University of Windsor

Scholarship at UWindsor

CRRAR Publications

Centre for Research in Reasoning, Argumentation and Rhetoric (CRRAR)

2016

Contested Cases of Statutory Interpretation

Douglas Walton
University of Windsor, Centre for Research in Reasoning, Argumentation and Rhetoric

G. Sartor

F. Macagno

Follow this and additional works at: https://scholar.uwindsor.ca/crrarpub

Part of the Arts and Humanities Commons

Recommended Citation

Walton, Douglas; Sartor, G.; and Macagno, F.. (2016). Contested Cases of Statutory Interpretation. *Artificial Intelligence and Law*, 24 (1), 51-91.
https://scholar.uwindsor.ca/crrarpub/33

This Article is brought to you for free and open access by the Centre for Research in Reasoning, Argumentation and Rhetoric (CRRAR) at Scholarship at UWindsor. It has been accepted for inclusion in CRRAR Publications by an authorized administrator of Scholarship at UWindsor. For more information, please contact scholarship@uwindsor.ca.

An Argumentation Framework for Contested Cases of Statutory Interpretation

D. Walton, G. Sartor and F. Macagno, *Artificial Intelligence and Law*, 24(1), 2016, 51-91.

Abstract: This paper proposes an argumentation-based procedure for legal interpretation, by reinterpreting the traditional canons of textual interpretation in terms of argumentation schemes, which are then classified, formalized, and represented through argument visualization and evaluation tools. The problem of statutory interpretation is framed as one of weighing contested interpretations as pro and con arguments. The paper builds an interpretation procedure by formulating a set of argumentation schemes that can be used to comparatively evaluate the types of arguments used in cases of contested statutory interpretation in law. A simplified version of the Carneades Argumentation System is applied in a case analysis showing how the procedure works. A logical model for statutory interpretation is finally presented, covering pro-tanto and all-things-considered interpretive conclusions.

Keywords. Argumentation systems, interpreting statutes, ordinary language meaning, argument from purpose, ambiguity, abductive reasoning, argumentation schemes.

1. Introduction

This paper is part of a larger research project that has the goal of showing how argumentation tools can be used to represent and assess claims that a word or term in a legal text should be interpreted in one way rather than in another.[1] One goal of the project is to extend the existing literature on argumentation schemes by building new schemes specially designed for textual interpretation in law and testing them on cases of arguments used in problem cases of statutory interpretation in law. Another goal is to adapt and develop existing computational resources for conflicts in legal argumentation where one interpretation of a statutory text is pitted against another. The ultimate goal is to provide a framework in which such arguments can be evaluated.

Interpretation is regarded as the passage from a legal text to a legal rule (Hage, 1996, p. 214; Tarello, 1980) – corresponding to the passage in linguistics from the linguistic content to communicative intention (Bezuidenhout, 1997; Capone, 2009; Carston, 2002, 2013; Dascal, 2003; Levinson, 2000; Soames, 2008; Sperber & Wilson, 1986; Wilson & Sperber, 2004) – and more precisely "an ascription of meaning to a linguistic sign in the case its meaning is doubtful in a communicative situation, i.e. in the case its 'direct understanding' is not sufficient for the communicative purpose at hand" (Dascal & Wróblewski, 1988, p. 204). In a case where is an "eventual 'mismatch' between the 'computed' utterance-meaning and some contextual factor" resulting from the background or the specific case to which the law is applied (Dascal & Wróblewski, 1988, p. 213; 216), interpretation needs to be justified through reasons (Atlas & Levinson, 1981; Atlas, 2008; Dascal, 2003, p. 635), namely arguments (Macagno & Capone, 2015) why a particular rule, rather than another, is valid on the basis of the statutory text (Hage, 1996, p. 215). In statutory interpretation, such arguments are usually based on specific maxims of interpretation, which can be translated into formal language (Hage, 1997).

In this paper, we will show how the canons of interpretation can be represented as argumentation schemes, namely prototypical patterns of defeasible argument (Macagno & Walton, 2015; Walton, Reed, & Macagno, 2008). These patterns can be classified depending on their purpose and preferential use, and can thus be modeled by combining the formal

[1] For previous work on the project see (Macagno, Walton and Sartor, 2012), (Sartor, Walton, Macagno and Rotolo, 2014), and (Walton, Macagno and Sartor, 2014).

argumentation system ASPIC+ with a logical language (Sartor, Walton, Macagno, & Rotolo, 2014). The focus of this work is the relationship between an interpretation of an expression (or rather the use of an expression) within a legal text and the reasons provided in its support. For this reason, the logical structure that will be developed will not be framed in deontic terms, but rather will concern terminological assertions concerning what should count as the best interpretations of the contested or potentially contested expressions. This technique can be extended by building a procedure of scheme-based argument mapping that summarizes and reveals the structure of complex networks of argumentation in legal cases. It is shown how formal argumentation systems such as ASPIC+ and the Carneades Argumentation System (CAS) can be extended to deal with interpretative argumentation.

The plan of the paper is to outline the legal theory resources first and then apply them to illustrate a typical example of a legal case where the outcome depends on competing statutory interpretations. Section 2 identifies eleven types of arguments recognized by MacCormick and Summers (Alexy & Dreier, 1991), and fourteen types from Tarello (Tarello, 1980) commonly used in statutory interpretation. Section 3 sets up a structure for modeling interpretive schemes generally. Section 4 analyzes the *a contrario* scheme using this structure. Section 5 shows how argumentation utilizing these schemes can be modeled in the Carneades Argumentation System (CAS). Section 6 and 7 offer two case studies using CAS. Section 8 sets in place a general structure on which to base the logical model for reasoning with interpretive canons built in section 9. Section 10 introduces preferences over interpretive arguments. Section 11 discusses the transition from interpretive claims to assertions concerning specific individuals. Section 12 offers some conclusions.

2. Types of Interpretive Arguments

2.1. The existing types of interpretive arguments

Macagno, Walton and Sartor (2012) compiled a list of eleven interpretive arguments identified by MacCormick and Summers (1991). Below, each type of argument recognized in that prior list is explained in a condensed manner to give the readers some idea of how each of them can be reconfigured as a distinct defeasible form of argument.

- *Argument from ordinary meaning* requires that a term should be interpreted according to the meaning that a native speaker would ascribe to it.
- *Argument from technical meaning* requires that a term having a technical meaning and occurring in a technical context should be interpreted in its technical meaning.
- *Argument from contextual harmonization* requires that a term included in a statute or set of statutes should be interpreted in line with whole statute or set.
- *Argument from precedent* requires that a term should be interpreted in a way that fits previous judicial interpretations.
- *Argument from analogy* requires that a term should be interpreted in a way that preserves the similarity of meaning with the term's occurrences in similar provisions of other statutes.
- *Argument from a legal concept* requires that a term should be interpreted in line with the way it has been previously recognized and doctrinally elaborated in law.

- *Argument from general principles* requires that a term should be interpreted in a way that is most in conformity with general legal principles already established.
- *Argument from history* requires that a term should be interpreted in line with the historically evolved understanding of it.
- *Argument from purpose* requires that a term should be interpreted in a way that fits a purpose that can be ascribed to the statutory provision, or whole statute, in which the term occurs.
- *Argument from substantive reasons* requires that a term should be interpreted in line with a goal that is fundamentally important to the legal order.
- *Argument from intention* requires that a term should be interpreted in line with the intention of the legislative authority.

These eleven types of interpretive argument are comparable to and overlap with the fourteen types previously identified by Tarello (Tarello, 1980, Chapter 8), listed as follows in (Sartor et al., 2014, pp. 12–13):

- *Arguments a contrario* reject interpretations of a term departing from the term's literal meaning.
- *Analogical arguments* support interpretations according to which a term is extended to cover entities that are not included in its literal scope, but present a relevant similarity with the entities literally included.
- *Arguments a fortiori* support interpretations extending the meaning of a term, which literally denotes a single class of entities, to other entities that deserve, to a higher degree, the same discipline as the entities literally included.
- *Arguments from completeness of the legal regulation* exclude interpretations that create legal gaps.
- *Arguments from the coherence of the legal regulation* exclude interpretations of different legal statements that make them conflicting.
- *Psychological arguments* support interpretations driven by the actual intent of the authors of legal text.
- *Historical argument*s support interpretations giving a legal statement the same meaning that was traditionally attributed to other statements governing the same matter.
- *Apagogical arguments* exclude interpretations that generate absurdities.
- *Teleological arguments* support interpretations contributing to it a purpose pertaining to the goals or interests that the law is supposed to promote.
- *Non-redundancy arguments* exclude interpretations that would make the interpreted expression redundant, under the assumption that the legislator does not make useless normative statements.
- *Authoritative arguments* support interpretations already given by authoritative courts or scholars.
- *Naturalistic arguments* support interpretations aligning a legal statement to human nature or the nature of the matter regulated by that statement.
- *Arguments from equity* support (exclude) (un)fair or (un)just interpretations.
- *Arguments from general principles* support (exclude) interpretations that are supported by (incompatible with) general principles of the legal system.

The two lists complement each other, even though Tarello's list emphasizes the kinds of input on which interpretive argument are based, such as ordinary language, technical language, and so forth, while MacCormick and Summers' list emphasizes the reasoning steps involved in the interpretive process.

In comparing the two lists of types of interpretive arguments some common elements stand out, but there are also significant differences. Some of the argument types inTarello's list—such as analogical arguments, teleological arguments and arguments from general principles— appear to be already included in the list of MacCormick and Summers. Tarello's psychological arguments seem to fit under McCormick and Summers' category of argument from intention. It looks like Tarello's authoritative arguments might fit under MacCormick and Summers' category of argument from precedent. Others types of argument are distinctively different, while in still other cases it is unclear how the type of interpretive argument described in the one list is related to the type described in the other list.

One of the crucial problems concerning types of interpretive arguments is their use (in training legal practitioners or scholars) and their relations with the works in argumentation theory and logic on argument analysis and reconstruction. Recently the canons or maxims that express the general principle characterizing each type of argument have been represented as defeasible rules, to be integrated within a prioritized defeasible logic system (Rotolo, Governatori, & Sartor, 2015). The purpose of this paper is to analyze types of interpretive arguments as argumentation schemes, or rather dialogical patterns of arguments, in which an interpretation is regarded as a defeasible viewpoint that needs to be supported by a pattern of reasoning and can be subject to default in case specific critical questions are successfully advanced. On this perspective, interpretive reasoning is framed within a broader dialectical framework, involving a specific burden of bearing out and defeating a specific interpretation (Gizbert-Studnicki, 1990).

Some of the interpretive argumentation schemes in both lists clearly relate to argumentation schemes already widely known and studied in argumentation that are not specifically designed to deal with interpretive issues (Macagno & Walton, 2015; Walton et al., 2008). Hence there are many questions about how some of the new interpretive schemes relate to these more general schemes that have been already widely recognized. For example, the category of authoritative arguments in Tarello's list might relate to the scheme for argument from expert opinion. Since laws formulated in statutes are binding on the courts, it can be said that the statement made in this context can be held to hold by reason of authority. But a legal scheme for argument from administrative authority that is a variant on argument from authority already has some recognition in the field of argumentation studies. Hence there are questions raised about how this new interpretive scheme proposed by Tarello distinguishes between the two kinds of argument from authority. As mentioned above there is also the question of how Tarello's version of interpretive argument from authority fits in with schemes from MacCormick and Summers' list such as argument from precedent, argument from a legal concept, argument from general principles and argument from history. None of these questions can be discussed in this paper, for reasons of length, but they need to be recognized here as problems for future research.

Another similar problem is how the interpretive argument from precedent, as it is called in MacCormick and Summers' list, is related to the general scheme for argument from precedent, already recognized in the argumentation literature. The problem is that there are great divisions of opinion on precisely how the scheme should be modeled. Many think that argument from

precedent is always based on argument from analogy, that is, on a comparison between and source case and a target case. But others might think that legal argument from precedent needs to be based on *ratio decidendi*. Another question raised by this difference of opinion is whether *ratio decidendi* represents some kind of analogy between the two cases where the rationale used to arrive at the conclusion in the source case is supposed to be similar to a comparable rationale that can fit the target case.

In this paper we recognize the existence of these problems without delving into a detailed analysis thereof, so that we can forge ahead with building a framework for interpretive argumentation schemes that can later be applied to studying specific schemes and issues. The starting point is to provide a general classification of the most important arguments of the two lists, identifying the more generic identities between them. Then we move through a sequence of examples of legal arguments where interpretation of a statute or law is an issue, applying the model to the examples. As always, the work of applying formal structures to real cases of argumentation in natural language discourse raises problems and difficulties in its own right.

2.2. Classifying the interpretive arguments

MacCormick (2005, pp. 124–125) proposed that there are three main categories of interpretive argument, over the above eleven categories of interpretive arguments acknowledged as persuasive in grounding a selected interpretation of a text in a disputed case in a broad variety of legal systems. First, there are so-called *linguistic arguments* that appeal to the linguistic context itself to support an interpretation (which we can call definitional arguments, Macagno & Walton, 2014). Second, there are the *systemic arguments* that take the special context of the authoritative text, within the legal system into account. Such schemes merge the authority of the source with the reconstruction of the definition from the text. Third, there are the *teleological-evaluative arguments* that make sense of the text in light of its aim or goal (which we can refer to as pragmatic arguments, see Macagno & Walton, 2015). A fourth category is what McCormick (2005) calls "*appeal to the lawmaker's intention*." McCormick does not consider this type of interpretive argument alongside the other main categories of interpretive argument, because of the ambiguity and indeterminacy of the notion of intention. He rather views it a trans-categorical type of argument that ranges across all the other categories and their types, as linguistic, systemic or teleological-evaluative considerations that can support the attribution of intentions to legislators.

If we try to analyze the lists of arguments in terms of patterns of argument, explaining the arguments of legal interpretation using the categories of argumentation schemes, we need to draw a first crucial distinction between arguments that support an interpretation and arguments that reject an interpretation. Some interpretive canons, however, are bivalent, in the sense that they provide for two interpretive schemes: one (positive or negative) when the canon's condition is satisfied, and the opposite (negative or positive) when the canon's condition is not satisfied. For instance, while the contextual coherence of an interpretation supports the adoption of an interpretation, lack of contextual coherence supports rejection. In such cases we use the symbol + and - to denote the use of a scheme to support and reject an interpretation.

The arguments supporting an interpretation are different in nature (Macagno, 2015). Pragmatic arguments, definitional (of different types, including the systemic ones), and analogical arguments represent distinct reasoning patterns, which are often merged with authority arguments. Such arguments are intended to back up a specific definition based on

previous interpretations (epistemic authority) or on the reconstruction of a possible "intention" of the lawmaker (deontic authority), or on the alleged "nature" of a concept (the commonly shared definition). Such categories often merge with each other, but they can be classified in Figure 1 based on a distinctive feature, namely their distinctive reasoning pattern.

