A Generalisation of Parsons' and Jennings Theory of Argument

Graham White
Computer Science
Queen Mary, University of London
London E1 4NS
graham@dcs.qmul.ac.uk

Abstract

We give a sequent-calculus based inference system which generalises the system of Parsons and Jennings.

Contents

1 The Theories

1.1 Parsons and Jennings

Parsons and Jennings ([1996]; see also [Parsons *et al.*, 1998]) have described a consequence relation, $\vdash_{\text{ACR}}$, which is intended to capture the practice of argumentation. It has the following properties:

1. the basic objects are pairs (p, A), where p is a sentence and A is a set of (labels for) sentences, the *grounds* of p
2. the inference rules of this system act on p: their effect on the grounds is minimal
3. the rules for p do not handle quantification.
4. the rules for p do not handle modal logic.

Although the system has a great deal of intuitive appeal, neither Feature 2, Feature 3, nor Feature 4 is tremendously satisfactory.

Feature 2 *Inferential* relations between grounds are not supported: for example, Parsons and Jennings define a relation of *undercutting* between arguments, which occurs when we have two arguments, and a conjunct of the conclusion of one of them is of the form Mp, whereas a conjunct of the conclusion of the other is of the form $M'q$, for modalities M and M', and where p is *equivalent to* the negation of q. Although this does represent an inferential relation between the premises of one argument and the conclusion of the other (or, rather, it does given appropriate assumptions on the modalities), it is far from being the only such inferential relation that there could be.

Example 1. Consider the arguments

$$(\text{this swan is white, all swans are white})$$
$$(\text{Australian swans are black, I read it in a book})$$

Now the second argument ought to undercut the first: however, the relation between the conclusion of one and the premise of the other is merely that of incompatibility, rather than one being equivalent to the negation of the other.

Feature 3 Parsons and Jennings admit that

> It should be noted that [our system] says nothing about quantification, so that it is not possible to instantiate rules such as $P(x) \to Q(x)$ given $P(a)$ in order to learn $Q(a)$. This, however, does not prevent us from handling first-order rules – instead, first-order rules are instantiated with every ground instance before arguments are constructed. In other words, if we have $P(x) \to Q(x)$ and $P(a)$ and $P(b)$, we replace $P(x) \to Q(x)$ with $P(a) \to Q(a)$ and $P(b) \to Q(b)$ before any arguments are constructed.

Although this solves the problem in simple cases, it cannot deal with propositions of the form $\forall x \exists y \Psi(x, y)$.

Example 2. We cannot deal, for example, with a rule like

> every picture must hang from something

which would be quite relevant in their picture-hanging example.

Feature 4 Parsons and Jennings want their base language to include the BDI modalities of Krause *et al.* [1993]. They also define relations – *rebuttal* and *undercutting* – between arguments, and these relations appear to come from inferential relations between, respectively, the conclusions of two argument, or between the conclusion of one argument and the grounds for another, or, possibly, between various argument conclusions and/or grounds and an implicit constraint on the joint compatibility of agents' intentions. However, these inferential relations are modal, and thus cannot be defined in terms of $\vdash_{\text{ACR}}$: Parsons and Jennings simply introduced them *ad hoc*.

Example 3. Consider the three arguments

$$(\mathbf{I}_i \neg p, a)$$
$$(\mathbf{I}_j \neg q, b)$$
$$(\mathbf{I}_k (p \vee q), c)$$

where p, q, a, b, and c are propositions, and $\mathbf{I}_i$ is the intention modality for agent i. The conclusions of these arguments clearly represent intentions which cannot be simultaneously satisfied, and can surely be regarded as rebutting each other just as much as do examples of Parsons and Jennings (see, for example, [Parsons *et al.*, 1998, Section 4.3]).

1.2 Our Proposal

We propose, instead, the modal system of [White, 2002; 2003]; this reformulates the work of McCain and Turner [1997], which was originally intended to deal with *causal* reasoning. However, in retrospect McCain and Turner's system seems to be much more to do with explanation in general, and it handles *causal* explanation as a special case (see, for example, [Lifschitz, 1997], which quite explicitly uses explanatory terminology). It is, then, no surprise that this system can also be applied to argumentation: explanation and argument are closely linked.

