Forthcoming in *Synthese*

A Positive Information Logic for Inferential Information[*]

Abstract

Performing an inference involves irreducibly dynamic cognitive procedures. The article proposes that a non-associative information frame, corresponding to a residuated pogroupoid, underpins the information structure involved. The argument proceeds by expounding the informational turn in logic, before outlining the cognitive actions at work in deductive inference. The structural rules of Weakening, Contraction, Commutation, and Association are rejected on the grounds that they cause us to lose track of the information flow in inferential procedures. By taking the operation of information application as the primary operation, the fusion connective is retained, with commutative failure generating a double implication. The other connectives are rejected.

KEYWORDS: information, inference, residuals, pogroupoids, structural rules

---

[*] A version of this chapter is forthcoming in *Synthese*, special section: *Knowledge, Rationality, and Action (KRA)*. This research was made possible by the generous support of the Harold Hyam Wingate Foundation. I am greatly indebted to everyone present at PIL-07, with special thanks due to all participants, especially Samson Abramsky, Patrick Allo, Alexandru Baltag, Johan van Benthem, Bob Coecke, Ed Mares, Mehrnoosh Sadrzadeh, and Sonja Smetts for their generous feedback. I am also grateful to Greg Restall, Tim Williamson, and Elia Zardini for setting me straight on more than one occasion. Any remaining errors remain my own.

1. Introduction

The informational turn in logic has resulted in logic being understood as essentially information-based. The points in semantic structures are understood as information states, with propositions being understood as sets of information states. These information states may be combined in such a way that the resulting sets of information states may be either inconsistent, incomplete, or both. Hence the informational turn extends the classical "consistent and complete" condition on points of evaluation. Accordingly, the turn has arisen naturally from research on substructural logics. The turn has also led to the logical analysis of information itself.

The starting point is as follows: We take a set S of information states, and a partial order $\sqsubseteq$ on S. Where x, $y \in S$, $x \sqsubseteq y$ is read as "the information in x is contained in the information in y", or "the information in y is a development of the information in x" and so forth. Accordingly, we have it that the heredity, or monotonicity condition: for all A, if $x \Vdash A$ and $x \sqsubseteq y$, then $y \Vdash A$, holds. '$\Vdash$' is the model-theoretic evaluation relation. In the present informational context, we may read $x \Vdash A$ as "x carries the information that A", or "x supports A". In a model $\mathbf{M} \coloneqq (S, \sqsubseteq, V)$, where V is a valuation, $x \Vdash A$ *iff* $x \in V(A)$. We may think of $x \Vdash A$ as a subtler analogue of the "makes true" relation of classical model theory. The difference is that our information states may be incomplete and/or inconsistent. There is no constraint on information states being consistent and complete, so we might not want to commit to something as strong as "makes true" in the case where x is inconsistent (baring a dialethic paraconsistentism that is). Similarly, if x is incomplete it does not follow from this alone that x will make $A \vee \neg A$ false, it will merely be the case that x does not carry the information that $A \vee \neg A$. The information ordering captured by the monotonicity condition leads immediately to ordered structures. Accordingly, the semantic structures so understood have been analysed in terms of both algebraic and frame semantics.[1]

The resources of such extra perspectives allow greater insight into the nature of the structure under investigation. As such, they become invaluable when it comes to evaluating a particular structure's suitability with respect to some intended application or other. The intended application here is information flow itself. However, we should resist the pull of an informational monism, initially tempting though it may be. Claude Shannon, the originator of the first formal theory of information, was dismissive of the plausibility of there ever being a single precisified notion that that would capture all of the subtleties at work.[2] This has been born out by the large array of logics of information present in the literature, few of which are in competition with each other.[3] As such, informational pluralism is our starting point.[4]

Despite being a fringe position in many philosophical debates, pluralism takes centre stage in research into logics of information. Shannon's insight into the ubiquity

---

[1] See Dunn and Hardegree (2001), Galatos *et. al.* (2007), Paoli (2002), and Restall (2000). This is not to suggest that only research explicitly within the informational turn examines semantic structures both algebraically and via frames. For just one example of this, see Blackburn *et. al.* (2001).

[2] See Shannon (1950, 80), and see Sequoiah-Grayson (2007) for a defense of Shannon's position on this point, as well as an analysis of what can go wrong when it is not adopted.

[3] See Abramsky (*forthcoming*), Barwise and Seligman (1993), Baltag and Sadrzadeh (2005), Coecke and Martin (2002), van Eijck and Visser (eds.) (1994), Floridi (2003), Hintikka (1970), Jago (2006), Mares (1997), Restall (1994, 2000), and Sequoiah-Grayson (2006). For an in-depth survey of many approaches to logics and information, see van Benthem and Martinez (*forthcoming*), and for a collection of recent articles on many aspects of the philosophy of information, see van Benthem and Adriaans (*forthcoming*).

