Argumentation Schemes for Statutory Interpretation: A Logical Analysis

Giovanni SARTOR[a], Doug WALTON[b], Fabrizio MACAGNO[c], Antonino ROTOLO[d]

[a]*EUI and CIRSFID, University of Bologna, Italy*
[b] *University of Windsor, Canada*
[c]*Universidade Nova de Lisboa, Portugal*
[d]*CIRSFID, University of Bologna, Italy*

Abstract. This paper shows how defeasible argumentation schemes can be used to represent the logical structure of the arguments used in statutory interpretation. In particular we shall address the eleven kinds of argument identified MacCormick and Summers [6] and the thirteen kinds of argument by Tarello [11]. We show that interpretative argumentation has a distinctive structure where the claim that a legal text ought or may be interpreted in a certain way can be supported or attacked by arguments, whose conflicts may have to be assessed according to further arguments.

1. Background: Arguments in Interpretation

This paper aims at developing a fresh formal analysis of interpretive arguments, i.e., arguments meant to support a particular interpretation of a statutory test, and to justify its choice over competing interpretations. Our theoretical framework is based on three different dimensions: legal theories on interpretive arguments, argumentation theories for analysing such arguments, and argumentation logics for formalising them.

In this section we will briefly present the legal doctrinal backgrounds for our formal analysis: MacCormick and Summers's [6] analysis of interpretive arguments, Tarello's [11] classification of interpretation canons, and Alexy and Dreier's [1] analysis of criteria for solving interpretive conflicts. Generally speaking, the so-called interpretation canons — i.e., the different rules to be applied to interpreting statutes — that are employed in legal systems can be viewed as patterns for constructing arguments aimed at justifying certain interpretations, while attacking other interpretations.

1.1. The Proposal of MacCormick and Summers

MacCormick and Summers [6, 464-5], summarising the outcomes of a vast study on statutory interpretation, involving scholars from many different legal systems, distinguish eleven types of arguments:

1. *Arguments from ordinary meaning* express the principle that if a statutory provision can be interpreted according to the meaning a native speaker of a given language would ascribe to it, it should be interpreted in this way, unless there is a reason for a different interpretation.

2. *Arguments from technical meaning* express the principle that if a statutory provision concerns a special activity that has a technical language, it ought to be interpreted in the appropriate technical sense, as opposed to its ordinary meaning.

3. *Arguments from contextual harmonization* express the principle that if the statutory provision belongs to a larger scheme in a statute or set of statutes, it should be interpreted in light of the whole statute it is part of, or in light of other statutes it is related to.

4. *Arguments from precedent* express the principle that if a statutory provision has a previous judicial interpretation, it should be interpreted in conformity with it. Where there is a hierarchy of courts, this principle needs to be applied in such a way to imply that the lower court must conform to the judgment of higher one.

5. *Arguments from analogy* express the principle that if a statutory provision is similar to provisions of other statutes, then it should be interpreted to preserve the similarity of meaning, even if this requires a departure from ordinary meaning.

6. *Arguments from a legal concept* express the principle that if the general legal concept has been previously recognized and doctrinally elaborated in law, it should be interpreted in such a way as to maintain a consistent use of the concept through the system as a whole.

7. *Arguments from general principles* express the principle that whenever general principles, including principles of law, are applicable to the statutory provision, one should favour the interpretation of that is most in conformity with these general legal principles.

8. *Arguments from history* express the principle that if the statute has come to be interpreted over a period of time in accord with the historically evolved understanding of a particular point, its application to a case should be interpreted in line with this historically evolved understanding.

9. *Arguments from purpose* express the principle that if a purpose can be ascribed to a statutory provision, or even to the whole statute, the provision should be interpreted as applied to a particular case in a way compatible with this purpose.

10. *Arguments from substantive reasons* express the principle that if there is some goal that can be considered to be fundamentally important to the legal order, and if the goal can be promoted by one rather than another interpretation of the statutory provision, then the provision should be interpreted in accord with the goal.

11. *Arguments from intention* express the principle that if a legislative intention concerning a statutory provision can be identified, the provision should be interpreted in line with that intention.