We will start with some logical system or other, which we call the *underlying logic*: this may be a BDI modal logic such as Parsons and Jennings use in [Parsons *et al.*, 1998], but our results are certainly not limited to such a choice of base logic. Suppose, then, that we take – as do Parsons and Jennings – a set Δ of basic arguments: each of them is of the form (ψ, ϕ), where ϕ is the *premise* and ψ the *conclusion*, and both are formulae of the base logic. (To be pedantic, ψ is here the conjunction of all the propositions in Parsons and Jennings' premise.) Then we can define a modal system, with an operator $\Box_\Delta$, satisfying the following three conditions:

basic arguments we have $\vdash \phi \rightarrow \Box_\Delta \psi$ for each basic explanation $(\psi, \phi) \in \Delta$,

K For any Δ, $\Box_\Delta$ is a **K** modality, and

minimality $\Box_\Delta$ is the minimal modality with these properties: that is, for any Δ, and for any other modality $\Box'_\Delta$ satisfying items 1 and 2 with respect to Δ, we have $\Box_\Delta p \vdash \Box'_\Delta p$, for any formula p.

In this system, we can express Parsons and Jennings' entailment $\Delta \vdash_{\text{ACR}} (q, p)$ as

$$p \vdash \Box_\Delta q.$$

The Missing Features

In this section, we show how our system can deal with the features that Parsons and Jennings's system lacks.

Feature 2 If we have an argument for p on grounds A, i.e. a valid entailment $A \vdash \Box_\Delta p$, and if, also, we know that $B \vdash A$, then we have an argument $B \vdash \Box_\Delta p$; that is, we can handle inferential relations between premises. Relations of undercutting can be handled similarly.

Feature 3 Quantification can be handled, because our system can have basic arguments with open formulae in them: consequently, our example "Every picture must hang from something" can be written as

$$\text{picture(x)} \vdash \Box_\Delta(\exists y.\text{hangsFrom}(x, y))$$

Feature 4 We can also handle more general entailment relations between heads of arguments: for example, if we have entails $a \vdash \Box\neg p$, $b \vdash \Box\neg q$, and $c \vdash \Box(p \vee q)$, we can derive $a, b, c \vdash \Box \perp$, and so, if we have $\vdash \neg\Box \perp$ as a premise, we can derive $\neg(a \wedge b \wedge c)$.

1.3 Relation With Parsons and Jennings: Soundness

In this section, we show that our interpretation of the Parsons and Jennings rules is sound: that is, that the translation of each of the Parsons and Jennings rules is valid in our system.

We first show

Lemma 1. *If $\Box_\Delta$ is the modality associated to a set Δ of explanations, and if $\tilde{\Box}_{\Delta\cup\{(\psi,\phi)\}}$ is associated to $\Delta \cup \{(\psi, \phi)\}$, then we have*

$$\Box_{\Delta\cup\{(\psi,\phi)\}} p \cong (\phi \wedge \Box_\Delta(\psi \rightarrow p) \vee \Box p$$

for any p.

Proof. For convenience, call the expression on the right $\tilde{\Box} p$. We will also write $\Box$ for $\Box_\Delta$ in the remainder of this proof. Note that, for any p, $\Box p \vdash \tilde{\Box} p$. We first show that $\tilde{\Box}$ is, in fact, a **K** modality. We have to show the following three properties:

1. $p \rightarrow q \vdash \tilde{\Box} p \rightarrow \tilde{\Box} q$. After the obvious $\rightarrow$ R stages, we have:

$$\dfrac{\dfrac{\Pi}{p \rightarrow q, \phi \wedge \Box(\psi \rightarrow p) \vdash \phi \wedge \Box(\psi \rightarrow q) \vee \Box q} \qquad \dfrac{\Pi'}{p \rightarrow q, \Box p \vdash \phi \wedge \Box(\psi \rightarrow q) \vee \Box q}}{p \rightarrow q, \phi \wedge \Box(\psi \rightarrow p) \vee \Box p \vdash \phi \wedge \Box(\psi \rightarrow q) \vee \Box q}$$