[4] For an extended exposition and defense of this starting point, see Allo (2007).

of the notion notwithstanding, many logicians working on logics of information have backgrounds in theoretical computer science, where a logical pluralism has been the norm for decades. Witness van Benthem's central qualification that his system of dynamic logic of information processing, although unified in that it presents a single basic picture or framework, "does not stand or fall with the adoption of some unique preferred "base calculus" or information-processing orientated logic" (1995, 187). In a similar spirit, Wansing notes "From the point of view of the fine structure of information processing it becomes clear that the general aim of our investigation cannot be a single formal system, being the one and only logic of information structures" (1991, 18).

Different applications, in this case different domains of information processing, will guide us towards particular families of logics. The application in question is the information flow between different doxastic states of an agent as the agent reasons deductively. This application brings with it cognitive-psychological notions that are, from traditional philosophical perspectives since Frege, thought best kept at arms length. That the information gain from inference must be cognitive-psychological in some sense seems clear enough. It was certainly clear enough, for instance, for Bar-Hillel and Carnap, the originators of the theory of *classical semantic information*. In their (1952, 223), they note that the information yield of logical truths may be "rather high" for resource bounded agents. They explicitly refer to the type of information at work in such a gain as "psychological information". Whether or not the information gain from inference is entirely psychological-cognitive in nature is a contested issue. Although a detailed examination of the related controversies would take us too far afield, we will take a very brief look.

Hintikka (1973, 222-4) famously referred to the positivist position expressed by

Bar-Hillel and Carnap as reducing logic to "merely psychological conditioning, some sort of intellectual psychoanalysis". Quite apart from the failure of Hintikka's positive contribution to the issue, for which see Sequoiah-Grayson (2008a), Hintikka's negative contribution goes by too briefly. We may fix the psychological-cognitive condition more firmly via a *criterion for non-psychological information*: an information-gain is non-psychological *iff* were the agent to be logically omniscient, then the gain would still occur. This is precise enough to get us going. A consequence of precisifying things in such a manner is that it then becomes very difficult to see just how the condition might be satisfied. Sequoiah-Grayson (2008b) shows that a useful notion of *metasemantic information* satisfies this condition in at most those scenarios where the agent is learning logic. However, in the case where logic is, once learned, being used, then even metasemantic information will fail to satisfy it. The task of capturing, or even making sense of information gain in the "doing logic" scenario such that the criterion is satisfied is left to those who wish to satisfy it. Admittedly, the desire to satisfy it is a compelling one. As such, something must be said with regards to the natural aversion to understanding the information gain from inference as being constitutively psychological in character.

Dummett (1991, p. 195) captures nicely the reaction when he writes:

> If that were correct, all that deductive inference could accomplish would be to render explicit knowledge that we already possessed: mathematics would be merely a matter of getting things down on paper, since, as soon as we had acknowledged the truth of the axioms of a mathematical theory, we should thereby know all the theorems. Obviously, this is nonsense: deductive inference has here been justified at the expense of its power to extend our knowledge and hence of any genuine utility.

Put in precisely this way, it is indeed nonsense. However, putting things in precisely this way is highly misleading, if not false. Inference does not merely render explicit knowledge that we already possess. Rather, it renders explicit knowledge that we already have *access* to, and having access to something is very different from having actually accessed it. It is not true that understanding the information gain from inference as being constitutively psychological in character implies that "as soon as we had acknowledged the truth of the axioms of a mathematical theory, we should thereby know all the theorems". Instead, it implies that as soon we acknowledge the truth of the axioms, we thereby satisfy a necessary condition on our ability to *gain* knowledge of the theorems. Another necessary condition on our gaining knowledge of the theorems is that we submit our knowledge of the axioms (or natural deduction rules or whatever) to our faculties of reason. This is not to say that the content of mathematical and logical propositions is psychological. Rather, it is to say that there is a process involved in getting the information in the premises to generate the information in the conclusion. This process is a cognitive/psychological one, and requires that we *compose* the information in the premises in the correct manner. Just what the correct manner is, and how we may capture such correctness conditions in a fully-fledged informational way is something that will be held off until §2 below. For now it may left as an intuitive notion, and is presumably robust enough for us to move on.

Recognising this psychological aspect of the information gain from inference demands that we also recognise the dynamic aspect that is involved. This is due to the fact that performing an inference involves just that, a *performance.* The performance is of a procedure, which in the case is a cognitive action. Moves towards the recognition of dynamic operations in areas traditionally understood in uniquely static terms are not restricted to inferential phenomena. For a linguistic example, note that van Benthem (1995) states "In the final analysis, it seems to me now, the crucial issue is not to 'break the syntactic code' of natural languages, but rather to understand the cognitive functioning of the human mind" (ix). And later, "[I]t turns out that, in particular, the Lambek Calculus itself permits of procedural re-interpretation, and thus, categorical calculi may turn out to describe cognitive procedures just as much as the syntactic or semantic structures which provided their original motivation" (186). The analogy with inferential information runs deeper than it might first appear. The moral of this essay is that the structures underlying the Lambek Calculi, and in particular the *non-associative* Lambek Calculus (**NL**) captures important properties at work in the information update resulting from inferential procedures.

More specifically, the application in question is the *positive* information flow between different doxastic states of an agent as the agent reasons deductively. We engage in deductive inference because such procedures are informative for us. This is to say that we deduce in order to gain new information. But what properties does the information gain in this context possess, and what behaviour does it exhibit? In frame semantics terms, what structural rules do we want to include on our information frame $\langle S, \sqsubseteq \rangle$?