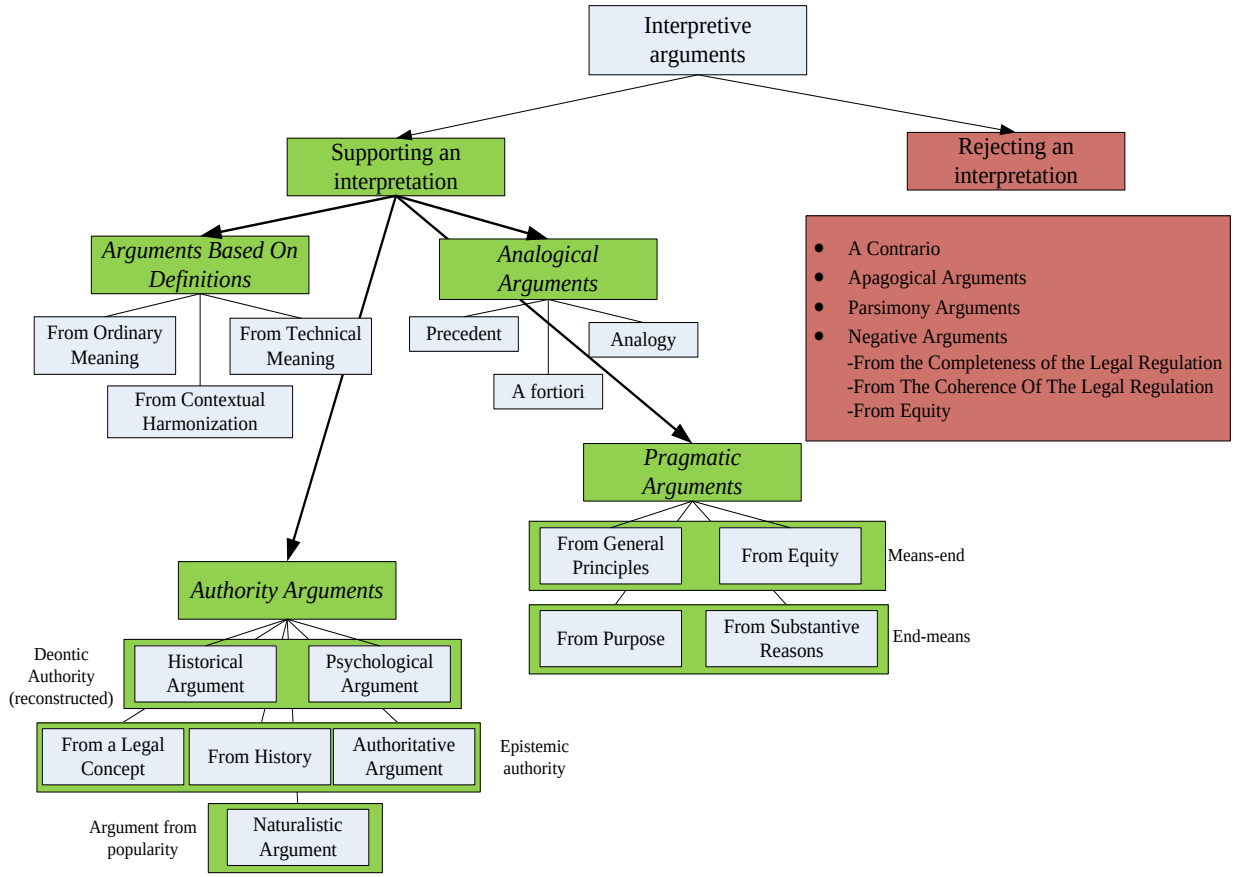

Figure 1: Classifying the arguments of interpretation

It was recognized by MacCormick (2005) that there can be conflicts between interpretive arguments, pitting one form of interpretive argument against another (Rotolo et al., 2015). Some legal traditions provide general criteria for dealing with conflicts of this sort based on certain kinds of priorities. Alexy and Dreier (Alexy & Dreier, 1991, pp. 95–98) have cited criteria such as the following: (a) in criminal law arguments from ordinary meaning have priority over arguments from technical meaning; (b) in criminal law generic arguments based on the intention of the legislator have priority over arguments not based on authority, but not over linguistic arguments. In this paper, we will use argumentation tools to represent such conflicts and priorities.

3. The Basic Interpretative Argumentation Schemes

Statutes are written in natural language. Our concern is with the interpretation of sentences expressed in natural language that are susceptible to differing interpretations (Atlas, 2005; Horn, 1995). The major philosophical concern is how the notion of meaning is to be defined in relation to the task of finding the evidential basis for preferring one interpretation or another (Atlas &

Levinson, 1981; Atlas, 2005; Dascal, 2003, p. 635). In this paper, we find it most highly suitable to adopt a pragmatic approach to meaning, namely to understand statutory meaning as the intention expressed through the legal text (Carston, 2013), an approach that corresponds to the trans-category understanding of interpretation in McCormick (2005). The syntax representing the structure of a sentence, as well as the individual semantic meanings of each term contained in the sentence, are important. But over and above such factors, it needs to be acknowledged that the meaning of the sentence composed of these elements, especially in the examples considered in this paper, need to be placed in the context of a broader text or corpus in which it is embedded. For example, the issue of whether a contested word should be taking it as expressing and ordinary meaning or a technical meaning is a dispute about whether the word can be interpreted the one way or the other in a special context of use. For these reasons, although we acknowledge the importance of semantics and syntax in matters of statutory interpretation, we need to study the notion of meaning in a broad manner to include not only these aspects, but also the aspect of the placement of the sentence in a broader context of use in different kinds of discourse.

From our perspective, making an interpretation consists in associating a linguistic occurrence and a meaning within a specific context and use, i.e., in claiming that a certain expression *E* in certain document *D* has a certain meaning *M*. Interpretations are not necessarily correct. They may be right or wrong, preferable or not to other interpretations.

We shall model the application of interpretation canons by using a uniform template, so that for each canon we obtain an argument scheme including a major premise, a minor premise and an interpretive conclusion.

- The major premise is a general canon: if interpreting an expression (word, phrase, sentence) in legal document (source, text, statute) in a certain way satisfies the condition of the canon issue, then *the expression* should/ should-not be interpreted (depending on whether the canon is a negative or positive one) in that way.
- The minor premise is a specific assertion: interpreting a particular expression in a particular document in a certain way satisfies the condition of the canon.
- The conclusion is a specific claim: the particular expression in that document indeed should / should not be interpreted in that way.

In this paper we shall apply this template to provide schemes for the following canons: (1) argument from ordinary language (*OL*), (2) argument from technical language, whose requirement is correspondence to technical language, (*TL*), (3) *a contrario* argument (*AC*), (4) argument from purpose, (*Pu*), (5) argument from precedent, (*Pr*), (6) argument from contextual harmonization (*CH*). This list of schemes will be added to as new schemes are formulated. Here is our system of notation for labeling the nodes in an argument diagram in sections 6 and 7 to indicate a scheme. We use + for schemes uses to argue for an interpretation. We use - for schemes used to argue against an interpretation, +e for exclusion, and +i for inclusion. So, for example the notation +*iPr* labels a pro inclusive argument from precedent.

3.1 Positive interpretive schemes

As mentioned above, two fundamental macro-categories of interpretive argument schemes need to be distinguished, the positive ones supporting an interpretation, and the negative onesrejecting an interpretation. Here is the template for positive interpretive argument schemes. In presenting this template we shall use uppercase letters for variables and lowercase letters for constants

Major Premise:	*C*: If the interpretation of *E* in a *D* as *M satisfies C's condition*, then *E* in *D* should be interpreted as *M*.
Minor premise:	The interpretation of *e* in *d* as *m satisfies C's condition*.
Conclusion	*e* in *d* should be interpreted as *m*.

In applying this template we need to substitute in the major premise the condition that characterizes a particular canon, for instance, fitting ordinary language (OL).

In order to show how positive interpretive canons can be applied with this pattern, we use the case of *Dunnachie v Kingston-upon-Hull City Council*, also used by MacCormick (2005), as a running example. This case concerns an employee who claimed to have been unfairly dismissed, and as a result to have suffered humiliation, injury to feelings and distress. The employer argued that the relevant section of the current UK legislation, called the Employment Rights Act of 1996, only permits recovery of *financial loss*. The employee argued that a proper construction of the relevant section of the statute allows for recovery of *losses* other than financial losses narrowly construed. The question posed was whether the term '*loss*', as used in the statute, referred only to financial loss or could be given a more extended meaning so that it included losses such as emotional loss that are not strictly financial.

If we use the canon *Ordinary Language*, we obtain the following structure:

Major Premise:	*OL*: If The interpretation of *E* in *D* as *M* fits *ordinary language*, then *E* in *D* should be interpreted as *M*.
Minor premise:	The interpretation of "*loss*" in *Employment Relations Act* as *PecuniaryLoss* fits *Ordinary Language*
Conclusion	"*loss*" in the *Employment Relations Act* should be interpreted as *PecuniaryLoss*.

Note that we use inverted commas for linguistic occurrences ("*loss*") and a single word, with capitalized initials for meanings (*PecuniaryLoss*).

By substituting the conditions of the *OL* canon, with the requirement of other canons listed above it is possible to generate other interpretation schemes. For instance, we can obtain the following scheme for *Technical Language* (*TL*):

Major Premise:	*TL*: If the interpretation of *E* in *D* as *M fits technical language*, then *E* in *D* should be interpreted as *M*.
Minor premise:	The interpretation of "*loss*" in the *Employment Relations Act* as *PecuniaryOrEmotionalLoss* fits technical language.
Conclusion	"*loss*" in the *Employment Relations Act* should be interpreted as *PecuniaryOrEmotionalLoss*.

Obviously, our interpretive schemes only provide the top-level step in the reasoning that is needed to apply an interpretive canon. For supporting the application of a canon we need to establish the minor premise of the corresponding scheme, namely, to show that the interpretation we are proposing indeed satisfies the canon we are considering. This requires specific arguments, according to particular scheme being considered. For instance, for establishing that interpretation "*pecuniary loss*" of expression "*loss*" in document *Employment Relations Act* fits canon

Ordinary Language, we will have to establish, by providing adequate evidence, that this interpretation matches the current linguistic usage. Thus, for instance, to support the application of the *Ordinary Language* canon, we would need an inference like the following:

Major Premise:	If *E* is commonly understood as *M*, then the interpretation of *E* in *D* as *M* fits ordinary language.
Minor Premise:	The "*loss*" is commonly understood as *PecuniaryLoss*.
Conclusion	The interpretation of "*loss*" in *Employment Relations Act* as *PecuniaryLoss* fits ordinary language.

Here the minor premise is a substitution instance of the antecedent of the major premise.

4. Negative interpretive schemes

According to negative canons, if an interpretation meets the canon's condition, then it is to be rejected.

Major Premise:	*C*: If the interpretation of *E* in *D* as *M satisfies condition of C's canon*, then *E* in *D* should not be interpreted as *M*.
Minor Premise:	The interpretation of *e* in *d* as *m satisfies condition of negative canon*.
Conclusion	*e* in *d* should not be interpreted as *m*.

The most common negative canon is the *a contrario* (*AC*), which rejects an interpretation which is over- or underinclusive with regard to the usual semantic meaning of that expression, according to the idea that *Ubi lex voluit, dixit; ubi noluit, tacuit* (what the law wishes, it states, what the law does not want, it keeps silent upon). The *A Contrario* canon can also be viewed as a counterfactual appeal to the intention of the legislator: if the legislator had meant to express a meaning that is different from the usual meaning (the semantic meaning) of the expression at issue, he would have used a different expression. Here is for instance an example of application of the *a contrario* canon.

Major Premise:	*AC*: If the interpretation of *E* in *D* as *M* conflicts with the usual meaning of *E* (is over or under-inclusive), then *E* in *D* should not be interpreted as *M*.
Minor Premise:	The interpretation of the expression "*loss"* in the *Employment Relations* as *PecuniaryOrEmotionalLoss* conflicts with the usual meaning of "*loss"*.
Conclusion	"*loss*" in *Employment Relations Act* should not be interpreted as *PecuniaryOrEmotionalLoss*.

There is also a more specific kind of *A Contrario* argument, which we may call subclass *a contrario*: rather than rejecting an interpretation as a whole, it addresses the exclusion or inclusion of a certain subclass *S* in the interpretation at issue, on the basis of the fact that the subclass is included in or excluded from the usual meaning. Here are the two variants: the exclusionary *a contrario* (*eAC*) and the inclusionary *a contrario* (*iAC*). Note that the *iAC* has a positive interpretive conclusion, as the non-exclusion, i.e., the non-non-inclusion is an inclusion.

Here is the first variant, namely, the exclusionary *a contrario* argument.

Major Premise:	*eAC*: If the interpretation of *E* in *D* as including *S* conflicts with the usual meaning of *E*, then *E* in *D* should be interpreted as excluding *S*.
Minor Premise:	The interpretation of "*loss"* in the *Employment Relations* as including *EmotionalLoss* conflicts with the usual meaning of "*loss"*.
Conclusion	"*loss*" in *Employment Relations Act* should be interpreted as excluding *EmotionalLoss*.

Here is the second variant, the inclusionary *a contrario* argument.

Major Premise:	*iAC*: If the interpretation of *E* in *D* as excluding *S* conflicts with the usual meaning of *E*, then *E* in *D* should be interpreted as including *S*.
Minor Premise:	The interpretation of "*loss"* in the *Employment Relations* as excluding *EmotionalLoss* conflicts with the usual meaning of "*loss"*.
Conclusion	"*loss*" in *Employment Relations Act* should be interpreted as including *EmotionalLoss*

The *a contrario* scheme can also be used in a meta-dialogical sense that concerns the choice of the scheme. A clear example is the following argument taken from *R. v. Barnet London Borough Council* (1 All ER 97, 2004):

> The words 'ordinarily residing with' are common English words and here there is no context requiring that they should be given other than their natural meaning in accordance with the accepted usage of English. Even in such circumstances, however, there can be difficulty and doubt as to their applicability to particular sets of facts, because the conception to which the words have reference does not have a clearly definable content or fixed boundaries.

The reasoning can be represented as follows, where *mAC* stands for Meta-*a contrario*.