   Here Π' is simply the proof of the corresponding rule for $\Box$, followed by $\vee$R. Π applies $\vee$R followed by $\wedge$R, and then has to prove the two branches

$$p \rightarrow q, \phi \wedge \Box(\psi \rightarrow p) \vdash \phi$$

   (which is trivial), and

$$p \rightarrow q, \phi \wedge \Box(\psi \rightarrow p) \vdash \Box(\psi \rightarrow q),$$

   which follows from the corresponding rule for $\Box$ together with the tautology

$$p \rightarrow q \vdash (\psi \rightarrow p) \rightarrow (\psi \rightarrow q).$$

2. $\vdash \tilde{\Box}\top$. This is immediate, since $\vdash \Box\top$, and $\Box p \vdash \tilde{\Box} p$ for any p.
3. $\tilde{\Box} p, \tilde{\Box} q \vdash \tilde{\Box}(p \wedge q)$. We apply $\vee$L twice, once to the disjunction in each $\tilde{\Box}$ on the left, and have to prove the following sequents:
   (a) $\phi \wedge \Box(\psi \rightarrow p), \phi \wedge \Box(\psi \rightarrow q) \vdash (\phi \wedge \Box(\psi \rightarrow p \wedge q)) \vee \Box(p \wedge q)$

(b) $\phi\wedge\Box(\psi \to p),\ \Box q \vdash (\phi\wedge\Box(\psi \to p\wedge q))\vee\Box(p\wedge q)$
(c) $\Box p,\ \phi\wedge\Box(\psi \to q) \vdash (\phi\wedge\Box(\psi \to p\wedge q))\vee\Box(p\wedge q)$
(d) $\Box p,\ \Box q \vdash (\phi \wedge \Box(\psi \to p \wedge q)) \vee \Box(p \wedge q)$

The first is clear: $\psi \to p, \psi \to q \vdash \psi \to (p \wedge q)$, so, since $\Box$ is a **K** modality,

$$\Box(\psi \to p),\ \Box(\psi \to q) \vdash \Box(\psi \to (p \wedge q)$$

which yields the result. The second and third are similar, using instead the entailments $\psi \to p, q \vdash \psi \to (p \wedge q)$ and $p, \psi \to q \vdash \psi \to (p \wedge q)$. For the last one, we simply use the fact that $\Box$ is a **K** modality.

We next show that $\tilde{\Box}$ satisfies the basic arguments condition, i.e. that $\vdash \phi' \to \tilde{\Box}\psi'$ for every basic argument (ψ', ϕ'). There are two cases:

1. $(\psi', \phi') \in \Delta$. In this case, we have $\vdash \phi' \to \Box\psi'$, by the properties of $\Box$: however, $\vdash \Box p \to \tilde{\Box}p$ for any p, so we have the result.
2. $(\psi', \phi') = (\psi, \phi)$. We prove this as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\phi \vdash \phi,\ \Box\psi} \qquad \begin{array}{c}\Pi\\ \vdots\\ \phi \vdash \Box(\psi \to \psi), \Box\psi\end{array}}{\phi \vdash (\phi \wedge \Box(\psi \to \psi)),\ \Box\psi}}{\phi \vdash (\phi \wedge \Box(\psi \to \psi)) \vee \Box\psi}}{\vdash \phi \to (\phi \wedge \Box(\psi \to \psi)) \vee \Box\psi}$$

where Π uses the entailment $\vdash \psi \to \psi$ and the fact that $\Box$ is a **K** modality.

