by a counterargument B of A that is IN or by a direct subargument of A that is OUT.

Note that the semantic just described does not always deliver a single labelling. This happens in particular in cases involving “team defeat”, or “team strict defeat”, i.e., in cases where argument A strictly attacks C, while being attacked by D, and B strictly attacks D, while being attacked by C. In such a case, both a labelling where A and B are IN and C and D are OUT and a labelling where all such arguments are UND fits the semantics. In all of the following examples, we will focus on the IN-minimal labelling, i.e., on the labelling where such arguments are labelled UND.

Example 4 (Civil law example) According to the description of Example 2, let us consider the following rules (note that we assume that evidence is provided to establish the factual claims at issue, i.e., that the corresponding burdens of production are satisfied).

$e1 : ev_1$ $e2 : ev_2$ $e3 : ev_3$
 $er1 : ev_1 \Rightarrow \neg guidelines$ $er2 : ev_2 \Rightarrow guidelines$ $er3 : ev_3 \Rightarrow harm$
 $r1 : guidelines \Rightarrow dueDiligence$ $r2 : harm, \sim dueDiligence \Rightarrow liable$

We can then build the following arguments:

$A1 : \Rightarrow ev_1$ $A2 : \Rightarrow ev_2$ $A3 : \Rightarrow ev_3$
 $A4 : A1 \Rightarrow \neg guidelines$ $A5 : A2 \Rightarrow guidelines$ $A6 : A3 \Rightarrow harm$
 $A7 : A5 \Rightarrow dueDiligence$ $A8 : A6 \Rightarrow liable$

The argumentation graph and its grounded $\{IN, OUT, UND\}$ -labelling are depicted in Figure 1 (left), in which all arguments are UND except arguments for undisputed facts. The result is not satisfactory, according to the law, since it does not take into account

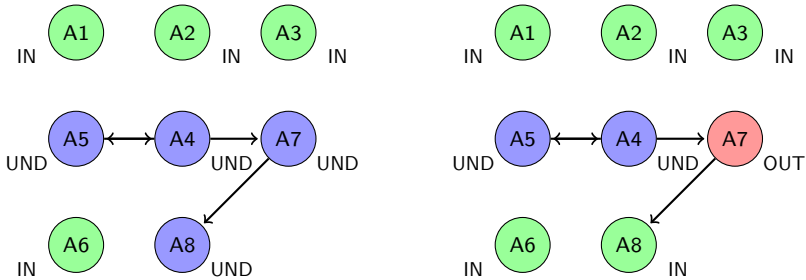


Figure 1. Grounded $\{IN, OUT, UND\}$ -labelling of Example 2 in the absence of burdens of persuasion (left) and its BP-labelling with $BurdPers = \{dueDiligence, liable\}$ (right).

the applicable burdens of persuasion. The doctor should have lost the case – i.e., be found liable – since she failed to discharge her burden of proving that she was diligent (non-negligent). The doctor’s failure results from the fact that it remains uncertain whether she followed the guidelines. To capture this aspect of the argument, we need to specify burdens of persuasion. Let us assume that (as under Italian law) we have $BurdPers = \{dueDiligence, liable\}$ (i.e., the doctor has to provide a convincing argument that she was diligent, the patient has to provide a convincing argument for the doctor’s liability). As the burdened doctor’s argument for *dueDiligence* is OUT, her liability can be established even though it remains uncertain whether the guidelines were followed. □

This example shows how the model here presented allows us to deal with the *inversion of the burden of proof*, i.e., a situation in which one argument A is presented for a claim ϕ burdened with persuasion, and A (or a subargument of it) is attacked by a counterargument B whose conclusion ψ is also burdened with persuasion. If no convincing argument for ψ can be found, then the attack fails, and the uncertainty on ψ does not affect the status A .

Example 5 (Criminal law example) According to the description in Example 1, let us consider the following rules (for simplicity's sake, we will not specify the evidence here, but we assume that all factual claims are supported by evidence):

$$\begin{array}{ll}
 f1: \Rightarrow \textit{killed} & f2: \Rightarrow \textit{intention} \\
 f3: \Rightarrow \textit{threatWithWeapon} & f4: \Rightarrow \neg \textit{threatWithWeapon} \\
 r1: \textit{threatWithWeapon} \Rightarrow \textit{selfDefence} & r2: \neg \textit{threatWithWeapon} \Rightarrow \neg \textit{selfDefence} \\
 r3: \textit{selfDefence} \Rightarrow \neg \textit{murder} & r4: \textit{killed}, \textit{intention} \Rightarrow \textit{murder}
 \end{array}$$

with $r3 \succ r4$. We can build the following arguments:

$$\begin{array}{lll}
 A1 : \Rightarrow \textit{killed} & B1 : \Rightarrow \textit{threatWithWeapon} & C1 : \Rightarrow \neg \textit{threatWithWeapon} \\
 A2 : \Rightarrow \textit{intention} & B2 : B1 \Rightarrow \textit{selfDefence} & C2 : C1 \Rightarrow \neg \textit{selfDefence} \\
 A3 : A1, A2 \Rightarrow \textit{murder} & B3 : B2 \Rightarrow \neg \textit{murder} &
 \end{array}$$