1.2. The Proposal of Tarello

Some years before the comparative inquiry of MacCormick and Summers, a list of interpretive arguments was developed by Tarello [11, Ch. 8]. While being based mainly on the Italian tradition, this categorisation has had a broad influence also outside Italy, being adopted in particular by Perelman [9, 55-9]:

1. *Arguments a contrario* exclude interpretations according to which a legal statement explicitly saying "if A then B" is given the meaning "if A or C then B", where C is any proposition not entailed by A.

2. *Arguments a simili ad simile (or analogical)* support interpretations according to which, a term in a legal statetment is extended to include entities that

are not literally included in its scope, but present a relevant similarity with the entities literally included.

    3. *Arguments a fortiori* support interpretations of a legal statement according to which a term in the statement, which apparently denotes a single class of subjects or acts, is extended to other subjects or acts, since these additional subjects or acts deserve to a larger extent the normative qualification linked to that term.

    4. *Arguments from completeness of the legal regulation* exclude interpretations that create legal gaps.

    5. *Arguments from the coherence of the legal regulation* exclude interpretations of different legal statements that make them conflicting.

    6. *Psychological arguments* support interpretations driven by the actual intent of the authors of legal text (for instance, on the basis of the *travaux préparatories*).

    7. *Historical argument*s support interpretations giving a legal statement the same meaning that was traditionally attributed to other statements governing the same matter.

    8. *Apagogical arguments* exclude the interpretation of a normative statement that generates an absurdity.

    9. *Teleological arguments* support the interpretation of a legal statement by attributing to it a rational purpose which is identified from the goals or interests that the law is supposed to promote;

    10. *Parsimony arguments* exclude interpretations that are redundant under the assumption that the legislator does not make useless normative statements.

    11. *Authoritative arguments* support interpretations already given by any authoritative judicial or doctrinal subject.

    12. *Naturalistic arguments* support interpretations aligning a legal statement to human nature or the nature of the matter regulated by that statement.

    13. *Arguments from equity* support (exclude) (un)fair or (un)just interpretations.

    14. *Arguments from general principles* support (exclude) interpretations that are suggested by (incompatible with) general principles of the legal system.

Tarello's list complements MacCormick and Summers' list, since the latter focuses on the kinds of input on which the interpretive argument is based (ordinary language, technical language, statutory context, precedent, etc.) while the first focuses on the reasoning steps by which the interpretive argument is constituted.

1.3. Priorities and Conflicts between Interpretive Arguments

Interpretive argument can be in conflict one against another, leading to opposite conclusions. In fact, as MacCormick [5] observes, "there may be arguments of many types available, and each is capable of generating an interpretation of a given text at variance with that generated by some other possible argument". To address such conflicts, we need to assume or argue that one of the conflicting arguments is stronger than its competitors. Some legal traditions provide indeed general criteria for addressing conflicts of arguments on the basis of their priorities. For instance, Alexy and Dreier [1, 95ff.] indicate various criteria according to which conflicts between interpretive arguments are adjudicated in German law:

1.   In criminal law, arguments based on the wording of the text to be applied have strong priority.

2.   In criminal law arguments based on ordinary meaning take priority over arguments which refer to technical terminology.

3.   A strong priority for the wording obtains where the state wishes to interfere with individual rights of freedom.

4.   A strong priority for the wording obtains for prescriptions on time limits.

5.   A weak priority for arguments based on wording obtains in general.

6.   Genetic arguments, based on the intention of the historical legislator, prevail over argument not based on authority (i.e., argument not based on the historical legislator's intention, on precedent, or on dogmatic consensus), though not over linguistic arguments.

7.   Rightness reasons based on the constitution or on superior sources have precedence over those who are not so grounded.

8.   The idea that a scrutiny is required when limitations to individual liberties are at issue has led some to the idea, refused by others, that substantive arguments based on individual rights prevail over arguments based on collective goals.