This is a more specific task than specifying the properties of information flow in general. Deductive inferential procedures operate on information pieces that are both necessary and *a priori*. By contrast, most existing theories of information flow, Restall (1994), Barwise and Seligman (1997), van Ditmarsch *et. al.* (2008) are concerned with the information flow between different points in distributed or multi-agent systems. In such systems, the "signals" or "information pieces" or "propositions" are usually contingent, or at least *a posteriori*. These information types bear important relations to each other, as we often reason on the basis of observational information. Indeed, our generating inferential information on the basis of observational information probably underpins most, if not all, of the reasoning involved in the natural sciences. Unifying these different systems of information flow, as well as exploring the issues involved in the relationship itself, is one of the major tasks for modelling information flow.[5] One of the prior tasks is getting clearer on the behaviour of inferential information, and that is the task undertaken here.

The obvious question then, is "which logic is fit for purpose?" Obvious or not, it is slightly misplaced. Adopting a *strong* logical pluralism, then not only will the logic required be contingent on the task at hand, there will also be more than one justifiable choice. Moreover, putting the question in precisely these terms is dangerously misleading. It is misleading because it might be read as *syntactically-biased* such that we are looking for a particular proof-theory whereby the "correct" list of axioms or derivable sentences are identified, and the semantics is part of the service-industry who's duty it is to keep the proof-theory running smoothly. By contrast, the approach here inverts this traditional picture; we have proof-theory in the service of semantics. There is an information-structure underpinning inferential information, and this structure will permit of certain procedural operations. Deductive inference is ubiquitous enough that such procedural operations and structures deserve study in their own right. To this end, proof-theory will be used to assist in the identification of

[5] One way in which unification might be carried out is via appeal to *dynamic epistemic logic* (DEL). Instead of multi-agent systems with information flowing between agents, we would have a single agent system with information flowing between different states of the agent. Different versions of DEL have been constructed with various logics, from coarse-grained systems such as S5 to more expressive varieties of linear logic. It would be a worthy task to investigate a hierarchy of dynamic epistemic logics in terms of their expressive capabilities and the reasoning contexts that they capture.

this information-structure, as well as the relevant procedural operation. Our perspective on the proof theory is one such that what we are looking for is the correct type of operation of information-application that serves an information processing mechanism in the appropriate way for the buttressing of the information-structure that underpins the phenomena (inferential information). As it stands this is somewhat vague, but it will have to remain so until it is cashed out in the sections below. It is at least enough to get us going for now. The information-application operation will take centre stage in the context we are working in. The toil is in working out just what sort of information-application operation we want. Or to put this more accurately, given that we *do* want the operation, what properties do we want it to take? In this sense, even our proof-theoretical concerns take on a robustly procedural-semantics hue. The information-structure and information-application operation ultimately identified will turn out to be the same that are identified by the non-associative Lambek calculus **NL**, albeit under an informational and procedural interpretation. We will simplify the task as much as possible by restricting ourselves to positive information.[6]

We begin in §2 by examining the effect of the presence of the structural rules of Weakening, Contraction, Commutation, and Association. All four are rejected, hence the structure we arrive at is a *non-associative information frame*. In §3, we move on to examining the effect of extensional conjunction, extensional disjunction, fusion, fission, material implication, and intensional implication. Only fusion and intensional implication will be kept. This is not arbitrary, and will be seen to be a direct

[6] How it is that we might understanding negation post the informational-turn, that is, how we might understand *negative information* (Deletion? Retraction? Failure? Informational *debt*? Procedural *prohibition*?) is a hard issue, and one that deserves independent treatment. For a suggestion along the lines of negative information as procedural prohibition, see Sequoiah-Grayson (2009).

consequence of the procedural semantics that underpin the frame conditions of these connectives. The failure of Commutation for the frame will split intensional implication and give us a double implication. The fllowing section is largely proof-theoretic. As we will see however, there is an information-based procedural interpretation of the proof theory that brings it directly into line with our information-based procedural concerns.

2. Inferential Information and Structural Rules

In this section we introduce structures, substructures, as well as introduce the information-application operation for both structures and formulae. The structural rules of Weakening, Contraction, Commutation, Association, are introduce and rejected, delivering a non-associative information frame. A robustly information-based procedural interpretation of the proof theory is developed, with specific properties of inferential information being preserved.

Structures X, Y…, are made up out of formulas A, B…, concatenated via the binary punctuation mark ';'. The substructures of a structure are as fine-grained as the formula of the structure, but not as fine-grained as the subformula. For example, take $(A \rightarrow B); (A; C)$. Here, the structures $A \rightarrow B$, $A; C$, A, C, and $(A \rightarrow B); (A; C)$ are substructures of the structure $(A \rightarrow B); (A; C)$. Note that every structure is a substructure of itself. Importantly, B is not a substructure of $(A \rightarrow B); (A; C)$, rather it is a subformula of a substructure of $(A \rightarrow B); (A; C)$. An important note: when we are operating at the level of formulae as opposed to structures, we will replace the structural level ';' via the formula-level analogue to ';', fusion, '$\otimes$'. Fusion inherits its properties directly from the semicolon. The semicolon may be read in differing ways depending on the context of application.