In the $\{\text{IN}, \text{OUT}, \text{UND}\}$ -labelling of Figure 2 (left), all arguments are UND except for the undisputed facts. Thus, in the absence of burdens of persuasion, we do not obtain the legally correct answer, namely, acquittal. To obtain acquittal we need to introduce burdens of persuasion. The prosecution has the burden of persuasion on murder: it therefore falls to the prosecution to persuade the judge that there was killing, that it was intentional, and that the killer did not act in self-defence. The BP-labelling is depicted in

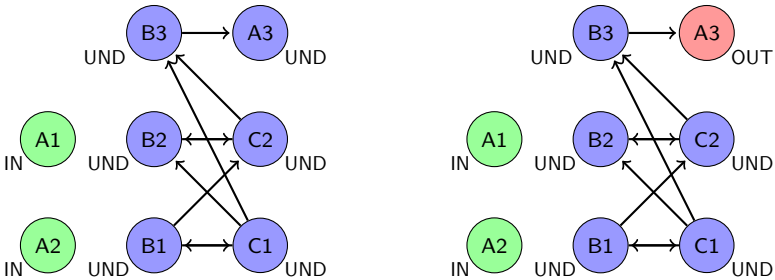


Figure 2. Grounded $\{\text{IN}, \text{OUT}, \text{UND}\}$ -labelling of Example 1 in the absence of burdens of persuasion (left) and BP-labelling with the burden of persuasion $\text{BurdPers} = \{\textit{murder}\}$ (right).

Figure 2 (right). The prosecution failed to meet its burden of proving murder, i.e., its argument is not convincing, since it remains undetermined whether there was self-defence. Therefore, murder is OUT and the presumed killer is to be acquitted. \square

3.3. Adversarial BP

Adversarial BP expands a BP-labelling approach with the idea that failure to meet a burden of persuasion on ϕ does not only mean that any argument for ϕ which fails to be

IN will be OUT. This also means that failure to provide an IN argument for ϕ will lead to $\neg\phi$ being established. For instance, failure to show that the accused is guilty will entail that he should be found innocent. Similarly, the plaintiff's failure to provide a convincing argument that he has a right to compensation for a certain event will entail that he has no right to be compensated. Or the burden of providing a convincing argument that a genetically modified crop is not harmful will entail that the crop is deemed to be harmful. Thus an adversarial burden of persuasion on a claim ϕ entails not only that arguments for ϕ will be OUT if they are not IN, but also that failure to establish ϕ entails ϕ 's complement: " $\sim\phi \Rightarrow \neg\phi$ ". For instance, by adding a rule "*abp1* : $\sim murder \Rightarrow \neg murder$ " we would conclude in the criminal law example above that there is no murder. This is indeed what happens in criminal and other legal cases: failure to establish the prosecution's claim that a murder was committed or the plaintiff's claim that a compensation is due leads to the conclusion that there is no crime or that no compensation is due.

3.4. Reasoning with BPs

In the model described above, BPs are defined outside the legal knowledge base used. What if BPs become part of that rule base, so that we can reason to establish whether or not there is a BP on a literal ϕ .

Notation 3.5 *To specify, within our rule language, that there is a burden of persuasion on a literal ϕ , we write $bp(\phi)$.*

We propose the following definition.