1.4. A General Structure for Interpretive Arguments

By analysing the different interpretive canons we can may identify a shared structure including the following elements: an expression E (word, phrase, sentence, etc.) occurs in a document D (statute, regulation, contract, etc.), E has a certain setting S (of ordinary language, technical language, purpose, etc., that) relevant to interpretation, E in D would fit this setting by having interpretation I. If all these elements are satisfied we are licensed to derive the interpretive conclusion that *E ought to be interpreted as I.*

Thus a general pattern for interpretive canons can be expressed as follows: If expression E occurs in document D, E has a setting of S and E would fit this setting of S by having interpretation I, then, E ought to be interpreted as I

By substituting S with the name of any of the argument schemata provided above, we obtain corresponding canons. For instance the ordinary language canon can be expressed as the following schema: If expression E occurs in document D, E has a setting of *ordinary language* and E would fit this setting of *ordinary language* by having interpretation I, then, E ought to be interpreted as *I.*

Consider the issue in the *Dunnachie v Kingston-upon-Hull City Council* case [5, Ch. 7], concerning whether the term "loss" in section 123(1) of the Employment Rights Act 1996 —granting to constructively dismisses employees a right to compensation— only covers pecuniary losses or also includes moral injuries, such as humiliation, injury to feelings and distress. By instantiating the ordinary language canon to the case of the term *loss* in section 123(1) of the Employment Rights Act 1996, we obtain the following conditional: If expression $'loss'$ occurs in document 123(1)ERA, E has a setting of *ordinary language, and E* would fit this setting of *ordinary language* by having interpretation *pecuniary loss*, then, $'loss'$ ought to be interpreted as *pecuniary loss.*

2. A Logical Model for Interpretative Arguments

In this section we shall provide a logical model for interpretive argumentation. We shall model interpretive canons as defeasible rules licensing deontic claims, namely, the claim

that a certain expression ought, ought not, may or may not be interpreted in a certain way. We shall also model meta-canons expressing preferences between interpretive canons. We shall argue that the meaning of interpretive claims is also defined by certain consistency conditions among interpretive claims concerning the same expression. For expressing such conditions we shall use strict rules. Finally we shall address the connection between interpretive claims and common-sense unquestioned meaning postulates, by representing such postulates as modal necessities.

The logical ingredients in our model are the following: (a) an argumentations system providing for strict rules, defeasible rules and priorities between them and defining attack relations between interpretive arguments; (b) a deontic logic for representing interpretive claims; (c) a modal logic for common sense postulates. We shall not provide a full definition of all such logical components, since we aim at providing an abstract approach that can be implemented in different ways according to the chosen logical frameworks.

2.1. Interpretive Canons in an Argumentation System

The first ingredient for our model of arguments is an argumentation system including strict rules, defeasible rules, and preference between rules, such as the system developed by Prakken and Sartor [10], the ASPIC+ system [8], or the Carneades system [3].

We express strict inference rules in the form $\varphi_1, \dots, \varphi_n \rightarrow \psi$, where $\varphi_1, \dots, \varphi_1$ and ψ are formulas in an underlying logical language $\mathcal{L}$. Similarly, we express defeasible inference rules in the form $r: \varphi_1, \dots, \varphi_n \Rightarrow \psi$, where $\varphi_1, \dots, \varphi_1$ and ψ are formulas in an underlying logical language $\mathcal{L}$, and r is the rule name. A defeasible rule r may be read as "if φ_1 and ... and φ_n then presumptively ψ. We assume that all defeasible inference rules have a unique name r. Preferences between rules are expressed by formulas such as $r_1 \succ r_2$ for r_1 is stronger than r_2 and $r_1 \succcurlyeq r_2$ for r_1 is at least as strong as r_2.

As usual, the fact that a rule $r: \varphi_1, \dots, \varphi_n \Rightarrow \psi$ is defeasible means that r, or rather its use in an argument, can be defeated. This may happen in three ways. This first consists in rebutting, i.e., by producing an argument that denies r's conclusion ψ, using a rule $s: \chi_1, \dots, \chi_n \Rightarrow \neg\psi$ that is not weaker than r. The second consists in undercutting r, i.e., in arguing that r does not apply to the case at hand, which we may express by denying that r holds in that case. The third consists in undermining the application of r, i.e., in denying one of its antecedent conditions $\varphi_1, \dots, \varphi_n$, i.e., in claiming that one of these antecedents does not hold.

2.2. Interpretive Oughts

We shall model interpretive conclusions as claims concerning what ought or may be the case in the interpretive domain, i.e., what interpretations should or may be adopted by legal reasoners. For this purpose we employ the usual deontic operators - **O**, for obligation, and **P**, for permission. Thus **O**(φ) is to be read as "it ought to be that φ", and **P**(φ) is to be read as "it may be that φ". We understand **O** and **P** according to standard deontic logic (see [4]): **O**(φ) means that φ is required, being true in all perfect worlds, so that $\neg\varphi$ entails suboptimality (violation), while **P**(φ) means that φ is acceptable, being true in at least one perfect worlds, so that that φ does not entail suboptimality.