Major Premise:	*mAC*: If *E* in *D* is an ordinary English expression, and *E* in *D* has no context requiring a technical meaning, then the *Technical Language* is inapplicable to expression *E* in a document *D*.
Minor Premise 1:	*"Ordinarily residing with"* in the *Local Education Authority Awards Regulations* is an ordinary English expression.
Minor Premise 2:	*Ordinarily residing with"* in the *Local Education Authority Awards Regulations* has no context requiring a technical meaning
Conclusion	The *Technical Language* canon is inapplicable to expression *"Ordinarily residing with"* in the *Local Education Authority Awards*

In this case, the absence of a context requiring a technical language (such as a definition, or the technical nature of the object of the regulation at issue), leads to the inapplicability of the *Technical Language* canon. This scheme is not a mere rebuttal (exclusion of a determinate

meaning), but an undercutter (an attack to the grounds of an argument, in this case the possibility of using a major premise) (Pollock, 1995, pp. 40–41; Walton, 2015, pp. 70–71). Thus, the fact that the technical language argument cannot be used to support that interpretation, does not exclude that the same interpretation can be successfully proposed through a different argument, such as the teleological one (argument from purpose).

The meta-dialogical analysis of the *a contrario* argument raises two issues concerning its nature. The first one is the relationship between the exclusion of alternative canons of interpretations and the idea of default. According to Alexy and Dreier (1991, pp. 95–98), the *Ordinary Language* scheme should be taken as the default setting. The general principle at work here is the following conditional: any expression in a legislative document should be interpreted using *Ordinary Language*, unless there are superior reasons to interpret the expression as fitting one of the other ten schemes. However, all interpretive canons are defaults. The difference here is that for any expression we can raise the defeasible claim that it should be interpreted according to its ordinary language meaning, while claims based on other canons can only be raised under specific conditions (e.g., a technical context is requires to substantiate the claim that a term should be interpreted in a technical meaning).

The second controversial issue about the *a contrario* argument is whether it ought to be treated only as an argumentation scheme or also as a meta-level principle that can be applied in conjunction with interpretive argumentation schemes. Argument from ignorance has traditionally been treated as an argumentation scheme in logic (Macagno & Walton, 2011; Walton, 2013), whereas the closed world assumption has been treated in AI as a meta-level principle rather than as a specific form of argument in its own right (Reiter, 1980). The *a contrario* argument is similar to the argument from lack of evidence as is supports an inference from a negative finding to a positive conclusion.

5. Attacking, Questioning and Defending Interpretive Arguments

Since the basic defeasible schemes share a general pattern for interpretive arguments, there is no need to formulate critical questions for each of these schemes individually. The critical questions for each of them follow the general pattern indicated by the three critical questions presented below.

(CQ_1) What alternative interpretations of *E* in *D* should be considered?
(CQ_2) What reasons are there for rejecting alternative explanations?
(CQ_3) What reasons are there for accepting alternative explanations as better than (or equally good as) the one selected?

The function of the critical questions is to help someone dealing with interpretive issues to probe into an interpretive argument in order to get an initial idea of what some of the weak points and it might be. They have a heuristic function of suggesting to an arguer who is at a loss on how to respond by suggesting possible avenues of attack. In this instance the CQs are not independent of each other, and they have an ordering. CQ_1 should be asked first.

The way we will analyze interpretive arguments, as well as critical questions matching them and counter-arguments attacking them, is to build an argumentation graph which includes a contested interpretive argument and provides an analysis of how the chains of argumentation on both sides of the dispute connect with each other and to the ultimate claim at issue. This can be done using tools from formal argumentation systems such as the Carneades Argumentation

System (CAS) or the ASPIC+ system. Both ASPIC+ and CAS are based on a logical language comprising both strict and defeasible inference rules that can be used to build arguments, and both systems use argumentation schemes. Sartor et al. (2014) have applied ASPIC+ to build a logical analysis of interpretative schemes, and we will use here a simplified version of CAS which will prove to have some tools that can be applied to examples illustrating our distinctive argumentation approach to interpretative arguments.

Both ASPIC+ and CAS use a scheme called defeasible *modus ponens*, also used in the DefLog argumentation system of Verheij (2008). This scheme is a variant of *modus ponens* in which the antecedent of the conditional premise takes the form of a conjunction. Verheij (Verheij, 2008, p. 24) observed that if you look at the typical argumentation scheme with eyes slightly narrowed, it appears to have a *modus ponens* format in outline. In the formalism that will be used in the second part of the present contribution, a scheme fits the following type of argument structure, where the major premises is a defeasible conditional with a conjunctive antecedent.

Major Premise: $A, B, C, \ldots \Rightarrow Z$
Minor Premise: $A, B, C, \ldots$
Conclusion: Z

It was shown in Walton (2004, pp. 134–139) how a majority of the schemes recognized in the argumentation literature can be tailored to fit this defeasible *modus ponens* form. In all three systems, arguments are modeled as graphs containing nodes representing propositions from the logical language and edges from nodes to nodes. In these systems an argument can be supported or attacked by other arguments, which can themselves be supported or attacked by additional arguments. The outcome in a typical case of argumentation is a graph structure representing a series of supporting arguments, attacks and counterattacks in a sequence that can be represented using an argument map, also often called an argument diagram.

CAS models arguments as directed graphs consisting of argument nodes connected to statement nodes. The premises and conclusions of an argument graph represent the edges of the graph, connecting the statement and argument nodes (Gordon, 2010). Argument nodes represent different structures of different kinds of arguments, such as linked or convergent arguments. A linked argument is one where two or more premises function together to support a conclusion. In the argument maps below the name of the argumentation scheme is inserted in the node (the circle) joining the premises to the conclusion. As will be shown in the figures, there can be two kinds of arguments shown in the node, a pro (supporting argument) or a con (attacking) argument. A supporting argument is represented by a plus sign in its argument node whereas a con argument is represented by a minus sign in the nodes that contain argumentation schemes such as *modus ponens*, argument from expert opinion and so forth (http://carneades.github.com). Conflicts between pro and con arguments can be resolved using proof standards such as including preponderance of the evidence (Gordon & Walton, 2009b). Argument graphs are evaluated, relative to audiences, modeled as a set of assumptions and an assignment of weights to argument nodes. An audience is defined as a structure <*assumptions*, *weight*>, where *assumptions* $\subseteq L$ is a consistent set of literals assumed to be acceptable by the audience and *weight* is a partial function mapping arguments to real numbers in the range 0.0...1.0. These numbers represent the relative weights assigned by the audience to the arguments (Gordon & Walton, 2011).

In CAS there can be compound arguments consisting of several argument nodes joined together by edges in the graph so that an argument represents a chain of reasoning from the supporting premises down to the ultimate proposition to be proved, the so-called statement at issue. Arguments are evaluated on the basis of whether the audience accepts the premises or not, and on how strong the various arguments making up the graph are. A very simple example of how an argument evaluation works in the CAS system is shown in Figure 2. The rounded nodes represent argumentation schemes accepted by the audience. A pro argument is indicated by the plus sign in its node. A con argument is represented by a minus sign in its argument node. A green node means the proposition in it is accepted by the audience. A red node means the proposition in it is rejected by the audience. If the node is white (no color), the proposition in it is neither accepted nor rejected. In the printed version, green appears as light gray and red appears as dark gray.

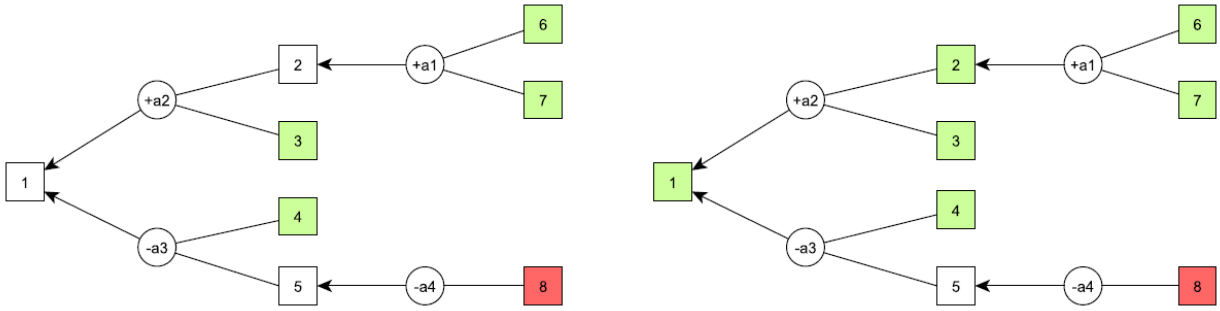

Figure 2: CAS Graphs Displaying an Argument Evaluation

In both argument diagrams shown in Figure 2, the ultimate conclusion, statement 1, is shown on the far left of the diagram. First, let's consider which premises the audience accepts or rejects, as shown in the argument diagram on the left. Argument 2 is a pro-argument supporting statement 1, while argument 3 is a con argument attacking statement 1. The audience accepts proposition 3 as a premise in argument 2, but the other premise, statement 2, is neither accepted nor rejected by the audience. Both premises of this additional argument, argument 1, are accepted by the audience. Argument a3 is a con argument but one of its premises, statement 5, is not accepted. Moreover, this premise is attacked by a con argument, but the only premise in this con argument statement 6, is rejected.

To see how this conflict is resolved, look at the diagram on the right. Since both statements 6 and 7 are accepted by the audience, CAS automatically calculates that the conclusion 2, is accepted. However, what about the con argument against statement 1 shown at the bottom, namely argument 3? This con argument could defeat statement 5, but its premise 8 is rejected by the audience. Therefore, pro argument a2 wins out over con argument a3, and so the ultimate conclusion 1 is shown in green as acceptable.

CAS also formalizes argumentation schemes. Schemes can be used to construct or reconstruct arguments, as well as to determine whether a given argument properly instantiates the types of argument deemed normatively appropriate according to the scheme requirements.

The critical questions matching an argumentation scheme cannot be modeled in a standard argument graph straightforwardly by representing each critical question as an additional implicit

premise of the scheme. The reason is that there are two different variations on what happens when a respondent asks a critical question (Walton & Gordon, 2005). These variations concern the pattern of how the burden of proof shifts from the proponent to the respondent and back as each critical question is asked by the respondent in a dialogue. With some critical questions merely asking the question is enough to defeat the proponent's argument, because the burden of proof is shifted onto the proponent's side, and if the proponent fails to meet this burden of proof, the initial argument is immediately defeated. With other critical questions, merely asking the critical question is not enough by itself to defeat the proponent's argument. For example, if the respondent asks the bias critical question when the proponent has put forward an argument from expert opinion, the proponent can simply reply, "What proof do you have that my expert is biased?" On this approach, merely asking the question does not defeat the proponent's argument until the respondent offers some evidence to back it up. CAS deals with this problem of burden of proof for critical questioning by distinguishing three types of premises in an argumentation scheme, called ordinary premises, assumptions and exceptions. Assumptions are assumed to be acceptable unless called into question. Exceptions are modeled as premises that are not assumed to be acceptable and which can block or undercut an argument as it proceeds. Hence an exception, which is modeled in CAS as an undercutter, only defeats the argument it was attacking if it is supported by other arguments which offer reasons to back up the undercutting argument. Ordinary premises of an argumentation scheme are treated as assumptions. They are assumed to be acceptable in case they are put forward, but must be supported by further arguments to remain acceptable after being challenged by critical questions.

For any one of these critical questions to be effective in defeating the original interpretive argument, the respondent must give some indication of what he takes this alternative interpretation to be. Thus it would appear that each of these critical questions only defeats the original interpretive argument if some evidence is presented by the respondent pinpointing an alternative interpretation which might challenge the one originally appealed to by the proponent's argument.

Like ASPIC+, CAS has three ways in which one argument can attack and defeat another. An opponent can attack one or more of the premises of an argument. This is called an undermining attack. Or an opponent can attack the conclusion by presenting an argument to show it is false or unacceptable. This type of attack is called a rebutter. But thirdly, the opponent can attack the inferential link joining the premises to the conclusion. This type of attack is called an undercutter. For example, if the inference is based on a rule, the attack could claim that there is an exception to the rule that applies in the present case at issue. This way of modelling argumentation is based on Pollock's distinction (Pollock, 1995, p. 40) between two kinds of argument attacks called rebutters and undercutters. On Pollock's view, a rebutter is a counter-argument that attacks the conclusion of a prior argument whereas an undercutter is a counter-argument that attacks the argument link between the premises and the conclusion. For example, an argument that fits the argumentation scheme for argument from expert opinion can be critically questioned by asking whether the expert is biased. In CAS such a critical question is modeled as an undercutter, and an undercutter is modeled as an argument that defeats the original argument it was aimed at only if it is backed up by some additional evidence that supports it.

Next, we use CAS to show how the interpretative statutory schemes can be applied to an extended sequence of argumentation in a typical case using large argument graphs to connect the individual interpretive arguments to each other.

6. The Education Grants Example

According to the account of the following case described in Cross (2005, p. 90), Section 1 of the Education Act of 1962 required local education authorities to make grants to students who were 'ordinarily resident' in their area, so that the student could attend higher education courses. A requirement in the Education Act stipulated that in order to be eligible, the student had to have been ordinarily resident in the UK for three years prior to application. The following issue arose: could someone who had come to the UK for education count the period spent in education as ordinary residence in order to qualify for a mandatory grant under the Education Act?

There were two sides to the issue. The Court of Appeal held that such a person could not count this period as ordinary residence, offering the following argument quoted from Cross (2005, p. 90). Lord Denning MR and Everleigh LJ (see the quotation above Fig. 4) related this Act to the policy of the Commonwealth Immigrants Act 1962 and its successor, the Immigration Act 1971. According to the latter Act, students coming only for study had a conditional leave to stay in the country limited to the purpose of study and this conditional leave did not involve ordinary residence for the general purposes of everyday life. They held that consistency with this Act required the term 'ordinarily resident' in the Education Act to be interpreted as living as an ordinary member of the community would, which could not include residence for the limited purpose of study.

Arriving at a different interpretation, the House of Lords unanimously reversed this decision. They felt that the Court of Appeal had given too much weight to arguments drawn from the Immigration Act. They offered the following argument, quoted from Cross (2005, p. 91).

> Parliament's purpose expressed in the Education Act gave no hint of any restriction on the eligibility for a mandatory award other than ordinary residence in the United Kingdom for three years and a satisfactory educational record. There was nothing expressed in the Immigration Act which gave guidance as to the interpretation of the Education Act and, indeed, despite a series of immigration measures since 1962, nationality had not formed part of the regulations under the Education Act until 1980. Accordingly, the ordinary natural meaning of the Education Act prevailed to make the students eligible for a mandatory grant if they had resided in the United Kingdom for the purposes of study.

In this case it was concluded that the role of the judge should not be to reconcile legislative provisions. Instead, it was proposed that the basis for interpretation should be that of the ordinary language meaning of the expression 'ordinarily resident'.