Finally we show that $\tilde{\Box}$ is minimal, i.e. that, if $\Box'$ is another modality with these properties, $\tilde{\Box}p \vdash \tilde{\Box}'p$, for any p. Note first that, since, by assumption, $\Box'$ satisfies the defining properties for $\Delta \cup \{(\psi, \phi)\}$, it satisfies the defining properties for Δ: consequently, since $\Box$ is minimal with these properties, we have $\Box p \vdash \Box' p$, for all p. Now we prove the result:

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\Pi_1\\ \vdots\\ \phi \vdash \Box'\psi\end{array} \qquad \dfrac{\begin{array}{c}\Pi_2\\ \vdots\\ \Box(\psi \to p) \vdash \Box'(\psi \to p)\end{array} \quad \begin{array}{c}\Pi_3\\ \vdots\\ \Box'\psi, \Box'(\psi \to p) \vdash \Box' p\end{array}}{\Box'\psi, \Box(\psi \to p) \vdash \Box' p}}{\phi,\ \Box(\psi \to p) \vdash \Box' p}}{\phi \wedge \Box(\psi \to p) \vdash \Box' p} \qquad \begin{array}{c}\Pi_4\\ \vdots\\ \Box p \vdash \Box' p\end{array}}{(\phi \wedge \Box(\psi \to p)) \vee \Box p \vdash \Box' p}$$

Here Π_2 and Π_4 come from the minimality of $\Box$, as explained above: Π_1 and Π_3 come from the assumptions on $\Box'$.

Consequently, $\tilde{\Box}$ satisfies the defining properties for $\Delta \cup \{(\psi, \phi\}$, and is minimal among the modalities which do; however, $\Box_{\Delta\cup\{(\psi,\phi\}}$ is also minimal among such modalities, and so, for any p, $\tilde{\Box}p \cong \Box_{\Delta\cup\{(\psi,\phi\}}p$. $\Box$

Proposition 1. *The modal interpretation is* sound*: that is, each of Parsons and Jennings' axioms is translated into a tautology.*

Proof. **Ax** is

$$\dfrac{}{\Delta \vdash_{\text{ACR}} (p, \{a\})}\,(p, a) \in \Delta$$

and this follows from the basic arguments condition.

$\wedge$–I, $\neg$–E, $\to$–E $\wedge -$ I, for example, is

$$\dfrac{\Delta \vdash_{\text{ACR}} (p, A) \quad \Delta \vdash_{\text{ACR}} (q, B)}{\Delta \vdash_{\text{ACR}} (p \wedge q, A \cup B)}$$

and this follows from the modal tautology $\Box p \wedge \Box q \vdash \Box p \wedge q$. $\neg -$ E and $\to -$E are similar.

$\wedge$–E1, $\wedge$–E2, $\vee$-I1, $\vee$-I2, EFQ $\wedge -$ E1, for example, is

$$\dfrac{\Delta \vdash_{\text{ACR}} (p \wedge q, A)}{\Delta \vdash_{\text{ACR}} (p, A)}$$

and this follows from the **K** tautology $\vdash (a \to b) \to \Box a \to \Box b$. $\wedge -$ E2, $\vee -$ I1, $\vee$I2, and EFQ are similar.

$\top$–I This is just

$$\vdash \Box\top,$$

a **K** tautology.

$\vee$–E This is

$$\dfrac{\begin{array}{l}\Delta \vdash_{\text{ACR}} (p \vee q, A)\\ \qquad\Delta, (p, A) \vdash_{\text{ACR}} (r, B)\\ \qquad\qquad\Delta, (q, A) \vdash_{\text{ACR}} (r, C)\end{array}}{\Delta \vdash_{\text{ACR}} (r, B \cup C)}$$

and this corresponds, in our system, to

$$\begin{array}{l}A \to \Box_\Delta(p \vee q),\\ \qquad B \to \Box_{\Delta\cup\{(p,A)\}}r,\ C \to \Box_{\Delta\cup\{(q,A)\}}r\\ \qquad\qquad\qquad\qquad \vdash B \wedge C \to \Box_\Delta r \qquad (1)\end{array}$$

We can use the lemma to express $\Box_{\Delta\cup\{(p,A)\}}$ and $\Box_{\Delta\cup\{(q,A)\}}$ in terms of $\Box_\Delta$; some routine but tedious computation then reduces this case to

$$\Box_\Delta(p \vee q), \Box_\Delta(p \to r), \Box_\Delta(q \to r) \vdash \Box_\Delta r$$

which is a **K** tautology.