Thus, the interpretive claim **O**[E is interpreted as I] is to be read as "it ought to be that expression E is interpreted as I", or "E should be interpreted as I". By making this ought-to-be claim, a legal reasoner asserts that interpreting E as I is required, so that any

different interpretations would be suboptimal, i.e., inferior to I according to the applicable legal standards and values. Correspondingly, the interpretive claim $\mathbf{P}[E$ is interpreted as $I]$ expresses the view that interpreting E as I is acceptable, being not inferior to any other possible interpretations, according to the applicable legal standards and values. Similarly, $\mathbf{O}\neg[E$ is interpreted as $I]$ means that E ought not to be interpreted as I , $\mathbf{P}\neg[E$ is interpreted as $I]$ that E does not need to have this interpretation.

In conclusion, we do not understand interpretive claims as stating enforceable legal obligations or rights, but rather as assertions concerning what interpretations are required or acceptable according to the applicable standards and values. This perspective is not linked to a particular jurisprudence of interpretation, such as the view that there is always or usually a single best interpretation (one right answer); it only assumes that it makes sense to argue why one interpretation should, may, shouldn't, or need not be adopted by legal reasoners and decision-makers, and that it makes sense to provide reasons supporting such claims.

2.3. The Logical Structure of Interpretive Canons

We are now able to provide a general structure for interpretive canons. An interpretive canon is a defeasible rule instantiating the pattern:

$$C: [E \text{ Occurs in } D], [E \text{ has a setting of } S], [\text{interpretation } I \text{ would fit } S]$$
$$\Rightarrow \begin{pmatrix} \mathbf{O} \\ \mathbf{P} \end{pmatrix} \pm [E \text{ in } D \text{ is interpreted as } I]$$

This pattern links an antecedent condition to an interpretive conclusion, through the connective $\Rightarrow$, expressing defeasible entailment:

- The antecedent condition indicates the interpretive condition that has to be satisfied according to the scheme being provided, with regard to the statutory expression E in document D. This condition, according to the general pattern presented in Section 1.4, specifies that the expression E occurs in document D, E has a setting of S, and E would fit this setting of S by having interpretation I.
- The pattern for the conclusion provides for all possible interpretive conclusions, according to the syntax defined by Makinson [7]: $\begin{pmatrix} \mathbf{O} \\ \mathbf{P} \end{pmatrix}$ denotes the alternative between $\mathbf{O}$ and $\mathbf{P}$, while $\pm$ denotes the alternative between negation (represented by the symbol $\neg$ and affirmation, which requires no special symbol.

For example, the canon of ordinary-language interpretation (OL) can be represented, according to the above patterns, as:

$$OL: E \text{ occurs in } D, E \text{ has a setting of } ordinary\ language, \text{interpretation } I \text{ would fit}$$
$$ordinary\ language \Rightarrow \mathbf{O}[E \text{ is interpreted as } I]$$

2.4. Conditions on Interpretations

We need to capture the idea that there is a tension between possible interpretations. In other worlds, the fact that a certain term ought to be interpreted in a certain way (that this is the required interpretation) seems to imply the text should not be interpreted in any different way, neither giving to it a broader coverage, or a smaller one. In fact by affirming that E ought to be interpreted as I_1, we claim that E should be applied to an

object or situation exactly when I_1 applies to it, which is incompatible with the claim that E ought to be interpreted as any I_2, having different extensions or different truth-values than I_1. Thus, we may assert that in general, the claim that an expression ought to be interpreted in certain ways entails that it ought not to be interpreted differently.

$$\mathbf{O}[E \text{ is interpreted as } I_1] \wedge I_1 \neq_m I_2 \rightarrow \mathbf{O}\neg[E \text{ is interpreted as } I_2]$$