The argumentation in this case can be analyzed as an interpretive argument put forward by its proponents Denning and Everleigh, countered by an interpretive argument put forward in the House of Lords. Below we use a sequence of three argument maps to model the structure of the argumentation in the case.

The first argument, shown in Figure 3, cites the Immigration Act of 1971, which stated that students coming to a country for study only had a conditional leave to stay in the country, adding that this conditional leave does not involve ordinary residence for the general purposes of everyday life. Because a related document is cited as the basis for drawing a conclusion in support of statutory interpretation, the argumentation scheme which is the basis of this argument is the one for argument from contextual harmonization (CH), recognized by MacCormick and Summers. For present purposes, this scheme is taken to represent the following kind of argument: a certain expression that occurs in a document is best interpreted as fitting with its usage in a set of related documents, therefore in this document it will interpreted in the same

way. In other words, if there is an issue about how to interpret an expression in a given document, such as a statute, then it can be argued that the best way to interpret it is within a context of related documents so that it fits with the way the term has been interpreted in these other documents.

Let's apply the scheme for the argument from contextual harmonization to the first part of this example. The notation *+CH*, referring to a supporting use of argument from contextual harmonization, has been inserted in the node linking the two premises in the middle of Figure 2 to the ultimate conclusion shown at the left. Here is a textual representation of the arguments, which corresponds to the graph of Figure 3. Let us first examine the top argument by Lord Denning.

Major Premise:	*eCH*: If the interpretation of *E* in *D* as excluding *C* fits the context, then *E* in *D* should be interpreted as excluding *C*.
Minor Premise:	The interpretation of "*residence"* in the *Education Act* as excluding *ResidenceForTheLimitedPurposeOfStudy* fits the context.
Conclusion	"*residence*" in *Education Act* should be interpreted as excluding *ResidenceForTheLimitedPurposeOfStudy*.

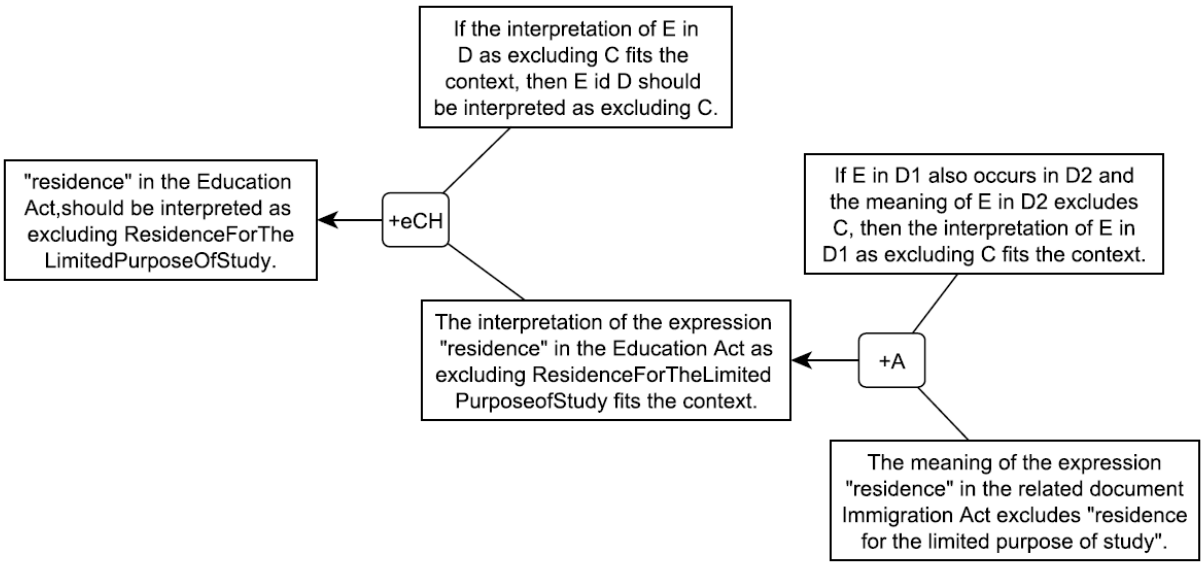

Figure 3: Proponent's Argument in the Educational Grants Example

The supporting argument may appeal to the fact that in other pieces of legislation "*ordinary residence*" excludes indeed "*residence for the limited purpose of study"*.

Major Premise:	*eCH*: If an expression *E* in document D_1 also occurs in a related document D_2, and the meaning of *E* in D_1 excludes a concept *C*, then the interpretation of the expression *E* in D_2 as excluding *C* fits the context.
Minor Premise:	The meaning of "*residence"* in the related document *Immigration Act*

	excludes concept "*residence for the limited purpose of study"*
Conclusion	The interpretation of an expression "*residence"* in the *Education Act* as excluding *ResidenceForTheLimitedPurposeOfStudy* fits the context.

The ultimate conclusion is the statement that non-UK students cannot count the period as ordinary residence.

Next we turn to an analysis of the argumentation in the second quoted text above, where the opponent, in this instance the House of Lords, put forward a counterargument.

> Parliament's purpose expressed in the Education Act gave no hint of any restriction on the eligibility for a mandatory award other than ordinary residence in the United Kingdom for three years and a satisfactory educational record.

This argument fits the scheme for inclusionary argument from intention (+*iAI*):

Major Premise:	+*iAI*: If the interpretation of *E* in *D* as excluding *S* conflicts with legislative purpose, then *E* in *D* should be interpreted as including *S*.
Minor Premise:	The interpretation of an expression "*residence"* in the *Education Act* as excluding *ResidenceForTheLimitedPurposeOfStudy* conflicts with legislative purpose.
Conclusion	"*residence*" in *Education Act* should be interpreted as including *ResidenceForTheLimitedPurposeOfStudy*.

The reason why the minor premise holds is provided by the following supporting counterfactual argument.

Major Premise:	If the linguistic meaning of *E* in *D* includes *S*, and there are no hints that the legislator intended to exclude *S* from the meaning of *E* in *D* , then the interpretation of *E* in *D* as excluding *S* conflicts with legislative intention.
Minor Premise 1	The linguistic meaning of "*residence*" in the *Education Act* includes *ResidenceForTheLimitedPurposeOfStudy*.
Minor Premise 2	There are no hints the legislator intended to exclude *ResidenceForTheLimitedPurposeOfStudy* from the meaning of "*residence*" in *Education Act.*
Conclusion	The interpretation of an expression "*residence"* in the *Education Act* as excluding *ResidenceForTheLimitedPurposeOfStudy* conflicts with legislative intention.

This argument is shown in Figure 4 as a counterargument to the one in figure 3.

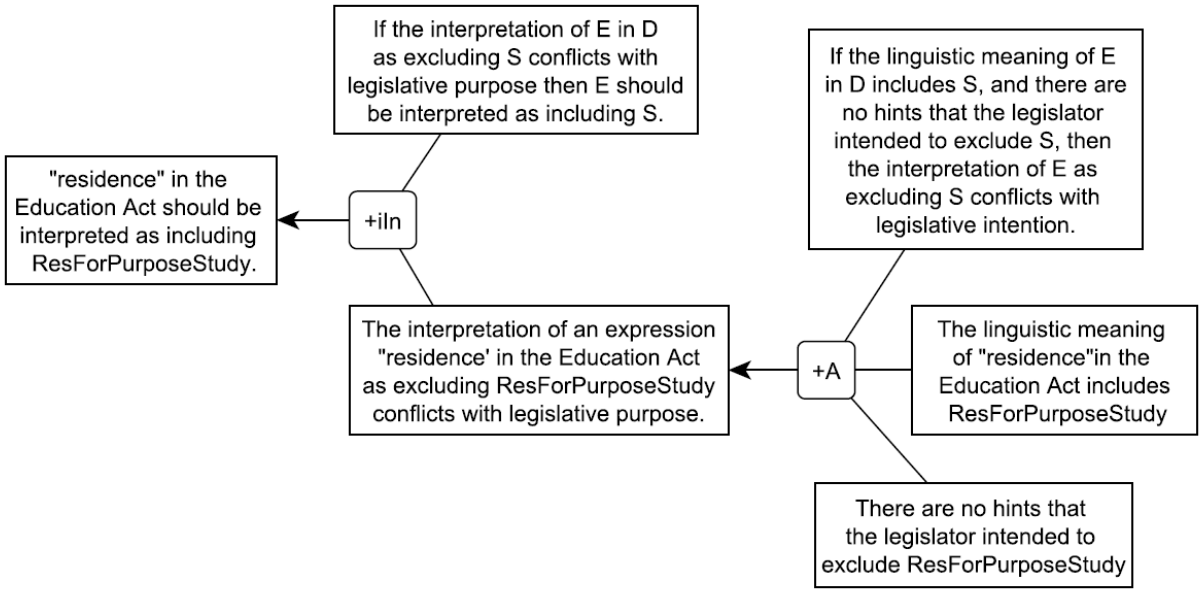

Figure 4: Respondent's Rebuttal to the Educational Grants Example

We leave it as an open problem how the argument on the right could be more fully represented, for example by including the "there are no hints" statement as a premise in an *a contrario* argument. This would make the argument on the right more complex. Hint: it is possible to solve this problem by invoking the notion of an enthymeme.

Next let's look at the other argument just below this one. Cross (2005, 91-92) offers this account of this part of the case.

Lord Denning MR and Everleigh LJ were impressed by the need to relate this Act to the policy of the Commonwealth Immigrants Act 1962 and its successor, the Immigration Act 1971. Under the latter act, students coming only for study had a conditional leave to stay in the country limited to the purpose of study and this did not involve ordinary residence for the general purposes of everyday life. They considered that consistency with this Act requires the term 'ordinarily resident' in the Education Act to be interpreted as living as an ordinary member of the community would, which could not include residence for the limited purpose of study.

We are told in the quoted part of the text that Denning and Everleigh considered that consistency with the Education Act requires living as an ordinary member of the community, and that being an ordinary member of the community does not include residence for the limited purpose of study. Accordingly, we have represented these two propositions as premises in a linked argument supporting the conclusion that conditional leave does not involve ordinary residence, as shown in Figure 5 at the bottom right. The rightmost argument supports one premise of the argument to the left of it. It is labelled as a supporting argument labeled *+iPr* in figure 5. The conclusion of this argument is the opposite of the conclusion shown in figure 4.

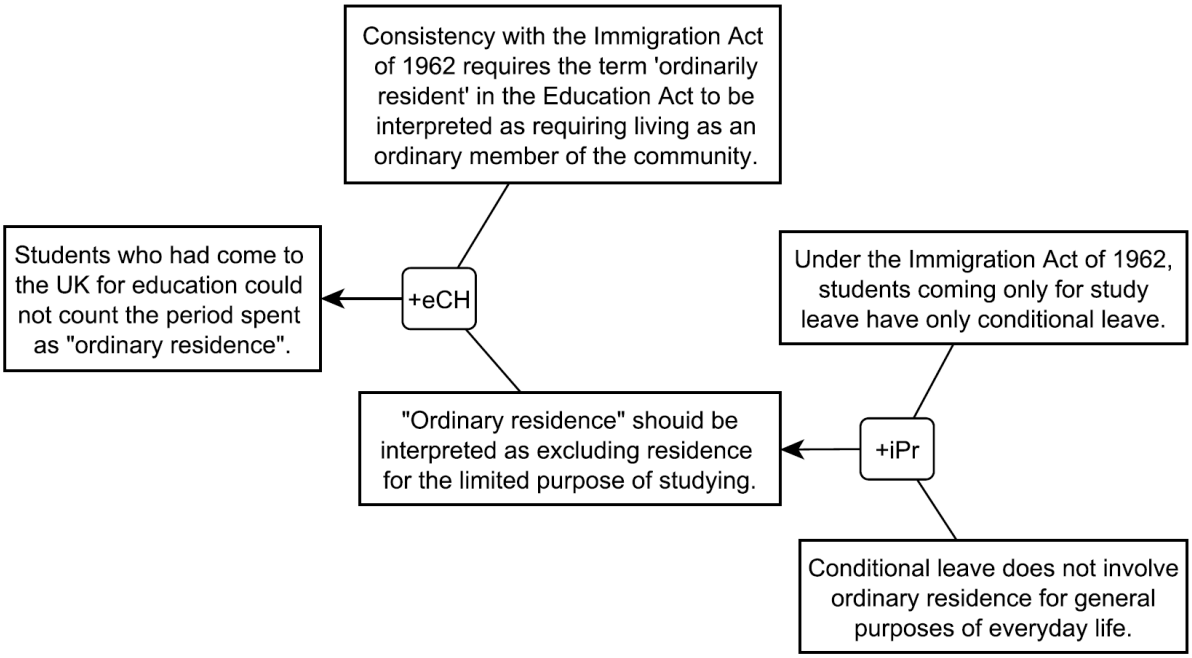

Figure 5: Respondent's Premise Attack in the Educational Grants Example

What we see in Figure 5 therefore a rebuttal because it presents an argument that attacks the ultimate conclusion of the original argument shown in Figure 4. There is a conflict between the argument shown in Figure 5 and the previous two arguments shown in Figures 3 and 4.

We have chosen to use the term 'interpretation' instead of 'meaning', because the latter term is not only vague but is itself susceptible to many contested interpretations. Nevertheless it can be said generally that what the interpreters of the statue are generally seeking is an interpretation that they contend that represents the genuine, true or real meaning of the textual item they are discussing. This notion that there is what is called a real meaning underneath the vagaries in the text being examined or deconstructed has however been subject to some abuse in philosophy. For all these reasons we generally prefer using the term 'interpretation' to the term 'meaning'.

The evaluation system of CAS compares the set of pro arguments against the set of con arguments if the two sets of arguments are independent of each other. However, summing the weights of arguments to check if the sum of the weights of the pro arguments outweighs the sum of the weights of the con arguments is only feasible if it be assumed that the two arguments are independent of each other. This can be done with CAS but it requires an additional evaluation.

As with all arguments found in natural language texts, it is possible to analyze the given text in further depth by bringing out more implicit assumptions and more subtle inferences. However, building an argument map of a real argument expressed in natural language is very often a difficult interpretive task requiring learned skills, and often itself providing many challenges of textual interpretation. Generally, one finds there are alternative interpretations opened up as the text of the cases is analyzed in greater depth and more implicit premises and arguments are brought out. Building an argument diagram can often raise important questions of argument interpretation and analysis that might not be initially visible to someone who is trying

to deal with the argument or find out what to do with it. To illustrate some of the problems inherent in such as task we go back to the *Dunnachie* example.

7. Fitting Interpretive Schemes to Cases

Dunnachie, following the commentary of MacCormick (MacCormick, 2005, p. 128), offers an example of argument from contextual harmonization. The scheme for argument from contextual harmonization requires that a particular sentence in a statute should be interpreted in light of the whole statute and any set of related statutes that are available. In line with the model of interpretive schemes introduced in section 2, the scheme for contextual harmonization as applied to *Dunnachie* takes the following form.