This is an important point. In the most general terms, $X; Y$ is read as "X taken together with Y", where this "taking together" is akin to set-union such that $X; Y$ is no different to $Y; X$, and $X; X$ is no different to X and so on. We are working within the context of information structures, so immediately we know that we are reading $X; Y$ as 'the result of the application of the information in X to the information in Y'. In this context we then ask what properties it is that the semicolon may take. What does "applied to" mean exactly? For just one example, is the information resulting from the application of the information in X to the information in Y the same information as that resulting from the application of the information in Y to the information in X? That is, is it the case that $X; Y = Y; X$? Given logical pluralism, the answer to this question will depend on the information-type under consideration. The information type under consideration here is (deductive) inferential information. The purpose of the present section is to specify which properties operate in this context. The first step is to make the notion of *information application* clear.

Information application is a restricted notion of information composition. The difference is that raw composition, at least as it stands, is ambiguous between the concatenation of bodies of information on the one hand, and the *interaction* between bodies of information on the other. Another way of putting this is that it amounts to the difference between a mere *list* (ordered or unordered, it does not matter at this stage) of bodies of information on the one hand, and the actual *integration* of the bodies of information in that list on the other. It is best to bring this out with a concrete example (one that will in fact play an important role later in this section): Suppose that we have the information that A, and the information that B conditional on the information that A. That is, we have the information that $A \to B$. In this case, the application of the information in A to the information in $A \rightarrow B$ will result in the information that B. Consider another case (unrelated to the first), where in this case we have the information that A as before, but we now have the information that C conditional on the information that B. That is, we have the information that $B \rightarrow C$. Let us also suppose that $A \neq B$. In this case, what is the information that results from the application of the information in A to the information in $B \rightarrow C$? The answer is straightforwardly that you get absolutely nothing. This is only straightforward on the basis of our restricting *composition* to *application*, and it is precisely for this reason why it is that this restriction is so crucial. If we were to sit with the notion of composition, then the result of composing the information in A to the information in $B \rightarrow C$ is not *straightforwardly* "nothing". Instead, it is left open that we might get the information in A collected with the information in $B \rightarrow C$, where this "collected with" is something like conjunction (and arguably just *is* conjunction). But we are not concerned with the mere collecting together of bodies of information. We are concerned with the *integration* of bodies of information, such that this integration may or may not result in new information. In short, we are concerned with information *processing*. This is brought out neatly via *structural rules*.

The structural rule notation is read as follows: $X \Leftarrow Y$ means that anything that follows from processing the structure X, also follows from processing the structure Y. In the present context, we may think of structures as structured bodies of information. We follow Gabbay (1996, 423) and refer to structures under an informational reading as *data structures*. In this case, we cash out a structural rule as stating that any piece of information A that follows from processing the data structure X also follows from processing the data structure Y. We can make this explicit by writing out a generalised structural rule in long-hand form:

$$\frac{X \vdash A}{Y \vdash A}$$

We only do this once in order to make the finer details clear, and shall continue to use the short-hand form. Note that a structural rule consists of two inferences, one on the top and one on the bottom (or one on the left hand side and one on the right hand side in the short hand form). For a *sequent* $X \vdash A$, we again follow Gabbay's lead and read it as stating something like “the piece of information expressed by A is contained in the data structure expressed by X” (*ibid*). Wansing makes a similar point concerning “deductive information” (whereby he just means what we mean by ‘inferential information’) when he states that deductive information processing takes the premises as databases and the consequence relation ‘$\vdash$’ as the information processing mechanism (*ibid*, 16). This is stricter than the mere “informational containment” of Gabbay, although Gabbay is careful to emphasise that “containment” is a first pass at a more refined understanding of the properties of information flow from data structures (*op. cit.*). The structural rules will specify the ways in which we may fix properties of the information flow expressed by the turnstile. This will, as Gabbay notes, depend on the structure of the data and the permitted ways of manipulating it. The structural rules tell us what these manipulations are. Which of them that we find permissible will depend on the type of data-structures under investigation (in our case premises) and the processing on them that takes place (in our case deductive inference as an instance of procedural reasoning).

The various properties are captured via structural rules. We may be immediately confident that we want to disallow both Weakening and Commuted Weakening (K and K′) and Strong Contraction (W):

$$(K): X \Leftarrow X; Y$$

$$(K'): X \Leftarrow Y; X$$

$$(W): (X; Y); Y \Leftarrow X; Y$$

The reasoning is straightforward. If K or K′ were to be present, then we would loose track of what sets of information states were in fact used in a deduction, as Y may contain arbitrary formulae. Disallowing the weakening rules is the mark of *relevant logics*. Unsurprisingly, relevant logics were developed in order to provide a realistic analysis of the conditional, and to circumvent the paradoxes of material implication.[7] Irrespectively of the success or otherwise of relevant logics accounting for natural language uses of the material conditional, we want to follow the relevantists insofar as the rejection of weakening is concerned. Weakening allows us to add irrelevant premises to an inference. In classical logic this is permissible, as there is nothing that one can do to a classically valid argument to make it classically invalid. However, if we are constructing a metalogic in order to capture the information flow involved in deductive inference (and we are), then weakening would allow us to destroy the bodies of information that were *actually used* in the processing for the conclusion. If $X \vdash A$, then it may not be true that $X; Y \vdash A$ in the strict sense of "following from" as the data structure Y, being arbitrary, has nothing to do with the processing that generates A.