RAA, $\neg$–I RAA is

$$\dfrac{\Delta, (p, \emptyset) \vdash_{\text{ACR}} (\bot, A)}{\Delta \vdash_{\text{ACR}} (\neg p, A)}$$

which corresponds to

$$A \to \Box_{\Delta\cup\{(p,\emptyset)\}} \bot \vdash A \to \Box_\Delta \neg p;$$

using the lemma on $\Box_{\Delta\cup\{(p,\emptyset)\}}$, and some computation, reduces this to

$$\Box \bot \vee \Box\,(p \to \bot) \vdash \Box\neg p$$

which is a **K** tautology. $\neg$–I is similar.

$\to$–I This is

$$\dfrac{\Delta, (p, \emptyset) \vdash_{\text{ACR}} (q, A)}{\Delta \vdash_{\text{ACR}} (p \to q, A)}$$

which corresponds to

$$A \to \Box_{\Delta\cup\{(p,\emptyset)\}}q \vdash A \to \Box\Delta(p \to q).$$

The usual moves reduce this to

$$\Box q \vee \Box(p \to q) \vdash \Box(p \to q)$$

again a **K** tautology.

$\Box$

1.4 Completeness: Two Approaches

We have shown how to associate, to every proof in Parsons and Jennings' system, a proof in ours. The question then arises of whether anything like a converse is valid. There are, basically, two approaches to this, model-theoretic and proof-theoretic: we can only sketch these results here, but we give references to fuller discussions.

Model Theory

This is a modal system, and questions of completeness can be settled by exhibiting the canonical model: a proposition will be a modal tautology iff it holds globally on the canonical model. We can show the following:

Proposition 2. *If the underlying system is non-modal, then the canonical model for our system has, for worlds, the set of models of the underlying system, and, for an accessibility relation, the relation R_Δ which is the largest relation such that*

$$wR_\Delta w' \quad \Leftrightarrow \quad (p, A) \in \Delta, w \Vdash A \Rightarrow w' \Vdash p.$$

Proposition 3. *If the underlying system is modal, then the canonical model for our system has, for worlds, the set of worlds of the canonical model of the underlying system, and, for an accessibility relation, the relation R_Δ which is the largest relation such that*

$$wR_\Delta w' \quad \Leftrightarrow \quad (p, A) \in \Delta, w \Vdash A \Rightarrow w' \Vdash p.$$

A proof of the first is sketched in [White, 2002]; the proof of the second needs only trivial modifications. The accessibility relation R_Δ is implicit in McCain and Turner's system of "causal reasoning" [1997].

This result shows that $\Box_\Delta$ captures a Kripke-style semantics, and it is a semantics with a good intuitive meaning: the accessibility relation is given by the following condition (which is simply a reformulation of the relation given implicitly in [McCain and Turner, 1997]; see [White, 2002]).

Proposition 4. *For worlds w and w', we have $wR_\Delta w'$ iff w' is a model of the set of conclusions of all the basic arguments whose premises are true in w.*

This may be intuitive: however, it is not, as it stands, tremendously suited to computation (although one could probably find a tableau method), and it is also hard to link it up with Parsons and Jennings' proof system.

Proof Theory

Suppose, now, that we have a sequent calculus formulation of the underlying system which satisfies cut elimination. We can – using results of Schröder-Heister [1992] – give a sequent calculus formulation of the modalised system which likewise satisfies cut elimination: the relevant right rule for $\Box_\Delta$ is

$$\frac{A \vdash \phi_1 \wedge \ldots \wedge \phi_k, B \quad \psi_1, \ldots, \psi_k \vdash p}{A \vdash \Box_\Delta p, B}$$

where A and B are sets of formulae, p is a formula, and where, for each i, $(\psi_i, \phi_i) \in \Delta$. There is a corresponding left rule: for the details, see, as usual, [White, 2002].

Using this cut elimination result, we can, by examining cut-free proof trees, prove the following:

Proposition 5. *Suppose that we have $A \vdash \Box_\Delta p$. Then there is a set of formulae $\{b_i\}_{i \in I}$ such that $A \vdash \bigwedge_{i \in I} b_i$, and such that, for each i, we have*

1. *a basic argument (ψ_i, b_i) and a proof of $\Delta \vdash_{ACR} (\psi_i, b_i)$, and*
2. *a proof of $\psi_i \vdash p$ in the underlying logic.*

This shows what role the $\vdash_{\mathrm{ACR}}$ entailment plays in our enlarged system.