Major Premise:	+*CH*: If the interpretation of *E* in *D* as *M* fits the context, then *E* in *D* should be interpreted as *M*.
Minor Premise:	The interpretation of "*loss"* in the *Employment Relations Act* as *PecuniaryLoss* fits the context.
Conclusion	"*loss*" in *Education Act* should be interpreted as *PecuniaryLoss.*

The reason why this interpretation fits context is provided by the following supporting argument, which addresses the case in which the same expression occurs in different positions in the document (for simplicity's sake we do not include in the scheme the possibility that there are multiple occurrences of the expression in the same document):

Major Premise:	If *E* besides occurring in position P_1 of document *D* also occurs in positions $P_1, \ldots, P_n$, where it has meaning *M*, then *E* in P_1 should also be interpreted as *M*.
Minor Premise:	"*loss*" besides occurring in *Section 2* of the *Employment Relations Act,* also occurs in *Section 4* where it has the meaning "*pecuniary loss*".
Conclusion	"*loss*" in *Section 2* of the *Employment Relations Act* should be interpreted as "*pecuniary loss*".

Again following the commentary of MacCormick (MacCormick, 2005, p. 128) on *Dunnachie*, the following example can be given to show how CAS models a pro argument supporting a claim in a case supports one premise of another pro argument supporting still another claim, as shown in Fig. 6.

The claim that "loss" should be interpreted as including both financial and emotional loss was partly based on a statement made in an earlier case. In this case, *Johnson Unisys Ltd*., Lord Hoffman had made the statement that an extension of the word 'loss' to 'emotional loss' could be made. So it would appear, at least initially, that the argument drawn from the statement can be classified as an instance of a pro-argument from precedent.

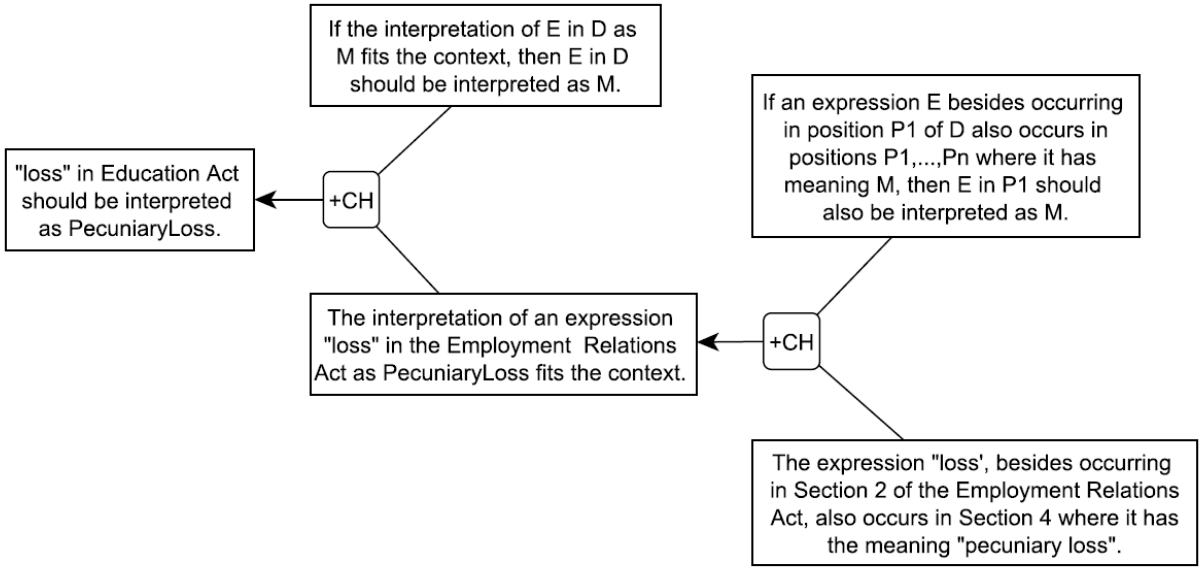

Figure 6: The Use of the Scheme for Argument from Contextual Harmonization in *Dunnachie*

The reader will recall from the list in section 2 that according to the description given by MacCormick and Summers, (1991) an interpretive argument from precedent requires that if a term has a previous judicial interpretation, it should be interpreted to fit that previous interpretation. In the previous case of *Norton Tool Co. v Tewson*, it had been ruled that "loss" was to be interpreted as signifying exclusively financial loss. Following the lines of the analysis of the structure of interpretative schemes in section, the scheme for interpretive argument from precedent can be cast in the following inclusionary and exclusionary forms.

Major Premise:	*ePr*: If the interpretation of *E* in *D* as excluding *S* fits precedents, then *E* in *D* should be interpreted as excluding *S*.
Minor Premise:	The interpretation of an "*loss"* in the *Employment Relations Act* as excluding *EmotionalDamage* fits precedents.
Conclusion	"*loss*" in *Education Act* should be interpreted as excluding *EmotionalDamage*.

The supporting argument is the following.

Major Premise:	If *E* in *D* was understood in precedent *P* as excluding *C*, then the interpretation of *E* in *D* as excluding *C* fits precedents.
Minor Premise:	"*loss"* in the *Employment Relations Act* was understood in Norton as excluding *EmotionalDamage*.
Conclusion	The interpretation of "*loss"* in the *Employment Relations Act* as excluding *EmotionalDamage* fits precedents.

Here is a positive application of the argument by precedent.

Major Premise:	*iPr*: If the interpretation of *E* in *D* as including *C* fits precedents, then *E* in *D* should be interpreted as *M*.
Minor Premise:	The interpretation of "*loss"* in the *Employment Relations Act* as including *EmotionalDamage* fits precedents.
Conclusion	"*loss*" in *Education Act* should be interpreted as including *EmotionalDamage.*

A supporting argument is the following:

Major Premise:	If *E* in *D* was understood in precedent *P* as including *C*, then the interpretation of *E* in *D* as including *C* fits precedents.
Minor Premise:	The interpretation of an expression "*loss"* in the *Employment Relations Act* was understood in precedent *Johnson vs Unisys* as including *EmotionalDamage*.
Conclusion	The interpretation of an expression "*loss"* in the *Employment Relations Act* as including *EmotionalDamage* fits precedents.

The arguments could be further developed by pointing to the clues which support this understanding of the precedent, using the argument diagram in Figure 7.

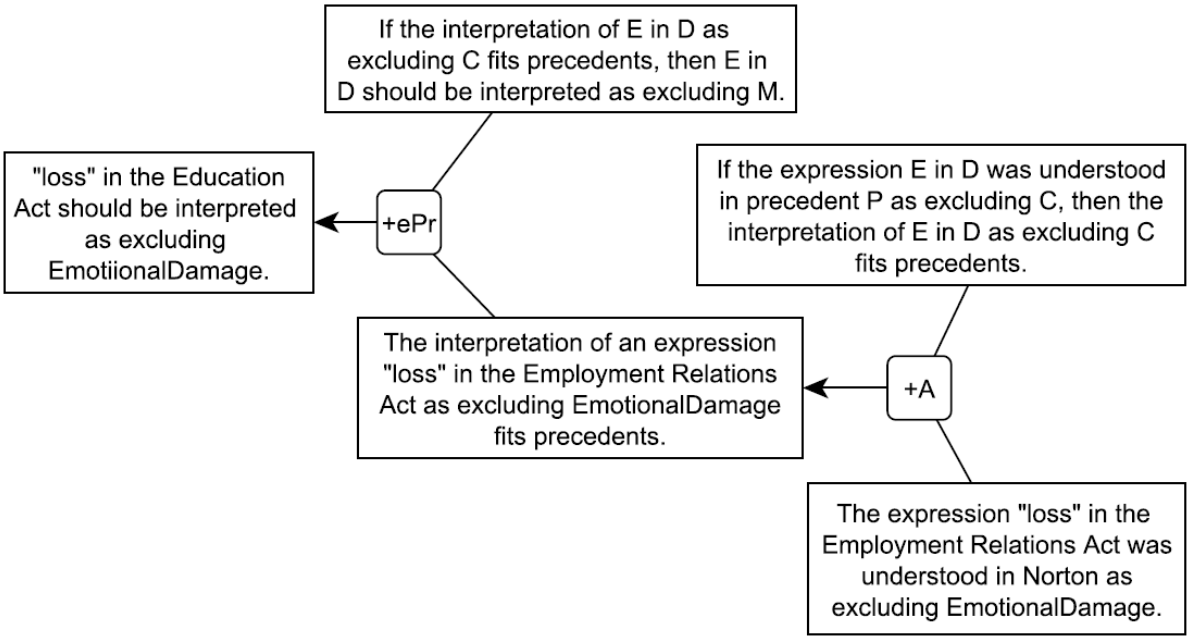

Figure 7: Use of a Prior Case as a Precedent Supporting a Textual Interpretation

But in *Dunnachie*, in addition to this pro instance of interpretive argument from precedent, there was also a con argument for the same conclusion. There is a conflict between the two interpretations shown in Figure 8.

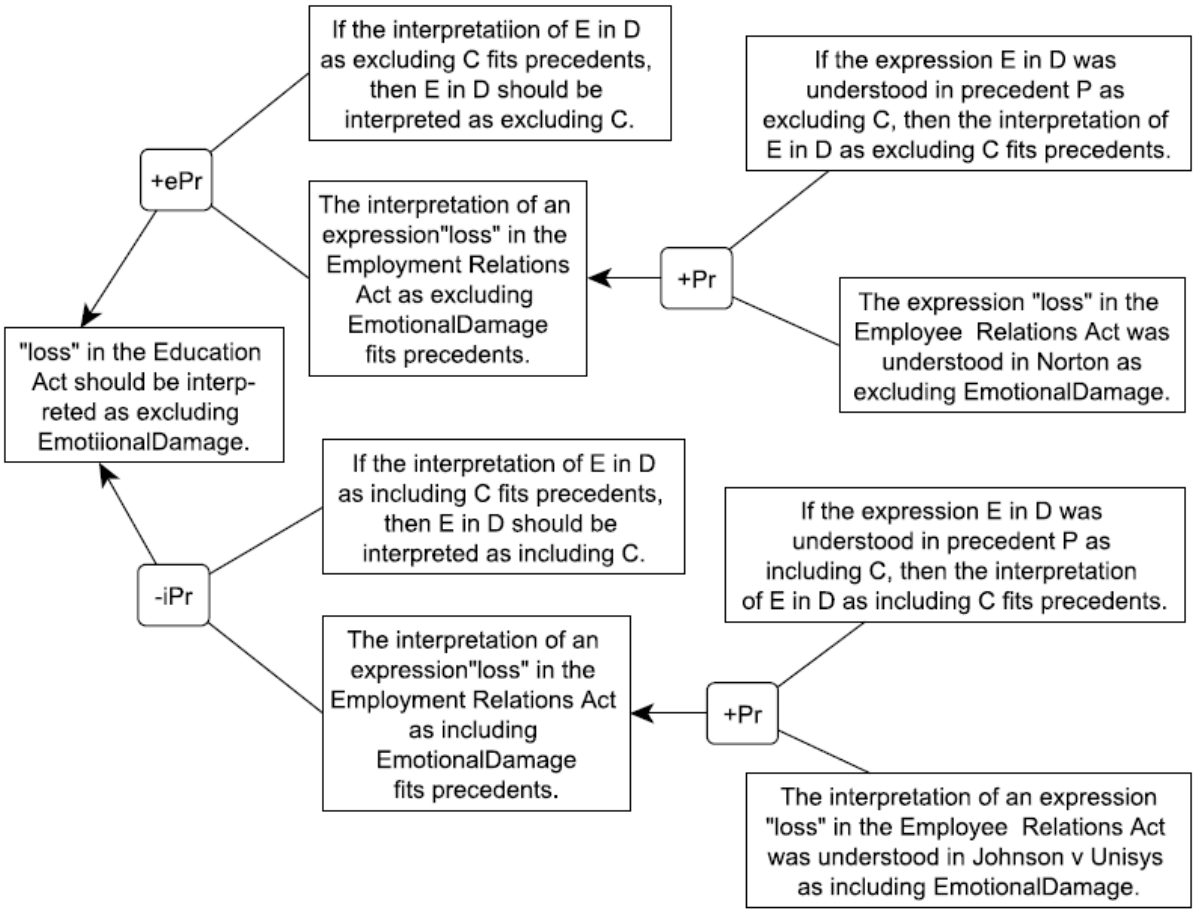

Figure 8: Conflicting Pro and Con Interpretive Arguments from Precedent

How could this conflict be resolved? The answer requires taking a closer look at the interpretive scheme for argument from precedent to see how one precedent can be stronger than another in supporting or attacking a claim about how a statute or law should be interpreted.

This way of modeling the scheme rests on the assumption that the user already has a clear idea of what a precedent is. Schauer (Schauer, 1987) has shown that arguments from precedent are already highly familiar in everyday conversational argumentation. This suggests that we need to begin with some intuitive understanding of what constitutes a precedent case. We could also build on the scheme for argument from precedent generally known in the argumentation literature, but there are differences of opinion on how that should be formulated (Walton, 2010), in particular on the issue of how that scheme is related to the one for argument from analogy.

In his commentary on the case, MacCormick (2005, p. 129) made the following argument to support seeing this statement by another court as a binding premise in an argument from precedent. First, this ruling had been followed and approved many times. Second, it contained an acceptable rationale for interpreting loss exclusively as financial loss. Therefore, MacCormick concluded that it was a better guide for future rulings than the *Johnson* case.

In contrast, MacCormick put forward arguments advancing several reasons why Lord Hoffman's statement in *Johnson* might not constitute a binding precedent. First, they were not

necessary to the decision reached in *Johnson*. Second, it had not been followed by other courts as a binding precedent. Third, although it was open to the House of Lords to have overruled *Norton Tool*, establishing a new ruling on the meaning of loss, this was not done. These arguments were used by MacCormick to question whether the remarks made by Lord Hoffman constitute a precedent binding on subsequent cases. These further arguments are shown in Figure 9. For simplicity and readability's sake we do not follow rigidly the structures illustrated above, and we omit to fully indicate the canons that are applied.