This strict inference rule states that if E ought to have interpretation I_1 and I_1 is different in meaning from to I_2 (this is expressed by the symbol $\neq_m$), then E ought not to have interpretation I_2 (we take identity of meaning, $=_m$ as a primitive relation, for the purpose of this paper). Thus for instance in *Dunnachie*, the requirement to interpret the term "loss" in in section 123(1) of the Employment Rights Act as "pecuniary loss" would entail that we should not interpret that term as the broader "pecuniary loss or injury to feelings": according to the axiom scheme above, and assuming that *pecuniary loss or injury to feeling* $\neq_m$ *pecuniary loss*, we obtain that

$$\mathbf{O}[E \text{ is interpreted as } pecuniary\ loss\,] \text{ entails}$$

$$\mathbf{O}\neg[E \text{ is interpreted as } pecuniary\ loss\ or\ injury\ to\ feelings]$$

This idea can be generalised to situations where the interpreter believes that an expression ought to be interpreted in certain alternative ways $I_1, \dots, I_n$, without committing to one of them. This belief entails that any interpretation I_i different from each of these ways ought not to be adopted.

I1: $\mathbf{O}([E \text{ is interpreted as } I_1] \vee \dots \vee [E \text{ is interpreted as } I_n]\,), I_i \neq_m I_1\,, \dots, I_i \neq_m I_n$
$\rightarrow \mathbf{O}\neg[E \text{ is interpreted as } I_i]$

According to this schema, a disjunctive interpretive claim such as

$$\mathbf{O}([E \text{ is interpreted as } I_1] \vee [E \text{ is interpreted as } I_2])$$

entails the rejection of any interpretation that is different from both I_1 and I_2. The claim that an interpretation I may be adopted, i.e., $\mathbf{P}[E \text{ is interpreted as } I\,]$ is consistent with all permissive interpretive statements, as well as with any statement that affirms the obligation of a disjunction of interpretations including also I, while being inconsistent with the obligation to a disjunction of interpretations not including I. Finally, by claiming that an expression ought to be interpreted in at least one of two ways, both being permissible,

$$\mathbf{O}([E \text{ is interpreted as } I_1] \vee [E \text{ is interpreted as } I_2])$$
$$\wedge \mathbf{P}([E \text{ is interpreted as } I_1] \wedge [E \text{ is interpreted as } I_2])$$

we state that all alternative interpretations are forbidden except for these ones.

We may say that our treatment of interpretive oughts assumes that there is an ordering over all possible interpretations of an expression E, and that $\mathbf{O}$ selects the interpretations that are better than all others, while $\mathbf{P}$ selects those that are not worse than any other. Thus $\mathbf{O}[E \text{ is interpreted as } I\,]$ would mean that I is the best of all interpretations of I , ie., that for any other interpretation I_i , $I_i < I$; similarly, $\mathbf{O}([E \text{ is interpreted as } I_1] \vee [E \text{ is interpreted as } I_2])$ would mean that for any interpretation I_i different from I_1 and I_2 either $I_i < I_1$ or $I_i < I_2$. Correspondingly, $\mathbf{P}[E \text{ is interpreted as } I\,]$ would mean that I is not inferior to any other interpretation, i.e., that for any other interpretation I_i, $I_i \leq I$.

2.5. Interpretive Arguments

An interpretive argument can be constructed by combining an interpretive canon with the corresponding interpretive condition. For instance, an argument from ordinary language would have the following form:

$$Argument\ A_1$$

1. $['Loss'$ occurs in $123(1)ERA]$, 2. $['Loss'$ has a setting of *ordinary language*]
3. [Interpretation 'pecuniary loss' would fit *ordinary language*]
4. OL: $[E$ Occurs in $D]$, $[E$ has a setting of *ordinary language*], [interpretation I would fit *ordinary language*] $\Rightarrow \mathbf{O}[E\ in\ D$ is interpreted as $I]$

---

5. $\mathbf{O}['Loss'$ in $123(1)ERA$ is interpreted as 'pecuniary loss']

Interpretive arguments can be attacked by counterarguments. For instance, the following counterargument based on technical language rebuts the argument based on ordinary language, by providing a different incompatible interpretation.

$$Argument\ A_2$$

1. $['Loss'$ occurs in $123(1)ERA]$, 2. $['Loss'$ has a setting of *technical language*]
3. [Interpretation pecuniary loss or injury to feelings would fit *technical language*]
4. TL: $[E$ occurs in $D]$, $[E$ has a setting of *technical language*], [interpretation I would fit *technical language*] $\Rightarrow \mathbf{O}[E\ in\ D$ is interpreted as $I]$

---

5. $\mathbf{O}['Loss'$in $123(1)ERA$ is interpreted as 'pecuniary loss or injury to feelings']

The interpretation based on ordinary language could also attacked by directly denying its conclusion, for instance by a parsimony argument claiming that *loss* should not be interpreted in this was, since this would make $123(1)ERA$ redundant.