If W were to be present, then although we would not lose track of what data structures were used in a deduction, we would lose track of how many *times* a data

[7] Relevant logics have received a large amount of attention over the last several decades. For up to date and staunchly philosophically motivated developments, see Mares (2004), and Priest (2001, ch. 9-10).

structure was used in a deduction. This is to say that contraction leaves us blind to *resource-use*, and resource-sensitivity is something that we would like to preserve in our context. There is a difference between the generation of A from one instance of processing on a data structure Y and, say, hundreds of instances of processing on Y. This difference comes down to one of sheer processing effort. By jettisoning the Weakening rules and the Contraction rules, we enter a family of logics weaker still than relevant logics, *linear logics*.[8]

Linear Logics were developed to capture computational complexity and cost.[9] At a minimum, they allow commutation and association.[10] Commutation is marked by the strong commutation rule (C), and association is marked by the rules of Associativity and Converse Associativity (B and B^C):

$$\text{C: } (X; Y); Z \Leftarrow (X; Z); Y$$

[8] The whole truth is a little subtler, as we also arrive at logics such as Ticket Entailment without Contraction, and the various types of Lambek Calculi among others. Neither have we yet mentioned the intensional connectives. However we are suppressing these details until the sections below. As mentioned above, it is the structure identified by the non-associative Lambek Calculi (**NL**) that will end up doing most of the labour.

[9] See Troelstra (1992) for a detailed introduction, albeit one that presupposes category theory.

[10] Again, the whole truth is a little subtler. By dropping the commutation rules we get the family of *non-commutative linear logics*. These should not be confused with a logic that rejects commutation, such as the Lambek Calculi. It is not only the absence of Weakening and Contraction that take us to linear logics, but also the presence of the exclusively linear exponentials '!' and '?'. Put extremely briefly, ! is the "obviously" or "of course" operator, such that $!A$ allows us as many instances of A as we require for some procedure, whereas ? is the "why not" operator, such that $?A$ is defined as $\neg!\neg A$. Hence, ! and ? allow us to recapture Contraction and Weakening respectively for specified fragments of the logics.

$$\text{B}: X; (Y; Z) \Leftarrow (X; Y); Z$$

$$\text{B}^{C}: (X; Y); Z \Leftarrow X; (Y; Z)$$

In the context of inferential information flow, what does Commutation amount to? In this case also, the result is harmful. This is easier to see at the level of formulae, hence we run through the example in terms of fusion ⊗. ⊗ mirrors the behaviour of the structure-level punctuation mark ‘;’ at the level of formulae. In this case $A \otimes B$ is the application of the information in the formula A to the information in the formula B. Set the following:

$$A: p,\ B: p \rightarrow q,\ C: q,\ D: q \rightarrow r,\ E: r.$$

In this case we have it both that:

$$A \otimes B \vdash C$$

and that:

$$C \otimes D \vdash E$$

hence:

$$(A \otimes B) \otimes D \vdash E$$

This last sequent follows from the previous two via a hitherto unstated rule – the *Cut* rule. Where we have $A \otimes B \vdash C$ and $C \otimes D \vdash E$, we can “cut out” the term C, getting $(A \otimes B) \otimes D \vdash E$. This is of course only an instance of the Cut rule. In its general form it is stated as follows: Where $X(Y)$ means that Y is a substructure of X, if we have it that both $X \vdash A$ and $Y(A) \vdash B$, then we have it via Cut that $Y(X) \vdash B$.

Having cut away C then, we can, via Commutation, now get the following:

$$(A \otimes D) \otimes B \vdash E.$$

This is a disaster. The original conclusion states the following: by applying the information that results from the application of the information in A to the information in B, to the information in D, we get the information in E. Given the setup, this makes perfect sense. Post the application of C however, we have something that states the following: by applying the information that results from the application of the information in A to the information in D, to the information in B, we get the information in E. Given the setup, this makes no sense at all. The result of applying the information in A to the information in D is nothing such that were it to applied to the information in B we would the information in E. By permitting Commutation, we "come off the trail" taken by the information flow through the proof.

What does Association amount to? Given the setup involved, the result appears to be harmless. We have it that:

$$A \otimes (B \otimes D) \vdash E.$$

This is a result of an application of cutting out F, since:

$$B \otimes D \vdash F, \text{ and } A \otimes F \vdash E.$$

From $A \otimes (B \otimes D) \vdash E$ via Association, we can get:

$$(A \otimes B) \otimes D \vdash E$$

and this is as we would expect having cut out C, as straightforwardly:

$$A \otimes B \vdash C$$

and

$$C \otimes D \vdash E$$

as desired.