Cut Elimination for BDI Parsons and Jennings use a modal BDI system as their underlying logic: we would like, then, to have cut elimination results for this. The dominant theoretical framework for BDI logics is model-theoretic, and uses tableau methods for a decision procedure [Rao and Georgeff, 1995a; 1991; 1995b]; nevertheless, there has been some work on sequent calculus formalisms, including the cut free formalism of Naoyuki and Takata [2002]. This allows us to prove a cut elimination result for our logic, at least for the BDI logic that Naoyuki and Takata discuss.

2 Defeat Between Arguments

Parsons and Jennings use their system to classify arguments into various grades of merit. However, the translation, though straightforward, is not entirely literal. The main problem is this: because, in our system, we can compose an entailment $A \vdash \Box p$ with an arbitrary inference $B \vdash \bigwedge_{a \in A} a$, our system has many more arguments than theirs. In particular, every proposition has a trivial argument for it, namely $\perp \vdash \Box p$.

We can, furthermore, generalise their relations of rebuttal and defeat:

Definition 1 (Parsons and Jennings). The argument (s, A) *rebuts* (s', B) if s and s' are conjunctions of sentences of the form $M^1(p_1) \wedge \ldots \wedge M^n(p_n)$ and $M'^1(q_1) \wedge \ldots \wedge M'^m(q_m)$, where the M and M' are belief, desire, or intention modalities, and if $p_i \cong \neg q_j$ for some i and j. (s, A) *undermines* (s', B) if one of the B is of the form $M''^1(r_1) \wedge \ldots \wedge M''^m(r_l)$, and if, for some i and j, $p_i \cong \neg r_j$.

This seems a little *ad hoc*. But the criteria they give, under suitable assumptions on the BDI modalities, entail respectively that $s, s' \vdash \perp$, or that $s, B \vdash \perp$. We can, then, generalise Parsons and Jennings' notion of defeat to the following:

Definition 2. An argument $A \vdash \Box s$ *defeats* an argument $B \vdash \Box s'$ if $s, s' \vdash \perp$: $A \vdash \Box s$ *undercuts* $B \vdash \Box s'$ if $s, B \vdash \perp$.

We can say more on this later, when we have dealt with the classification of arguments.

Classes of Acceptability

Parsons and Jennings – following [Elvang-Gøransson *et al.*, 1993; Krause *et al.*, 1995] – define the following. We fix a particular database Δ, and, consequently, we will omit subscripts on modal operators: we will simply write the operator $\Box$.

We will quote the following proposition, which follows from [White, 2002]; in effect, it gives a left rule for $\Box$.

Proposition 6. *Let s be a proposition in our base language: then*

$$\Box s \cong \bigvee_{\psi_1,\ldots,\psi_k \vdash s} \phi_1 \wedge \ldots \wedge \phi_k$$

where $\psi_1, \ldots, \psi_k$ range over the set of all collections of basic argument conclusions which entail s, and where the ϕ_i are the corresponding basic argument premises.

A1 The class of all arguments from Δ, i.e. the class of all valid entailments $A \vdash \Box s$.

A2 The class of all non-trivial arguments, i.e. the arguments for which $A \nvdash \bot$.

Lemma 2. *A proposition s is only supported by trivial arguments iff $\vdash \Diamond \neg s$.*

Proof. We have

$$\Box s \quad \cong \quad \bigvee_{\psi_1,\ldots,\psi_k \vdash s} \phi_1 \wedge \ldots \wedge \phi_k,$$

with the usual conditions on the ϕs and ψs. The right hand side of this entails $\bot$ iff every argument for s is trivial: however, $\Box s \vdash \bot$ iff $\vdash \Diamond \neg s$. □

A3 The class of all arguments for which there are no rebutting arguments.