An undercutting attack against the ordinary language argument could be mounted by arguing that the expression *loss* in the Employment Rights Act is used in a technical context, namely, industrial relations where arguments from ordinary language do not apply. Thus this canon is inapplicable to the expression *loss* in $123(1)ERA$.

$$Argument\ A_3$$

1. $['Loss'$ occurs in $123(1)ERA]$, 2. $[123(1)ERA$ is a technical context]
3. TC: $[E$ occurs in $D]$, $[D$ is a technical context] $\Rightarrow \neg[OL$ is applicable *to* $E]$

---

3. $\Rightarrow \neg[OL$ is applicable *to* $'Loss']$

Finally, the last attack would consist in undermining the argument, namely, in countering one of its premises. For instance it maybe argued that it is not the case that $'loss'$ is understood as *pecuniary loss* in ordinary language, since most people do not limit this term to this. In this case we would need to instantiate to the case of *Loss* and *Pecuniary loss* the common sense rule C1:

$$C1: \neg[E$ is generally understood as $I] \Rightarrow \neg[I$ would fit *ordinary language*]$$

$$Argument\ A_4$$

1. $\neg['Loss'$generally understood as *Pecuniary loss*]
2. $\neg['Loss'$ is generally understood as *Pecuniary loss*] $\Rightarrow$
$\neg[Pecuniary\ loss$ would fit *ordinary language*]

---

3. $\neg[Pecuniary\ loss$ would fit *ordinary language*]

2.6. Preferences between Interpretive Arguments

We may have preferences between interpretive arguments. For example, in Italy the Court of cassation revised its interpretation of the term *Loss* (*danno*) as occurring in the Italian Civil code (ICC) using an argument from substantive reasons (the constitutional value of health): the court thus rejected the previous scope of the term as including only *pecuniary damage*, and expand it to compensate also any *damage to health*:

$$Argument\ A_1$$

1. $[Loss$ occurs in $ICC]$, 2. $[Loss$ has a setting of $legal\ history]$
3. [Interpretation *pecuniary loss* would fit *legal history*]
4. LH: $[Loss$ occurs in $Era]$, $[E$ has a setting of $legal\ history]$, [interpretation I would fit *legal history*] $\Rightarrow \mathbf{O}[E\ in\ D$ is interpreted as $I]$
_____
5. $\mathbf{O}['Loss'$ in ICC is interpreted as *Pecuniary* loss]

$$Argument\ A_2$$

1. $['Loss'$ Occurs in $ICC]$, 2. $['Loss'$ has a setting of *substantive reasons*]
3. [Interpretation *pecuniary loss or damage to health* would fit *substantive reasons*]
4. SR: $[E$ Occurs in $D]$, $[E$ has a setting of *substantive reasons*], [interpretation I would fit *substantive reasons*] $\Rightarrow \mathbf{O}\ [E\ in\ D$ is interpreted as $I]$
_____
5. $\mathbf{O}[Loss$ in ICC is interpreted as *pecuniary loss or damage to health*]

Given these two conflicting arguments, the judges argued that the second argument defeats the first, since SR concerns constitutional values, and this canon, when applied to constitutional values, prevails over the historical argument.

2.7. Interpretations as Concepts

As Araszkiewicz has recently argued [2], interpretive statements may be viewed as concerning sets, rather than sequences of words. Let us assume that $'PecuniaryLoss'$ denotes the set of pecuniary losses and $'InjuryToFeelings$' denotes the set of injures to feelings. Then an interpretation of $123(1)ERA$ could be stated as

$$1.\mathbf{O}['Loss'\text{in }123(1)ERA \text{ is interpreted as } PecuniaryLoss]$$