Converse Associativity (B^C) simply takes us back in the other direction. We have it this time that:

$$(A \otimes B) \otimes D \vdash E.$$

This is a result of cutting C, since:

$$A \otimes B \vdash C, \text{ and } C \otimes D \vdash E.$$

From here via B^C, we can now get:

$$A \otimes (B \otimes D) \vdash E$$

and this is what we would expect. We already know from the reasoning above with respect to B that $B \otimes D \vdash F$ and that $A \otimes F \vdash E$. However, this appearance of harmlessness is mere appearance and nothing more. It is an artifact of the setup, and some simple changes show that both B and B^c have the same disruptive effect as Commutation. To see this we set the following:

$$A^*\text{: } p \rightarrow q,\ B^*\text{: } r \rightarrow p,\ C^*\text{: } r,\ D^*\text{: } p,\ E^*\text{: } q.$$

In this case we have it that:

$$(A^* \otimes (C^* \otimes B^*)) \vdash E^*$$

This is a result of cutting D^*, since:

$$C^* \otimes B^* \vdash D^*, \text{ and } A^* \otimes D^* \vdash E^*.$$

From $(A^* \otimes (C^* \otimes B^*)) \vdash E^*$ via B we can get the following:

$$((A^* \otimes C^*) \otimes B^*) \vdash E^*$$

and this is just the type of disaster that we are looking for. Via B, we have something that states the following: by applying the information that results from the application of the information in A^* to the information in C^*, to the information in B^*, we get the information in E^*. Given the new setup, this makes no sense at all. The result of applying the information in A^* to the information in C^* is nothing such that were it to applied to the information in B^* we would get the information in E^*. We can now see that by permitting B, we "come off the trail" taken by the information flow through the proof in exactly the way that we would were we to permit C. A similar result is got for B^C if we set the following:

$$A^{**}\colon r \rightarrow p,\ B^{**}\colon p \rightarrow q,\ C^{**}\colon r,\ D^{**}\colon p,\ E^{**}\colon q.$$

We have it this time that:

$$((A^{**} \otimes C^{**}) \otimes B^{**}) \vdash E^{**}$$

This is a result of cutting D^{**}, since:

$$A^{**} \otimes C^{**} \vdash D^{**}, \text{ and } D^{**} \otimes B^{**} \vdash E^{**}$$

From $((A^{**} \otimes C^{**}) \otimes B^{**}) \vdash E^{**}$ via B^C we can get the following:

$$(A^{**} \otimes (C^{**} \otimes B^{**})) \vdash E^{**}$$

and again we arrive at the type of disaster we are looking for (or not looking for) as no information results from applying the information in C^{**} to the information in B^{**} such that if the information in A^{**} is applied to it, we will get the information in E^{**}. B^C causes us to come off the trail of the information flow through the inference in the same way as does B and C.

This is not all that can be said about structural rules and information flow. On the contrary, given the informational turn, informational interpretations of various substructural logics are likely to be a fruitful area of research for some time to come. Given informational pluralism, we should expect that many different information structures, specified by different information frames, and hence by different collections of structural rules, will have useful applications.[11] Given strong informational pluralism, we should expect that more than one information logic will have a useful role to play in precisifying the behaviour of a particular information type. Nevertheless, the choice of structural rules dealt with here has not been arbitrary. The Weakening, Contraction, and Commutation rules have been at the forefront of concerns regarding information identification, tracking, order, and flow in general. Given the context specified by inferential information, our minimal information frame $\langle S, \sqsubseteq \rangle$ has been extended to a *non-associative information frame*. The rejection of associativity requires a few further words.

By rejecting commutation, we are preserving the order of the data structures. In other words, these structures are not combined into sets, but rather into *sequences*. Since sequences occur in a particular order, if we were to allow commutation, then the sequences themselves would be destroyed, and we would instead have "bags" or multisets (and if we had retained Contraction as well, then we would have regular sets). By rejecting association, we achieve a greater level of expressive power still. We move from pure sequences to *pairs*. Another way of thinking about this is that we have moved from merely preserving the order of the structures within the sequence, to

[11] Failing to appreciate this point will not only cause one to fail to appreciate the potential of research into information logics. It will also cause one to misunderstand much of the research into information logics currently under way. See Sequoiah-Grayson (2007).

preserving the order of the *application* of the structures (that form the sequence) to each other. Having fixed on the information frame with no structural rules, we now turn our attention to connectives.

3. Inferential Information and Connectives[12]

In this section we consider the addition of the following connectives to our non-associative information frame:

($\wedge$): Extensional conjunction

($\vee$): Extensional disjunction

[12] There are large variations in terminology here, as well as with notation. A brief, but far from complete list, is as follows: Negations of differing types are variously denoted by ‘~’, ‘¬’, ‘$^{\perp}$’, and ‘–’. *Extensional conjunction* is also referred to as *additive conjunction*, *classical disjunction*, and *Boolean conjunction*, and variously denoted by ‘∧’, ‘&’, ‘·’, and ‘⊓’. *Extensional disjunction* is also referred to as *additive disjunction*, *Boolean disjunction*, and *classical disjunction*, and variously denoted by ‘∨’, ‘⊔’, and ‘⊕’. *Material implication* is also referred to as *extensional implication*, and the *material conditional*, and variously denoted by ‘⊃’, ‘→’, ‘⇒’. *Fusion* is also referred to as *multiplicative conjunction*, *intensional conjunction*, *tensor*, and *times*, and variously denoted by ‘∘’, ‘⊗’, ‘·’, and ‘★’. *Fission* is also referred to as *multiplicative disjunction*, *intensional disjunction*, *dual of fusion*, and *par*, and variously denoted by ‘+’, ‘⊕’, and ‘⅋’. *Intensional implication* is also referred to as *relevant implication*, *linear implication*, and the *right conditional*, and the *right residual*, and variously denoted by ‘→’, ‘⇒’, ‘⊸’, ‘/’ and ‘\’. Leaving aside the well-known multitude conditional notations, note that ‘⊕’ is sometimes used to denote extensional disjunction, and other times used to denote fission. Moreover, ‘→’ and ‘⇒’ are often used to denote the consequence relation that we are denoting with ‘⊢’.

(⊃): Material implication

(⊗): Fusion

(⊕): Fission

(→): Right implication

(←): Left implication

The task of identifying the connectives relevant to our associative information frame in the context of inferential information gain is simplified by the decision to restrict our attention to the positive fragment of the logic. By leaving negations aside, we circumvent one of the most difficult, although also most interesting, issues in logics of information flow. An informational interpretation of negations is a topic unto itself, and there is not enough space here to investigate the issue.[13] Hence there will be no discussion of *falsum* 'f' or *bottom* '⊥'.[14] Of the seven connectives examined, only three of them will be retained. As was stated in §1, these will be fusion, right implication, and left implication. This cuts against the grain of many substructural logics, where the more minimal the set of structural rules, the greater the natural array of connectives.[15] Why should the present case be an exception to this general (although far from universal) moral? The answer is straightforward, although it takes some filling out: We are concerned with the information-application operation. At the

[13] See Heinrich Wansing's (1993) for a benchmark approach to the issues, and Sequoiah-Grayson (2009) for a counter approach.

[14] Relatedly, discussion of the truth constant *t* is held off until §4, where it will be discussed under its algebraic analogue, *identity*: 1.

[15] For example, consider the traditionally connective-heavy linear logics, or Ono's system of Full Lambek (**FL**) and its variants, for which see Galatos *et. al.* (*op. cit.*).

formula-level this is fusion. Fusion and right and left implication form a connective family, with fusion as the parent and right and left implication as "residuals". The three connectives interact in a manner that will be laid out below. What we will see is that they all have their frame conditions specified in terms of the same model-theoretic binary combination operator $\bullet$ on S, along with the partial order $\sqsubseteq$ on S of informational development. The form of the model-theoretic operation that will ultimately concern us is $x \bullet y \sqsubseteq z$. This is a robustly procedural operation, read as *the result of the application of the information in x to the information in y develops into the information in z.* As such, our three connectives will have a set of explicitly procedural, information-based model-theoretic conditions. Moreover, the binary combination operator has the same properties as the information-application operation, and this dynamic operation is the very one that we are concerned with. In fact the binary combination operator just *is* the information-application operation operating on information states (as opposed to fusion and the semicolon which operate on formulas and structures respectively). This should be enough for now to at least motivate the idea that fusion and right and left implication are privileged insofar as it comes to procedural information processing, so on to the details.

We begin with *extensional conjunction* ($\wedge$): It is difficult to see what, if any utility there is in considering extensional conjunction. In inference, we are concerned with the *application* of one body of information to another, not merely bodies of information being considered together. Although within formulae, and hence structures, there may of course be extensional conjunctions operating on subformulae, we are concerned here with the informational composition of formulae and structures. Extensional conjunction operates on *subformula* of substructures, but not on substructures or structures themselves.

The issue here is slightly more involved than it might first appear, as extensional conjunction reappears when we treat the metalogic that is modelling inferential information as the object language. The instance where this has already occurred is in our discussion of Cut. Taking the first occurrence as an example, from $A \otimes B \vdash C$ and $C \otimes D \vdash E$, we "cut out" the term C, getting $(A \otimes B) \otimes D \vdash E$. The "and" in "$A \otimes B \vdash C$ and $C \otimes D \vdash E$" is extensional conjunction. But this does not mean that extensional conjunction has a role to play *in* our non-associative information frame. Its role reappears when we are speaking *about* the operations on the frame. When we are cutting, the information-processing sequents are being mentioned, not used. This is no different to our natural language discussions of any logic without extensional conjunction where we are treating that logic as the object language. Putting things in an ugly yet accurate manner: extensional conjunction is present in our original object language (classical logic), not present in our meta-language (non-associative information frame), and present again in our meta-meta-language (when we speak about the meta-language).