Now having no rebutting arguments is purely a property of the proposition s: and, in fact, we have

Lemma 3. *A proposition s has no non-trivial rebutting arguments, iff $\vdash \Diamond s$.*

Proof. Suppose that s has no non-trivial rebutting arguments: that is, if, for a finite set of rule heads $\{\psi_i, \ldots, \psi_k\}$ with $\psi_i, \ldots, \psi_k \vdash \neg s$, we have $\phi_1, \ldots, \phi_k \vdash \bot$. So

$$\Box \neg s \vdash \bigvee_{\psi_i,\ldots,\psi_k \vdash \neg s} \phi_1 \wedge \ldots \wedge \phi_k \vdash \bot;$$

contraposing, we have the result.

Conversely, suppose that we have $\vdash \Diamond s$. Then, if $A \vdash \Box \neg s$ is a rebutting argument, then, by the assumption, $\Box \neg s \vdash \bot$, and so we have $A \vdash \bot$. So the rebutting argument is trivial. □

A4 The class of arguments for propositions for which there are no undercutting arguments.

Lemma 4. *If an argument $A \vdash \Box s$ has no non-trivial undercutting arguments, then we have $\vdash \Diamond \Box s$.*

Proof. Suppose that the argument has no non-trivial undercutting arguments: then there is no non-trivial argument for $\neg A$. By the previous lemma, we have $\vdash \Diamond A$; consequently, we have

$$\vdash \Diamond A \vdash \Diamond \Box s,$$

and hence the result. □

Lemma 5. *A proposition s is such that some argument for it has no non-trivial undercutting arguments iff $\vdash \Diamond \Box s$.*

Proof. Only if follows from the previous lemma. Conversely, suppose that we have a proposition s, and an argument $A \vdash \Box s$ for which there is a non-trivial undercutting argument: that is, a non-trivial argument B for $\neg A$. Thus, we have $B \vdash \Box \neg A$, and we may assume that B is a conjunction of rule bodies. Suppose that we have $\vdash \Diamond \Box s$: then, contraposing, we have

$$\Box \neg \Box s \vdash \bot$$

Now the left hand side of this is

$$\bigvee_{\psi_1,\ldots,\psi_k \vdash \neg \Box s} \phi_1 \wedge \ldots \wedge \phi_k,$$

so, for each such disjunct,

$$\phi_1 \wedge \ldots \wedge \phi_k \vdash \bot$$

Now the conjunction of the ϕs is an argument for $\neg \Box s$: that is, an argument undercutting an argument for s. □

A5 This is the class of tautological arguments: it follows immediately from our basic equivalence that we have

Lemma 6. *A formula s is supported by a tautologous argument iff we have*

$$\vdash \Box s$$

Proposition 7. *We thus have the following "dictionary" relating modal entailments and equivalent argument classes:*

A1	***A2***	***A3***	***A4***	***A5***
	$\nvdash \neg \Box s$	$\vdash \Diamond s$	$\vdash \Diamond \Box s$	$\vdash \Box s$

This is to be read as follows: a formula s satisfies the modal entailment in the bottom row iff there is a argument supporting s which belongs to the corresponding argument class in the top row.

2.1 Inclusions

Parsons and Jennings would like there to be inclusions

$$\mathbf{A1} \supseteq \mathbf{A2} \supseteq \mathbf{A3} \supseteq \mathbf{A4} \supseteq \mathbf{A5}$$

but do not make precise the conditions under which these are valid. (By contrast, Elvang-Gøransson *et al.* explicitly define each class as a subclass of the preceding one.)

In the terms in which Parsons and Jennings state the proposition, there are easy counterexamples.

Example 4 (A2 $\not\supseteq$ A3). Suppose that the database consists of the single basic argument

$$\bot \quad \vdash \quad \Box p$$

for some propositional atom p. Then the argument $\bot \vdash \Box p$ is valid, and it has no rebutting arguments (because there are no arguments with a head which entails $\neg p$). Consequently, it is in **A3**. However, it is also trivial: so it is not in **A2**. Note that the argument applies to Parsons and Jennings' notion of argument as well as our own.