The alternative interpretation could be stated as

$$2.\mathbf{O}['Loss'\text{in }123(1)ERA \text{ is interpreted as } PecuniaryLoss\ \cup\ InjuryToFeelings\ ]$$

where $\cup$ denotes set-theoretical union. The use of the set-theoretical language would enable us to express a broader set of interpretive claims, such as the claim that it ought to be than any interpretation of 'Loss' includes also injuries to self-respect.

$$3.\mathbf{O}(['Loss'\text{in }123(1)ERA \text{ is interpreted as } X]\ \Rightarrow\ InjuryToSelfRespect\ \subseteq X)$$

Now, assume that the first interpretation (as *PecuniaryLoss*) is claimed to be obligatory (according to 1 above). Given that in standard deontic logic $\mathbf{O}A$ and $\mathbf{O}$ (A $\Rightarrow B$) entail $\mathbf{O}$ B, from 1 and 3 we get

$$4.\boldsymbol{O}(InjuryToSelfRespect\ \subseteq\ PecuniaryLoss)$$

Injuries to self-respect are not a subset of pecuniary losses, this being a conceptual truth, which cannot be changed through interpretation[1].

_____

[1] Conceptual assumptions pertaining to unquestioned commonsense background are expressed by axioms in the form of necessary logical relations: such relations are indisputable in the domain of interpretation at hand.

5. $\mathbf{N}(InjuryToSelfRespect \nsubseteq PecuniaryLoss)$

Then we get a contradiction, since what is deontically required cannot violate (in standard deontic logic) a necessary constraint. Thus an argument to the clam that any interpretation of 'Loss' ought to include injuries to self-respect contradicts the claim according to which 'Loss' ought to be interpreted as pecuniary loss.

3. Conclusion

We have provided a fresh formal framework that can capture the interpretive arguments identified by MacCormick and Summers [6] and Tarello [11], as well as criteria for arguing about priorities over interpretive arguments, as discussed by Alexy and Dreier [1]. With regard to argumentation logics, we have not committed to any specific framework, although two features are technically needed for our purposes: (a) a language able to represent the internal logical structure of arguments, (b) a mechanism to argue about priorities over arguments (as in [8]). With regard to interpretive claims we have argued that standard deontic logic provides an adequate framework for modeling them, once integrated with constraints expressing the incompatibility of claims concerning interpretive oughts. Finally, we have considered that interpretations can also be viewed as concerning concepts rather than terms, following an idea in [2].

This work still is quite preliminary, but we hope that it can contribute to the analysis of the logical structures for addressing arguments for statutory interpretation, an aspect of legal reasoning that have so far been neglected in legal logic and in AI & law. Further research will include a more refined categorization, for instance, according to the typology of Walton, Macagno and Reed [12], as well as a development of the logical framework, which requires merging argumentation, deontics and conceptual constraints.

References

[1] R. Alexy and R. Dreier, Statutory interpretation in the Federal Republic of Germany, in [6, pp. 72-121].

[2] M. Araszkiewicz, Towards systematic research on statutory interpretation in AI and law, in *Proc. of JURIX 2013,* pp. 15-24, IOS Press, 2013.

[3] T. Gordon and D.N. Walton, The Carneades argumentation framework - using presumptions and exceptions to model critical questions, in *Proc. of COMMA-06*, pp. 195-207, IOS Press, 2006.

[4] R. Hilpinen, Deontic logic, in *The Blackwell Guide to Philosophical Logic*, pp. 159-82, Blackwell, 2001.

[5] D.N. MacCormick, *Rhetoric and the Rule of Law*, Oxford University Press, 2005.

[6] D.N. MacCormick and R.S. Summers, R.S. (eds.), *Interpreting Statutes: A Comparative Study*, Darthmouth, 1991.

[7] D. Makinson, On the Formal Representation of Rights Relations. *Journal of Philosophical Logic* 15:403–25, 1986.

[8] S. Modgil and H. Prakken, Reasoning about preferences in structured extended argumentation frameworks, in *Proc. of COMMA 2010*, pp. 347-58, IOS Press, 2010.

[9] C. Perelman, *Logique juridique: Nouvelle rhétorique*, Dalloz, 1979.

[10] H. Prakken and G. Sartor, A dialectical model of assessing conflicting arguments in legal reasoning, *Artificial Intelligence and Law* 4:331–68, 1996.

[11] G. Tarello, *L'interpretazione della legge*, Giuffrè, 1980.

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