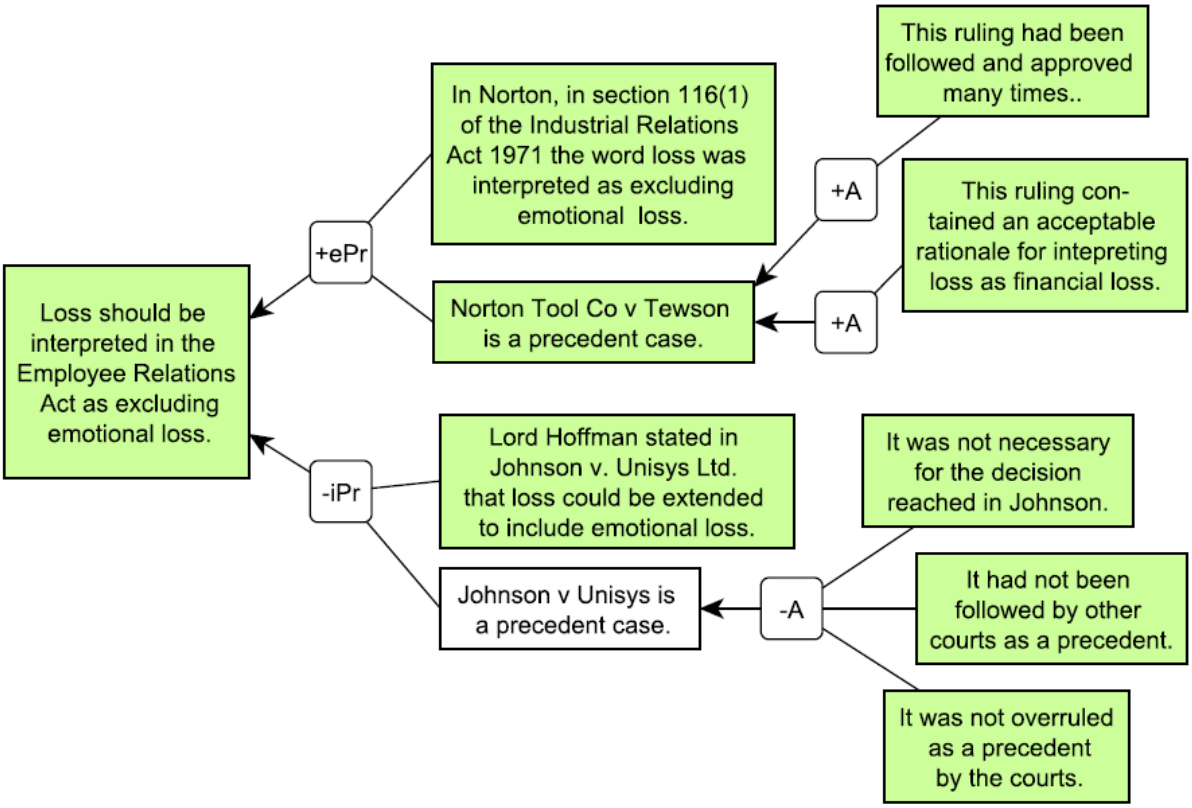

Figure 9: Conflict Resolved by Taking Other Arguments into Account

Let's say that all the propositions shown in the five rightmost rectangles are accepted by the audience. These five rectangles are shown in green backgrounds. Next, look at the pro argument from precedent at the top. Each of the two arguments supporting the proposition that *Norton Tool Co. v Tewson* is a precedent case has only one premise, and in both instances that premise is accepted. Therefore the proposition that *Norton tool Co. v Tewson* is a precedent case is automatically shown as accepted by CAS. Let's also assume that the other premise of this argument is accepted. Since both premises of the argument are now accepted the ultimate conclusion shown at the left of Figure 9 is now automatically shown as accepted.

But now let's look at the bottom argument, the con argument from precedent. Since all three of its premises are accepted, the con argument attacking the proposition that *Johnson v Unisys* is a precedent case is successful in defeating it. Hence this proposition is shown in a rectangle with a white background, indicating that it is not accepted. Actually, the additional evidence provided

by the two Pro arguments shown at the top right of Figure 9 are not needed for the pro-argument from precedent to defeat the con argument from precedent in the case. It is enough that because one premise of the con argument (shown in white at the bottom of Figure 9) is defeated, the pro argument from precedent at the top prevails.

Summing everything up, the pro argument from precedent at the top prevails over the con argument from precedent at the bottom, because one of the premises of the con argument is unacceptable. It is shown by CAS as not accepted because it is defeated by the applicable con argument -A. Only the pro argument is accepted, and so the conclusion is accepted. Hence the conflict is resolved.

There is another way of modeling the conflict between the two arguments from precedent.

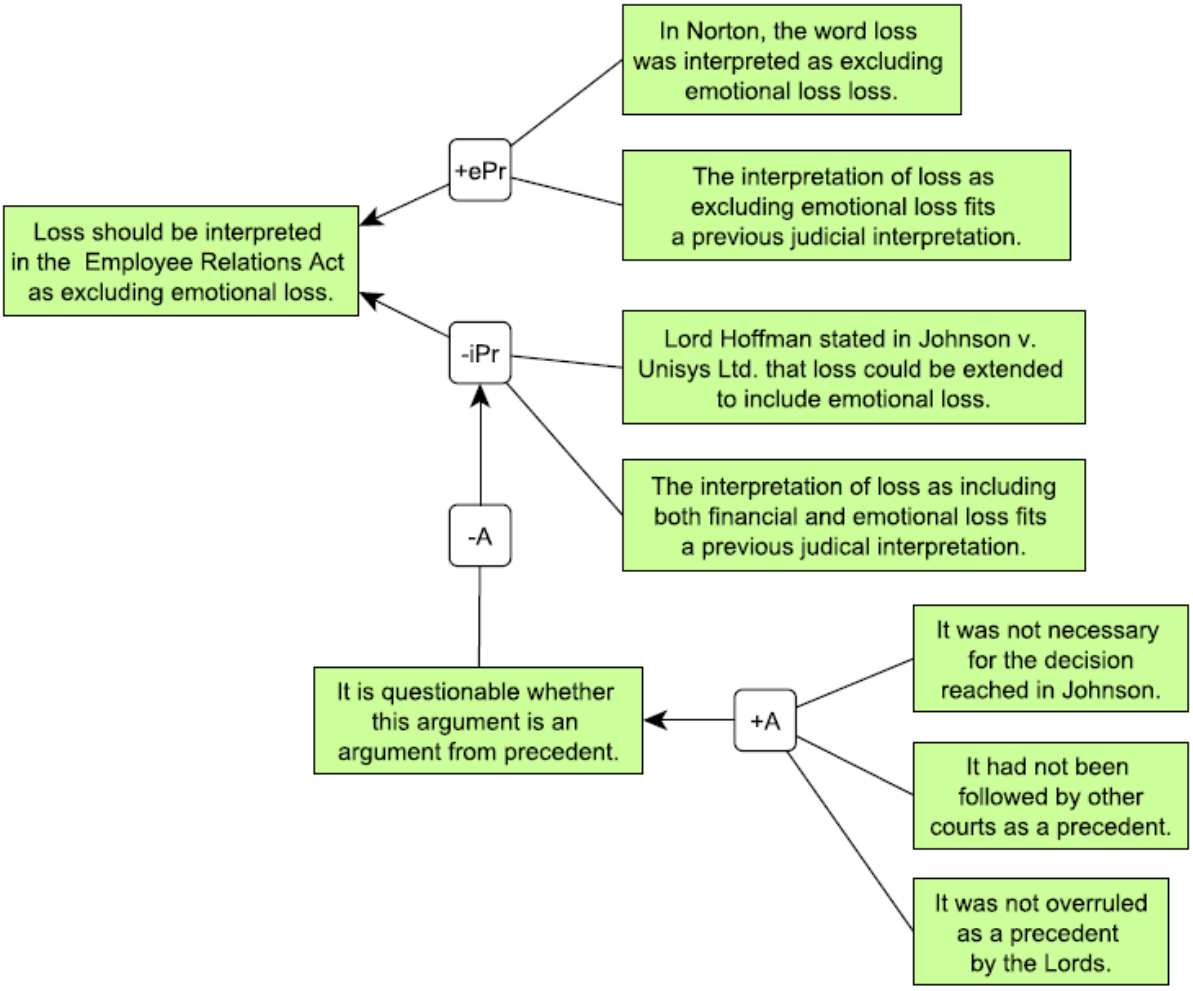

Figure 10: Attacking an Interpretive Argument from Precedent

Using the scheme for argument from precedent put forward in section 2, MacCormick's argument could be modeled as an undercutter critically questioning whether the top argument shown in Figure 10 fits the argumentation scheme for argument from precedent. This way of interpreting MacCormick's remarks on how to model the argumentation in this instance is to take

his argument above as an undercutter that attacks the argument used in the *Johnson* case by arguing that it is questionable whether the pro-argument shown in Figure 10 is a proper instantiation of the scheme for argument from precedent. Such an interpretation of MacCormick's evaluation of the argumentation is shown in Figure 10.

This case is an interesting one because the way MacCormick analyzes the argumentation in it, because there is still another alternative interpretation of it that is possible, judging from his remarks. It might be possible to argue that even though the ruling in Johnson on how to interpret loss was not a binding precedent, because it was not necessary to the decision made in that case, still it could be taken to be a weaker kind of precedent. MacCormick (MacCormick, 2005, p. 129) distinguishes between a binding precedent and a precedent that is persuasive but not binding. Honoring this distinction, interpretation of the word "loss" in Johnson could be taken as a weaker kind of precedent. Following this line of argument, the conflict between the two arguments from precedent no longer represents a deadlock because the stronger precedent from Norton would have priority over the weaker precedent from Johnson. CAS and ASPIC+, as well as other systems**,** recognize different kinds of priority orderings on rules, and so that would be another way that AI systems could model the argumentation in this case.

In section 2 we only proposed schemes for some of the interpretive arguments to give the reader an idea of what these schemes should ultimately look like. However, especially with some of the schemes, the descriptions of the different kinds of interpretive arguments given by MacCormick and Summers are not enough in themselves to definitively formulate the matching scheme. In particular, the scheme for argument from precedent needs more study by applying it to cases before a definitive version can be given.

8. Formalizing Interpretive Arguments – General Structure

In this section we shall provide a general formal structure for interpretive arguments, on the basis of the approach of interpretive arguments introduced and exemplified in the previous sections. Let us first summarize that approach.

Interpretive arguments can be distinguished along two different criteria: positive vs negative and total vs partial. The first distinction concerns whether they argue that a certain interpretation should be adopted or rather rejected. The second distinction pertain as to whether they address the whole interpretation of a term, or only the inclusion or exclusion of a subclass in the term's meaning. Correspondingly, partial interpretive arguments can be distinguished into exclusionary and inclusionary ones.

All interpretive arguments we shall consider are based on canons, namely, defeasible conditionals stating that, if a certain conditions are or are not met, a certain interpretive condition should or should not be adopted. Canons may be positive or negative dispending on whether their consequent is the obligation to adopt or not to adopt a certain interpretation. Positive canons can also have a negative counterpart, to the extent that the absence of the condition they require leads to the rejection of an interpretation.

In this section we shall propose appropriate formal structures for capturing all of these forms of interpretive arguments.

Let us start with positive and negative total interpretive arguments. Both structures have the following elements: an expression *E* (word, phrase, sentence, etc.) occurs in a document *D* (statute, regulation, contract, etc.), interpreting this occurrence as meaning *M* satisfies the condition of a certain interpretive scheme (of ordinary language, technical language, purpose, etc.). Positive canons state that if all these elements are satisfied we are licensed to derive the

interpretive conclusion that *E in D should be interpreted as M.* Negative canons state that if an interpretation *I* would not fit the scheme, then *E in D should not be interpreted as M.* In (Sartor et al., 2014) we modelled interpretive claims as deontic claims, stating the obligation to adopt a certain interpretation. Here we follow a different approach, focusing on the relationship between an interpretation and its justification, as a meta-linguistic discourse on why a meaning is the best interpretation of an expression. In this sense, we model interpretive claims as terminological assertions concerning best interpretations of the contested or potentially contested expressions within a legal text (for a similar idea, see Araszkiewicz, 2013).

All canons are modelled as defeasible rules expressed in the form $r: \varphi_1, \dots, \varphi_n \Rightarrow \psi$, where r is the rule name, where $\varphi_1, \dots, \varphi_1$ and ψ are formulas in a logical language, $\varphi_1, \dots, \varphi_1$ being the *antecedents* and ψ the *consequent* of the rule.

We express interpretive conclusions as claims concerning conceptual relations between a meaning *M* that is proposed and the outcome of the best legal interpretation of the linguistic occurrence at issue, namely, expression *E* in document *D* (Bezuidenhout, 1997; Carston, 2002, 2013; Soames, 2008; Sperber & Wilson, 1986; Wilson & Sperber, 2004). Such an outcome is denoted by the function expression $BestInt(E, D)$. Conceptual relations are expressed with description logic symbols: $\equiv$ for conceptual equivalence, $\not\equiv$ for difference, $\sqsupseteq$ for inclusion.

Thus a general pattern for positive-total interpretive canon can be expressed as follows:

C: expression E occurs in document D,
the interpretation of E in *D* as *M satisfies the condition of positive canon C* ⇒
$BestInt(E, D) \equiv M$

Here is an example:

OL: expression E occurs in document D,
the interpretation of E in D as *M* fits ordinary language ⇒
$BestInt(E, D) \equiv M$

Similarly, negative canons claim that the best interpretation is not the proposed one, as in the following example, based on the non-redundancy canon:

NR: expression E occurs in document D,
the interpretation of *E* in D as *M* is redundant ⇒
$BestInt(E, D) \not\equiv M$

Let us now provide examples for partial interpretations, such as, for exclusionary interpretative claims:

eSAC: expression *E* occurs in document *D*,
the interpretation of expression *E* in the *D* as including *S* conflicts with usual meaning ⇒
$BestInt(E, D)^C \sqsupseteq S$

where $BestInt(E, D)^C$ is the complement of $BestInt(E, D)$, and for inclusionary interpretive claims:

iSAC: expression *E* occurs in document *D*,
the interpretation of *E* in the *D* as excluding *S* conflicts with the usual meaning ⇒
$BestInt(E, D) \sqsupseteq S$

We can also identify a pattern for priority arguments between different (instances of) interpretive canons (we use ≻ to express priority).

C: concerning expression *E* in document *D*, the interpretation as M_1 according to canon C_1 *meets priority criterion with regard to* the interpretation as M_2 according to canon C_2 ⇒
$C_1\ (E, D, M) \succ C_2(E, D, M_2\ )$.

Consider, for instance, Alexy and Drier's idea that in criminal law ordinary language has priority over technical language.

P_1: expression *E* in document *D* concerns Criminal law ⇒
$OL(E, D, M_1) \succ TL(E, D, M_2)$.

In this sense, interpretive arguments can be ordered in hierarchies depending on the specific legal context.