Example 5 (A3 $\not\supseteq$ A4). Suppose that the database of arguments consists of the basic arguments

$$\begin{array}{ccc} p & \vdash & \Box q \\ p & \vdash & \Box \neg q \end{array}$$

for some propositional atom p. Then the argument $p \vdash \Box q$ is valid, has no undercutting arguments (since no argument has a head which entails $\neg p$), but it has a rebutting argument (namely $p \vdash \Box \neg q$).

However, we can, using our dictionary, provide conditions under which analogues of these inclusions are true. First some notation for classes of *propositions*:

Definition 3. We define the following classes of propositions:

B2 The class of propositions for which $\nvdash \Box \neg s$.

B3 The class of propositions s for which $\vdash \Diamond s$.

B4 The class of propositions s for which $\vdash \Diamond \Box s$.

B5 The class of propositions s for which $\vdash \Box s$.

Our dictionary can be restated as follows:

Proposition 8. *For $i = 2, \ldots, 5$, a proposition is in **B**i iff there is an argument for s which is in **A**i.*

We can now show:

Proposition 9. *1. **B5** $\subseteq$ **B4** in all cases*

*2. **B4** $\subseteq$ **B3** if $\Box s \vdash s$ for all s.*

Proof. We get both of these assertions by the valid inference

$$\frac{X \vdash Y}{\Diamond X \vdash \Diamond Y.}$$

$\Box$

There seem to be no significant relations between **B**2 and **B**3.

References

[Elvang-Gøransson *et al.*, 1993] Morten Elvang-Gøransson, Paul J. Krause, and John Fox. Dialectical reasoning with inconsistent information. In *Uncertainty in Artificial Intelligence: Proceedings of the Ninth Conference*, San Mateo, CA, 1993. Morgan Kaufmann.

[Krause *et al.*, 1993] S. Krause, M. Nirkhe, and K. Sycara. Reaching agreements through argumentation: A logical model (preliminary report). In *Proceedings of the Workshop on Distributed Artificial Intelligence*, 1993.

[Krause *et al.*, 1995] Paul Krause, Simon Ambler, Morten Elvan-Gøransson, and John Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11:113–131, 1995.

[Lifschitz, 1997] Vladimir Lifschitz. On the logic of causal explanation. *Artificial Intelligence*, 96:451–465, 1997.

[McCain and Turner, 1997] Norman McCain and Hudson Turner. Causal theories of action and change. In *Proceedings of AAAI97*, pages 460–465, Providence, RI, 1997. AAAI.

[Naoyuki and Takata, 2002] Nide Naoyuki and Shiro Takata. Deduction systems for BDI logics using sequent calculus. In *AAMAS 02*, Bologna, Italy, 2002. ACM.

[Parsons and Jennings, 1996] Simon Parsons and N. R. Jennings. Negotiation through argumentation: A preliminary report. In *Proceedings of the International Conference on Multi Agent Systems (ICMAS'96)*, Kyoto, Japan, 1996.

[Parsons *et al.*, 1998] Simon Parsons, Carles Sierra, and Nick Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292, 1998.

[Rao and Georgeff, 1991] Anand S. Rao and Michael P. Georgeff. Modeling rational agents within a BDI architecture. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–384, 1991.

[Rao and Georgeff, 1995a] Anand S. Rao and Michael P. Georgeff. BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems*, 1995.

[Rao and Georgeff, 1995b] Anand S. Rao and Michael P. Georgeff. Formal models and decision procedures for multi-agent systems. Technical Report Technical Note 61, Australian Artificial Intelligence Institute, 171 La Trobe Street, Melbourne, Australia, 1995.

[Schroeder-Heister, 1992] Peter Schroeder-Heister. Cut-elimination in logics with definitional reflection. In D. Pearce and H. Wansing, editors, *Nonclassical Logics and Information Processing*, volume 619 of *Lecture Notes in Artificial Intelligence*, pages 146–171, Berlin, etc., 1992. Springer-Verlag. International Workshop, Berlin 1990.

[White, 2002] G. Graham White. A modal formulation of McCain and Turner's theory of causal reasoning. In *Proceedings of JELIA2002*. Springer, 2002.

[White, 2003] G. Graham White. Interpolation for McCain-Turner causal theories. In *Proceedings of NRAC'03*, 2003.