Definition 3.12 *A BP-labelling of an argumentation graph G , relative to burdens of persuasion $BurdPers$, is a $\{IN, OUT, UND\}$ -labelling such that $\forall A \in \mathcal{A}_G$*

1. $A \in L(IN)$ iff
 - (a) *there is an IN argument for $bp(\bar{\phi})$ and*
 - i. $\forall B \in \mathcal{A}_G$ such that B strictly top attacks A : $B \in L(OUT)$ and
 - ii. $\forall A' \in DirectSub(A)$: $A' \in L(IN)$ or
 - (b) *there no IN argument for $bp(\bar{\phi})$ and*
 - i. $\forall B \in \mathcal{A}_G$ such that B top attacks A : $B \in L(OUT)$ and
 - ii. $\forall A' \in DirectSub(A)$: $A' \in L(IN)$;
2. $A \in L(OUT)$ iff
 - (a) *there is an IN argument for $bp(\phi)$ and*
 - i. $\exists B \in \mathcal{A}_G$ such that B top attacks A and $B \notin L(OUT)$ or
 - ii. $\exists A' \in DirectSub(A)$ such that $A' \notin L(IN)$ or
 - (b) *there is no IN argument for $bp(\phi)$ and*
 - i. $\exists B \in \mathcal{A}_G$ such that B strictly top attacks A and $B \in L(IN)$ or
 - ii. $\exists A' \in DirectSub(A)$: $A' \in L(OUT)$;
3. $A \in L(UND)$ otherwise.

Accordingly, bp-statements can be part of the knowledge base or be inferred from it.

Example 6 (Antidiscrimination la example) Consider, for instance, the following formalisation of the European nondiscrimination law in Example 3:

$$\begin{array}{lll}
 e1 : ev1 & e2 : ev2 & e3 : ev3 \\
 er1 : ev1 \Rightarrow \text{indiciaDiscrim} & er2 : ev2 \Rightarrow \neg \text{discrim} & er3 : ev3 \Rightarrow \text{discrim} \\
 r1 : \text{indiciaDscrim} \Rightarrow \text{bp}(\neg \text{discrim}) & &
 \end{array}$$

In this case, since there are *indicia* of discrimination, we can infer that there is the burden of proving nondiscrimination. Then, given that there is uncertainty about whether there was discrimination, the argument for nondiscrimination fails (it is OUT), which means that the argument for discrimination is IN. \square

4. Conclusion

In this paper we provide and discussed a formal model for the burden of persuasion. The model shows how an allocation of the burden of persuasion may lead to a single outcome (IN arguments) in contexts in which the assessment of conflicting arguments would otherwise remain undecided. Our model explores the intersection between the burden of persuasion and argumentation labelling frameworks and provides a starting point for further research. In particular, it combines the insight of [8,9], where the burden of persuasion provides a criterion for adjudicating conflicts of arguments, and the insight of [10,11], where the satisfaction of burdens of argumentation depends on the dialectical status of the arguments at issue. The proposed model also deals with situations in which we have to combine a general burden of persuasion for one party (concerning the top conclusion to be reached), with inversions of the burden relative to specific propositions.

References

- [1] D. Walton, *Burden of proof, presumption and argumentation*, Cambridge University Press, USA, 2014.
- [2] R. Calegari and G. Sartor, Burden of Persuasion in Argumentation, in: *Proceedings 36th International Conference on Logic Programming (Technical Communications), ICLP 2020*, Vol. 325, Open Publishing Association, 2020, pp. 151–163. doi:10.4204/EPTCS.325.21.
- [3] H. Prakken and G. Sartor, A Logical Analysis of Burdens of Proof, *Legal Evidence and Proof: Statistics, Stories, Logic* **1** (2010), 223–253.
- [4] H. Prakken, An Abstract Framework for Argumentation with Structured Arguments, *Argument and Computation* **1** (2010), 93–124.
- [5] S. Modgil and H. Prakken, The ASPIC⁺ framework for structured argumentation: a tutorial, *Argument & Computation* **5**(1) (2014), 31–62.
- [6] M. Caminada and L. Amgoud, On the Evaluation of Argumentation Formalisms, *Artificial Intelligence* **171**(5–6) (2007), 286–310.
- [7] G. Vreeswijk, Abstract Argumentation Systems, *Artificial Intelligence* **90**(1–2) (1997), 225–279.
- [8] H. Prakken and G. Sartor, More on Presumptions and Burdens of Proof, in: *21th Annual Conference on Legal Knowledge and Information Systems*, IOS, Groningen, The Netherlands, 2008, pp. 176–85.
- [9] H. Prakken and G. Sartor, On Modelling Burdens and Standards of Proof in Structured Argumentation, in: *24th Annual Conference on Legal Knowledge and Information Systems*, IOS, 2011, pp. 83–92.
- [10] T.F. Gordon, H. Prakken and D. Walton, The Carneades model of argument and burden of proof, *Artificial Intelligence* **171**(10) (2007), 875–896.
- [11] T.F. Gordon and D.N. Walton, Proof Burdens and Standards, in: *Argumentation in Artificial Intelligence*, I. Rahwan and G.R. Simari, eds, Springer, 2009, pp. 239–60.