For reasoning about interpretation we need an argumentation system including strict rules, defeasible rules, and preference between rules, such as the system developed by Prakken and Sartor (Prakken & Sartor, 1996), the ASPIC+ system (Prakken, 2010) or the Carneades system (Gordon & Walton, 2009a). We express defeasible rules in the form $r: \varphi_1, \dots, \varphi_n \Rightarrow \psi$, and strict rules in the form $\varphi_1, \dots, \varphi_n \longmapsto \psi$. We use arrows → and ↔ for material conditional and biconditional of propositional logic. We also assume that our system includes the inferences of classical logic, namely, that for any propositions of classical logic φ and ψ, if φ is derivable from ψ, then we have a strict rule $\varphi \longmapsto \psi$.

Here we assume that argument A including defeasible rules may be defeated in two ways. This first consists in successfully *rebutting* A , i.e., by contradicting the conclusion of a sub-argument of A, though an an argument that is not weaker that the attacked subarguments (we assume that A too is a sub-argument of itself). More precisely, B rebuts A when (a) B's conclusion is incompatible with the conclusion of a subargument A' of A, and (b) B is not weaker than A', i.e., $A' \not\succ B$ (see Prakken, 2010). Condition (b) corresponds to the idea that if A were stronger than B, it would resist to B's challenge.

With regard to comparative strength, we assume that the comparison between two arguments A and B is to be assessed according to two criteria:

(a) preference for strict arguments (those only contains strict rules) over defeasible ones (those also containing defeasible rules): if A is strict and B is defeasible then $A>B$.
(b) preference between defeasible arguments according to the last link principle: if A is preferable to B according to the last link principle, then $A>B$.

The *last link principle* assumes a partial strict ordering ≻ over defeasible rules and compares arguments A and B having incompatible conclusions by considering the sets of the last defeasible

rules which support such conclusions in the two arguments (see for a formal characterization, Prakken & Sartor, 1996; Prakken, 2010).

The second way of defeating an argument A consists in *undercutting* A, i.e., in producing an argument B concluding for the inapplicability of a defeasible rule in A, this being the top rule of a subarguments A' of A. Let us express the applicability of rule through a special predicate $appl$, so that an argument for the inapplicability of a rule r has the conclusion $\neg app(r)$. Then we can say in general terms that argument B undercuts argument A, if B has the conclusion $\neg app(r)$, where r is the top rule of a sub-argument A' of A. For instance argument $[\rightarrow a;\ r_1: a\ \Rightarrow b]$ is undercut by argument $[\rightarrow c;\ r_2: c\ \Rightarrow \neg appl(r_1)]$.

A semantics for an argumentation system can be constructed on the basis of the idea of an extension, namely, a set of compatible arguments, which includes resources (arguments) that respond to all defeaters of arguments in the set. Here we adopt the approach that consists in looking for most inclusive extensions, which are called preferred extensions (Dung 1995). An argument is then considered to be justified if is included in all such extensions. It is considered defensible if it is included in some (but not necessarily in all) extensions. The arguments that are defensible but not justified are only in some preferred extensions: their status remains undecided, as their inclusion in a preferred extension depends on what other arguments are already included in the extension, different choices being possible.

Consider for instance the following set of arguments:$\{[a], [b], [a, r_1:\ a \Rightarrow c],\ [b, r_2: b \Rightarrow \neg c]\}$. We have two preferred extensions $E_1 = \{[a], [b], [a, r_1:\ a \Rightarrow c]\}$ and $E_2 = \{[a]\{[a], [b], [b, r_2: b \Rightarrow \neg c]\}$. Each extension includes an argument that is defeated, but also defeats an argument in the other extension: $A_1 = [a, a \Rightarrow c]$ for E_1 and $A_2 = [b, b \Rightarrow \neg c]$ for E_2. So each one of the two extensions is able to respond to all defeaters of any argument it includes. A_1 and A_2 are merely defensible as they are incompatible, and we do not have, in the given set of arguments, reasons for preferring one to the other.

Assume that we add argument $[r_3: \Rightarrow\ r_1 \succ r_2\ ]$. Then we have just one preferred extension, namely $\{[a], [b], [a, r_1:\ a \Rightarrow c], [r_3: \Rightarrow\ r_1 \succ r_2\ ]\}$, since, according to the preference $r_3: \Rightarrow\ r_1 \succ r_2$, $A_1$is no longer defeated by A_2.

Moving from arguments to conclusion, we have two possibilities for defining what conclusions are justified. One option is to view a conclusion as justified when it is established by a justified argument. The other option consists in viewing a conclusion as justified when is supported in all preferred extensions through possibly through different arguments. More precisely, we get the following definition:

Definition (*Defensibility and justifiability*).

- *Defensibility*. Claim φ is defensible with regard to argument set $\mathcal{A}$ if there exists a preferred extension S of $\mathcal{A}$ that contains an argument with conclusion φ.
- *Strong-Justifiability*. Claim φ is strongly-justifiable with regard to argument set $\mathcal{A}$, if φ is the conclusion of an argument $\mathcal{A}$ that is contained in all preferred extensions.
- *Weak-Justifiability*. Claim φ is weakly- justifiable with regard to argument set $\mathcal{A}$ if all preferred extensions of contain arguments having conclusion φ.

Note that the weak definition of justifiability is broader than the strong, since it allows for a justifiable conclusion to be obtained through different incompatible arguments, included in different extensions. This is the notion that seems to be more appropriate to interpretation, as we shall argue in the following.

9. Interpretive Arguments

An interpretive argument can be constructed by combining an interpretive canon with the corresponding interpretive conditions. For instance, an argument from ordinary language would have the following form (in the argument for conciseness sake we put the general norm rather than its instantiation to the case at hand):

$Argument\ A_1$

1. expression "Loss" occurs in document 123(1)ERA
2. the interpretation of "Loss" in 123(1)ERA as *PecuniaryLoss* fits ordinary language
3. *OL*: expression E occurs in document D ∧
   the interpretation of E in D as *M* fits ordinary language ⇒
   $BestInt(E, D) \equiv M$

$BestInt$("Loss", 123(1)ERA) $\equiv$ $PecuniaryLoss$

Interpretive arguments can be attacked by counterarguments. For instance, the following counterargument based on technical language successfully rebuts the above argument based on ordinary language, by providing a different incompatible interpretation (assuming that no priority can be established, and that concepts are different when denoted with a different name):

$Argument\ A_2$

1. expression "Loss" occurs in document 123(1)ERA
2. the interpretation of "Loss" in 123(1)ERA as *PecuniaryOrEmotioalLoss* fits technical language
3. *TL*: expression E occurs in document D ∧
   the interpretation of E in D as *M* fits technical language ⇒
   $BestInt(E, D) \equiv$ *PecuniaryOrEmotionalLoss*

$BestInt$("Loss", 123(1)ERA) $\equiv PecuniaryOrEmotionalLoss$

The interpretation based on ordinary language could also attacked by directly denying its conclusion, for instance by a non-redundancy argument claiming that *"Loss"* should not be interpreted in this way, since this would make 123(1)ERA redundant.

$Argument\ A_3$

1. expression "Loss" occurs in document 123(1)ERA
2. the interpretation of "Loss" in 123(1)ERA as *PecuniaryLoss* makes the norm redundant
3. *NR*: expression E occurs in document D ∧
   the interpretation of E in D as *M* makes the norm redundant ⇒
   $BestInt(E, D) \not\equiv M$

$BestInt$("Loss", 123(1)ERA $\not\equiv$ $PecuiaryLoss$

A rebutting attack can also be played by using partial (inclusionary or exclusionary interpretive) arguments.

$Argument\ A_3$

1. expression "Loss" occurs in document 123(1)ERA
2. the interpretation of "Loss" in 123(1)ERA as $EmotionalLoss$ conflicts with usual meaning
3. *eAC:* expression *E* occurs in document *D*,
   the interpretation of expression *E* in the *D* as including *S* conflicts with usual meaning ⇒
   $BestInt(E, D)^C \sqsupseteq S$

$BestInt(\text{“Loss”}, 123(1)\text{ERA})^C \sqsupseteq EmotionalLoss$

where $BestInt(\text{“Loss”}, 123(1)\text{ERA})^C$ denotes the complement of $BestInt(\text{“Loss”}, 123(1)\text{ERA})$.

Given that $PecuniaryOrEmotionalLoss$ includes emotional loss, i.e.,

4. $PecuniaryOrEmotionalLoss \sqsupseteq EmotionalLoss$

we can conclude

5. $BestInt(\text{“Loss”}, 123(1)\text{ERA}) \not\equiv PecuniaryOrEmotionalLoss$

which contradicts the conclusion of the above argument A_2.

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, e.g., in the context of the discipline of industrial relations, where arguments from ordinary language do not apply. Thus this canon is inapplicable to the expression $Loss$ in $123(1)$ERA.

1. expression "Loss" occurs in document 123(1)ERA
2. 123(1)ERA is a technical context
3. *TC*: expression E occurs in document D,
   D is a technical context ⇒
   $\neg\text{Appl}(OL, E)$

$\neg\text{Appl}(OL, 123(1)\text{ERA})$

10. Preference Arguments over Interpretive Arguments

We may have preferences over 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 traditional interpretation as pecuniary damage, arguing that also damage to health should also be included in the scope of the term (and consequently compensated):

$$Argument\ A_1$$

1. expression "Loss" occurs in document Art2043ICC
2. the interpretation of "Loss" in Art2043ICC as $PecuniaryLoss$ fits legal history
3. OL: expression E occurs in document D,
   the interpretation of E in D as M fits legal history ⇒
   $BestInt(E, D) \equiv M$

$BestInt(\text{"Loss"}, \text{Art2043ICC} \equiv PecuniaryLoss)$

$$Argument\ A_2$$

1. expression "Loss" occurs in document Art2043ICC
2. the interpretation of "Loss" in Art2043ICC as $PecuniaryLossOrDamageToHealth$ contributes to substantive reasons
3. SR: expression E occurs in document D,
   the interpretation of E in D as M contributes to substantive reasons ⇒
   $BestInt(E, D) \equiv M$

$BestInt(\text{"Loss"}, \text{Art2043ICC}) \equiv PecuniaryLossOrDamageToHealth$

These two arguments conflict (rebut each other), as:

$$PecuniaryLoss \not\equiv PecuniaryLossOrDamageToHealth$$

To address the conflict, the judges argued that the second argument defeats the first, since SR in this context contributes to constitutional values.

$$Argument\ 3$$

1. The interpretation of expression "Loss" in Art2043ICC, as $PecuniaryLossOrDamageToHealth$ according to SR contributes to constitutional values
2. SR: The interpretation of expression E in D, as M according to SR contributes to constitutional values ⇒
   $SR(E, D, M) \succ LH(E, D, M')$

$SR(\text{“Loss”}, \text{Art2043ICC}, PecuniaryLossOrDamageToHealth\ ) \succ$
$LH(\text{“Loss”}, \text{Art2043ICC}, PecuniaryLoss)$

11. From Best Interpretations to Individual Claims

We must be able to move from interpretive claims to conclusions in individual cases, namely, from conceptual assertions to individual claims. For this purpose, we can adopt general patterns for strict rules, which provide for the transition from interpretive claims to assertions concerning individuals.

1. $BestInt(E, D) \equiv M \longmapsto \forall \boldsymbol{x}[E_D(\boldsymbol{x}) \leftrightarrow M(\boldsymbol{x})]$
2. $BestInt(E, D) \sqsupseteq M \longmapsto \forall \boldsymbol{x}[M(\boldsymbol{x}) \rightarrow E_D(\boldsymbol{x})]$
3. $BestInt(E, D)^C \sqsupseteq M \longmapsto \forall \boldsymbol{x}[M(\boldsymbol{x}) \rightarrow \neg E_D(\boldsymbol{x})]$

where $\boldsymbol{x}$ is sequence of variables which is required by concept *M*, $M(\boldsymbol{x})$ is the predicate corresponding to concept *M*, and E_D is a predicate representing the occurrence of *E* in *D* at issue. Consider for instance the above interpretive claim according to which

$$BestInt(\text{"}loss\text{"}, 125\text{ERA}) \equiv PecuniaryLoss$$

The corresponding instance of transition rule 1 would be:

$$BestInt(\text{"}loss\text{"}, 125\text{ERA}) \equiv PecuniaryLoss$$
$$\longmapsto \forall \boldsymbol{x}[Loss_{ERA}(x, y, z) \leftrightarrow PecuniaryLoss(x, y, z)]$$

To be read as: if the best interpretation of expression “loss” in document Section 125 of the Employment Relations Act is concept *PecuniaryLoss*, then a person *x* in an event *y* has a “loss” of amount *z* (as understood in Section 125 of the Employment Relations Act) if and only if *x* in *y* has a pecuniary loss of *z*.

Let us assume that John in his unfair dismissal by Tom had a pecuniary loss of Euro 100, i.e. $PecuniaryLoss(John, DismissalByTom, 100)$. Let us expand the ordinary language argument with the following: the latter assumption, the above instance of transition rule 1, and strict rules corresponding to an inference of classical logic. We get the following argument (where we list with the premises in the argument, and with letters the intermediate conclusions).

$Argument\ A_4$

1. expression “Loss” occurs in document 123(1)ERA
2. the interpretation of “Loss” in 123(1)ERA as *PecuniaryLoss* fits ordinary language
3. *OL*: expression E occurs in document D $\wedge$
   the interpretation of E in D as *M fits ordinary language* $\Rightarrow$
   $BestInt(E, D) \equiv M$

---

a. $BestInt(\text{“Loss”}, 123(1)\text{ERA} \equiv PecuniaryLoss$ (from 1, 2, and 3)

4. $BestInt(\text{"}loss\text{"}, 125\text{ERA}) \equiv PecuniaryLoss \longmapsto \forall \boldsymbol{x}[Loss_{ERA}(x,y,z) \leftrightarrow PecuniaryLoss(x,y,z)]$

---

b. $\forall \boldsymbol{x}[Loss_{ERA}(x,y,z) \leftrightarrow PecuniaryLoss(x,y,z)]$ (from a and 4)

5. $PecuniaryLoss(John, DismissalByTom, 100)$

---

c. $Loss_{ERA}(John, DismissalByTom, 100)$ (by classical logic) (from b and 5)

The mixture of interpretive and other arguments that are needed for a legal conclusion can also include additional conceptual relations. For instance, let us assume that we know that John has sustained a pecuniary loss of 100 Euros, as a consequence of his unfair dismissal. Since the concept of pecuniary loss is included in the concept of pecuniary or emotional loss, we can infer that he suffered a pecuniary or emotional loss. This conclusion would enable us to conclude that John has a loss in the sense of Section 125 ($Loss_{ERA}(John, DismissalByTom, 100)$), also on the basis of the interpretation of loss as $PecuniaryOrEmotionalLoss$, according to an argument $Argument\ A_5$ which includes this interpretation.

$Argument\ A_5$

1. expression "Loss" occurs in document 123(1)ERA
2. the interpretation of "Loss" in 123(1)ERA as $PecuniaryOrEmotionalLoss$ fits ordinary language
3. TL: expression E occurs in document D $\wedge$
   the interpretation of E in D as *M fits technical language* $\Rightarrow$
   $BestInt(E, D) \equiv M$

---

a. $BestInt$("Loss", 123(1)ERA $\equiv$ $PecuniaryOrEmotionalLoss$ (from 1, 2, and 3)

4. $BestInt(\text{"}loss\text{"}, 125\text{ERA}) \equiv PecuniaryOrEmotionalLoss \longmapsto \forall \boldsymbol{x}[Loss_{ERA}(x,y,z) \leftrightarrow PecuniaryOrEmotionalLoss\ (x,y,z)]$

---

b. $\forall \boldsymbol{x}[Loss_{ERA}(x,y,z) \leftrightarrow PecuniaryOrEmotionalLoss\ (x,y,z)]$ (from a, and 4)

5. $\forall \boldsymbol{x}[PecuniaryLoss(x,y,z) \rightarrow PecuniaryOrEmotionalLoss(x,y,z)]$

6. $PecuniaryLoss(John, DismissalByTom, 100)$

---

c. $PecuniaryOrEmotionalLoss(John, DismissalByTom, 100)$ (from 5, and 6)

---

d. $Loss_{ERA}(John, DismissalByTom, 100)$ (from 5 and c)

Arguments A_4 and A_5 are inconsistent, as they include incompatible interpretive conclusions (incompatible sub-arguments): according to conclusion (a) in A_4, the best interpretion of "loss" in Section 125 is *PecuniaryLoss*, while according conclusion (a) in $A_5$the best interpretation is a different concept, namely *PecuniaryOrEmotionalLoss*. However, the two arguments lead to the same conclusion in the case of John's dismissal: He suffers a loss of 100, as understood in Section 125 of the Employment Relations act.

Therefore, we may view this conclusion as legally justified, namely, as weakly justified. This is the case even though we are unable to make a choice between the two incompatible interpretations (the two competing interpretive arguments are both defeasible, and neither is justified), as the conclusion follows from both such interpretations. This view corresponds to the idea that only relevant issues have to be addressed in legal decision-making: the issue of whether "loss" is limited or not to pecuniary losses is irrelevant in John's case, since he has only suffered a pecuniary loss (this issue would be relevant if he had on the contrary suffered instead, or additionally, an emotional loss).

12. Conclusions

In this paper our goal was to show how the interpretive schemes can be formulated in such a manner that they can be incorporated into a formal and computational argumentation system such as CAS or APSIC+, and then applied to displaying the pro-contra structure of the argumentation using argument maps applied to legal cases. To this purpose, we have analyzed the most common types of statutory arguments and brought to light their common characteristics. We have shown how canons of interpretation can be translated into argumentation schemes, and we have distinguished two general macro-structures of positive and negative, total and partial canons, under which various types of schemes and rebuttals can be classified. This preliminary classification was then used for modeling the interpretive arguments formally, and integrating them into computational systems and argument maps.

The interpretive schemes can be applied initially when constructing an argument diagram to get an overview of the sequence of argumentation in a case of contested statutory interpretations. The schemes can be applied in order to help the argument analyst convey an evidential summary showing how the sub-arguments fit together in a lengthy sequence of argumentation in a case, as indicated in the main example of the educational grants case. The next step is to zoom in on parts of the argumentation sequence that pose a problem where critical questions need to be asked or refinements need to be considered. Here the critical questions can be applied in order to find further weak points in an argument by bringing out implicit premises that may have been overlooked and that could be questioned.

The function of the set of critical questions matching a scheme is to give the arguer who wants to attack the prior argument some idea of the kinds of critical questions that need to be asked in replying to it. Thus the critical questions can offer the respondent guidance as to where look for weak points that could be challenged. However, there are theoretical issues of how to

structure the critical questions. If critical questions can be modeled in the argument diagrams as additional premises, ordinary premises, assumptions or exceptions such as done in CAS or ASPIC+, they can be modeled in argument maps as undercutting or rebutting counter-arguments. The problem that always arises in attempts to fit critical questions into argument diagrams in this manner is one of burden of proof. Is merely asking a critical question enough to defeat a given argument? Or should a critical question be taken to defeat the given argument only if some evidence is given to back it up. CAS or ASPIC+ provides a way of dealing with this problem that has been shown to be applicable to interpretative schemes.

The danger with using such schemes to construct hypotheses about the best interpretation is one of jumping to a conclusion too quickly. This danger can be overcome by asking critical questions matching the scheme, and by considering possible objections to the argument fitting an interpretive scheme. For as we have seen in the example, a sequence of argumentation based on the application of interpretive argumentation schemes is defeasible, and can be attacked by undercutters and rebutters in an opposed sequence of argumentation. Indeed, it is this very situation of one sequence of interpretive argumentation being used to attack another one that is characteristic of the example we studied, a standard example of statutory interpretation.

We have also provided a fresh logical formalization of reasoning with interpretive canons. Rather that modelling interpretive conclusions as deontic claims, as we did in (Sartor et al., 2014), here we have modelled them as conceptual (terminological) claims concerning best interpretations.

We have then considered how interpretive arguments can be framed within argumentation systems, including defeasible and strict rules. We have argued that a semantics based on preferred extensions can provide an appropriate approach for drawing interpretive conclusions, and for distinguishing between defensible and justifiable interpretive claims. With regard to justification, we have argued for weak justifiability (derivation in all extensions, also through different arguments) to be more appropriate to interpretive reasoning in legal contexts.

This work still is quite preliminary, but necessarily so, since AI and law research has neglected issues pertaining to statutory interpretation, and more generally, the issue of determining the correct meaning of authoritative sources of the law. Further research should include a more refined classification system for interpretative schemes. Also the idea of merging argumentation with deontic logic as advanced in (Sartor et al., 2014; Walton, Macagno, & Sartor, 2014) needs to be reconsidered, and integrated with the different framework presented in this paper.

Acknowledgements

Douglas Walton would like to thank the Social Sciences and Humanities Research Council of Canada for award of Insight Grant 435-2012-0104 (2012-2018). Fabrizio Macagno would like to thank the Fundação para a Ciência ea Tecnologia for the research grants IF/00945/2013 and PTDC/MHC-FIL/0521/2014.

References

Alexy, R., & Dreier, R. (1991). Statutory interpretation in the Federal Republic of Germany. In N. MacCormick & R. Summers (Eds.), *Interpreting statutes. A comparative study*. Aldershot: Dartmouth.

Araszkiewicz, M. (2013). Towards systematic research on statutory interpretation in AI and Law. In R. Hoekstra (Ed.), *Proceedings of JURIX 2014: The Twenty-Seventh Annual Conference on Legal Knowledge and Information Systems* (pp. 15–24). Amsterdam: IOS Press.

Atlas, J. D. (2005). *Logic, meaning, and conversation: Semantical underdeterminacy, implicature, and their interface*. Oxford: Oxford University Press.

Atlas, J. D. (2008). Presupposition. In *The Handbook of Pragmatics* (pp. 29–52). http://doi.org/10.1002/9780470756959.ch2

Atlas, J. D., & Levinson, S. (1981). It-clefts, informativeness and logical form: Radical pragmatics (revised standard version). In P. Cole (Ed.), *Radical pragmatics* (pp. 1–62). New York: Academic Press.

Bezuidenhout, A. (1997). Pragmatics, semantic undetermination and the referential/attributive distinction. *Mind*, *106*(423), 375–409. http://doi.org/10.1093/mind/106.423.375

Capone, A. (2009). Are explicatures cancellable? Toward a theory of the speaker's intentionality. *Intercultural Pragmatics*, *6*(1), 55–83. http://doi.org/10.1515/IPRG.2009.003

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell Publishing Ltd.

Carston, R. (2013). Legal texts and canons of construction: A view from current pragmatic theory. In M. Freeman & F. Smith (Eds.), *Law and Language: Current Legal Issues* (Vol. 15, pp. 8–33). Oxford: Oxford University Press.

Cross, R. (2005). *Statutory Interpretation*. (J. Bell & G. Engle, Eds.). Oxford: Oxford University Press.

Dascal, M. (2003). *Interpretation and understanding*. Amsterdam: John Benjamins Publishing Company.

Dascal, M., & Wróblewski, J. (1988). Transparency and doubt: Understanding and interpretation in pragmatics and in law. *Law and Philosophy*, *7*(2), 203–224.

Dung, P. (1995).On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games. *Artificial Intelligence*, 77(2), 321–357.

Gizbert-Studnicki, T. (1990). The burden of argumentation in legal disputes. *Ratio Juris*, *3*(1), 118–129.

Gordon, T. (2010). An overview of the Carneades argumentation support system. In C. Reed & C. W. Tindale (Eds.), *Dialectics, dialogue and argumentation. An examination of douglas Walton's theories of reasoning and argument* (pp. 145–156). London: College Publications.

Gordon, T., & Walton, D. (2009a). Legal reasoning with argumentation schemes. In C. D. Hafner (Ed.), *Proceedings of the 12th International Conference on Artificial Intelligence and Law* (pp. 137–146). New York: ACM.

Gordon, T., & Walton, D. (2009b). Proof burdens and standards. In I. Rahwan & G. Simari (Eds.), *Argumentation in Artificial Intelligence* (pp. 239–258). Berlin: Springer.

Gordon, T., & Walton, D. (2011). A formal model of legal proof standards and burdens. In F. van Eemeren, B. Garssen, A. Blair, & G. Mitchell (Eds.), *7th Conference on Argumentation of the International Society for the Study of Argumentation (ISSA 2010)* (pp. 644–655). Amsterdam: Sic Sat.

Hage, J. (1996). A theory of legal reasoning and a logic to match. *Artificial Intelligence and Law*, *4*(3-4), 199–273.

Hage, J. (1997). *Reasoning with rules*. Dordrecht: Kluwer Academic Publishers.

Horn, L. (1995). Vehicles of meaning: unconventional semantics and unbearable interpretation. *Washington University Law Quarterly*, *73*, 1145–1152.

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, Mass.: MIT Press.

Macagno, F. (2015). A means-end classification of argumentation schemes. In F. van Eemeren & B. Garssen (Eds.), *Reflections on theoretical issues in argumentation theory* (pp. 183–201). Cham: Springer.

Macagno, F., & Capone, A. (2015). Interpretative disputes, explicatures, and argumentative reasoning. *Argumentation*, 1–24. http://doi.org/10.1007/s10503-015-9347-5

Macagno, F., Sartor, G., & Walton, D. (2012). Argumentation schemes for statutory interpretation. In J. Šavelka, M. Araszkiewicz, M. Myška, T. Smejkalová, & M. Škop (Eds.), *ARGUMENTATION 2012. International Conference on Alternative Methods of Argumentation in Law* (pp. 63–75). Brno: Masarykova univerzita.

Macagno, F., & Walton, D. (2011). Reasoning from paradigms and negative evidence. *Pragmatics & Cognition*, *19*(1), 92–116. http://doi.org/10.1075/pc.19.1.04mac

Macagno, F., & Walton, D. (2014). *Emotive language in argumentation*. New York: Cambridge University Press.

Macagno, F., & Walton, D. (2015). Classifying the patterns of natural arguments. *Philosophy & Rhetoric*, *48*(1), 26–53. http://doi.org/10.1353/par.2015.0005

MacCormick, N. (2005). *Rhetoric and the rule of law: a theory of legal reasoning*. Oxford: Oxford University Press.

MacCormick, N., & Summers, R. (Eds.). (1991). *Interpreting statutes: a comparative study*. Aldershot: Dartmouth.

Pollock, J. (1995). *Cognitive Carpentry*. Cambridge, Mass.: MIT Press.

Prakken, H. (2010). An abstract framework for argumentation with structured arguments. *Argument & Computation*, *1*(2), 93–124. http://doi.org/10.1080/19462160903564592

Prakken, H., & Sartor, G. (1996). A dialectical model of assessing conficting arguments in legal reasoning. *Artificial Intelligence and Law*, *4*, 331–368.

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, *13*(1-2), 81–132. http://doi.org/10.1016/0004-3702(80)90014-4

Rotolo, A., Governatori, G., & Sartor, G. (2015). Deontic defeasible reasoning in legal interpretation: two options for modelling interpretive arguments. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law* (pp. 99–108). New York: ACM.

Sartor, G., Walton, D., Macagno, F., & Rotolo, A. (2014). Argumentation schemes for statutory interpretation: A logical analysis. In R. Hoekstra (Ed.), *Legal Knowledge and Information Systems: JURIX 2014: The Twenty-Seventh Annual Conference* (pp. 11–20). Amsterdam: IOS Press.

Schauer, F. (1987). Precedent. *Stanford Law Review*, 571–605.

Soames, S. (2008). *Philosophical Essays, Volume 1: Natural Language: What It Means and How We Use It*. Princeton: Princeton University Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Tarello, G. (1980). *L'interpretazione della legge.* Milano: Giuffrè.

Verheij, B. (2008). About the logical relations between cases and rules. In E. Francesconi, G. Sartor, & D. Tiscornia (Eds.), *Legal Knowledge and Information Systems. JURIX 2008: The Twenty-First Annual Conference* (pp. 21–32). Amsterdam: IOS Press.

Walton, D. (2004). *Abductive Reasoning*. Tuscaloosa: University of Alabama Press.

Walton, D. (2010). Similarity, precedent and argument from analogy. *Artificial Intelligence and Law*, *18*(3), 217–246. http://doi.org/10.1007/s10506-010-9102-z

Walton, D. (2013). *Argumentation schemes for presumptive reasoning*. Mahwah: Lawrence Erlbaum Publishers.

Walton, D. (2015). *Goal-based Reasoning for Argumentation*. Ne: Cambridge University Press.

Walton, D., & Gordon, T. (2005). Critical questions in computational models of legal argument. In P. Dunne & T. Bench-Capon (Eds.), *Argumentation in Artificial Intelligence and Law, IAAIL Workshop Series* (pp. 103–111). Nijmegen: Wolf Legal Publishers.

Walton, D., Macagno, F., & Sartor, G. (2014). Interpretative Argumentation Schemes. *JURIX-2014. IOS*, 21–22.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. New York: Cambridge University Press.

Wilson, D., & Sperber, D. (2004). Relevance theory. In L. Horn & G. Ward (Eds.), *Handbook of Pragmatics* (pp. 607–632). Oxford: Blackwell. http://doi.org/10.1016/j.pragma.2009.09.021