Doing away with extensional conjunction is related to doing away with *conjunctive closure.*[16] Take some broadly epistemic operator $\Box$ and a related agent and proposition. In this case we have $\Box_{\alpha_i} A$ (an agent α_i knows/believes that A). The conjunctive closure principle is as follows: If $\Box_{\alpha_i} A$ and $\Box_{\alpha_i} B$, then $\Box_{\alpha_i} B$. It is entirely reasonable that conjunctive closure should fail in the context of inferential information. In fact, it would be odd if it were to hold. The contrast here is between *local reasoning* and *global reasoning*. We can think of global reasoning as analogous to logical omniscience. We are not considering logically omniscient agents in the inferential information gain context. Indeed, this is one way of getting at the entire

[16] Many thanks to Patrick Allo for getting me to think harder about this than I otherwise would have.

point. We can think of local reasoning as reasoning under constraints. These are the constraints imposed upon natural agents in virtue of our being resource-bound. In an act of inference, an agent is *situated in that act of inference*. An agent may well infer that *A*, and also infer that *B*. In this case, the agent may well believe that *A*, and also believe that *B*.[17] However, these two acts of inference are not identical to a single act of inference that results in *A and B*. Indeed, the agent may never infer to $A \wedge B$. The agent may never entertain the belief (and hence never know) that *A and B*. The parallel here with the reasoning concerning Cut in the paragraph above is object-language/meta-language distinction. In this case it is not true in the object language that the agent knows/believes that *A and B*, although it is true in the meta-language.

The absence of extensional conjunction trivially rules out conjunctive closure. However, even if extensional conjunction were present, we would understand conjunctive closure to fail. The failure of conjunctive closure in the inferential information context is to be expected. This expectation in only strengthened by the understanding that the informational turn has some of its roots in the *situation semantics* due in its original form to Barwise and Perry (1983). It was not long before Barwise (1993) made explicit the identity between the operations in situation semantics and the frame semantics of relevant logics. This was the first step from the original set-theoretic constructions of situation semantics to the now dominant modal understanding of such inconsistent and incomplete states as *impossible situations*. Restall (1996) and Mares (1997) brought this underlying semantic structure explicitly into the context of information flow. Sequoiah-Grayson (2006) made explicit the philosophical motivations for an understanding of such states as possibly existing

[17] The agent may also not hold such beliefs of course, should she be inferring from hypothetically posited premises for example. In this case we might say that she was *entertaining* the proposition.

doxastic states. Fagin (*et. al.*, 1995) developed a modal framework wherein local reasoning structures are used to circumvent problems of logical omniscience via their being interpreted as “frames of mind”. The local-situated/global distinction maps on to the distinction between an “internal perspective” and an “external perspective” on the information corresponding to the points of evaluation (situations) supporting propositions in a model. Levesque (1990) makes the same distinction in terms of a “subjective understanding of logic” and “an objective use of logic” (p. 266), and Buszkowski (1989) marks the distinction via “internal logic: versus “external logic”. Allo (2006) develops the internal perspective in terms of the notion of *local information*. This reasoning with respect to the internal and external perspective on extensional conjunction carries straight over to extensional *disjunction*.

With *extensional disjunction* ($\vee$) we have a near identical result: Similarly to extensional conjunction, extensional disjunction drops out of the context of inferential information, although it may of course reappear when we speak *about* components of our non-associative information frame. Similarly to extensional conjunction again, although extensional disjunction may operate on subformula of substructures, it will not operate on substructures or structures themselves.

Our reasons for rejecting *disjunctive closure* are analogous to our reasons for rejecting conjunctive closure. The disjunctive closure principle states that if $\Box_{\alpha_i} A$, then $\Box_{\alpha_i} (A \vee B)$ where B is arbitrary. Reiterating, in an act of inference, an agent is *situated in that act of inference*. An agent may well infer that A, in which case it follows that A *or* B, for any B. However, this act of inference to A is not identical to an act of inference to A *or* B. Indeed, the agent may never infer $A \vee B$. The agent may never entertain the belief (and hence never know) that A *or* B. Of course, it *is* the case that $\Box_{\alpha_i} (A \vee B)$ if we consider global reasoning, however we are by definition not.

Again, global reasoning is analogous to logical omniscience, and we are considering rationally bounded agents operating under naturalistic constraints. Namely, we are not logically omniscient.

There is a further consequence to rejecting the extensional connectives. The absence of extensional disjunction (and conjunction) means that the issues concerning the distributivity property ($A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)$) of these connectives do not arise (as they do in some quantum logics for example). Even more simple and straightforward than the task of rejecting extensional disjunction, is the task of rejecting *material implication*.

Rejecting *material implication* ($\supset$) is immediate: Material implication has no role to play in the context of inferential information gain. Using material implication to plug substructures together is for all practical purposes (or at least for ours) useless. Material implication notoriously leads to the *paradoxes of material implication*: *explosion*: $\neg p \vdash p \supset q$ and *irrelevance*: $p \vdash q \supset p$. As is well known, the intensional implication of substructural logics was originally conceived in order to circumvent the paradoxes. Material implication is rejected straightforwardly, as nothing could be more damaging to the project of modelling the information gain for an agent engaged in a situated inferential procedure.

There is a possible rejoinder here, along the lines of the following: "Look, we all know about the paradoxes, but so what? Just because material implication *allows* one to go and plug in irrelevant antecedents and so forth, one is under no compulsion to do so. As long as the material conditional is only used in a sensible way [which in this case means only linking permissible formulae] then no harm is done." This response is an instance of a more general type of response to non-classical connectives in general: "Why bother with the complexity of non-classical connectives? Why not just use the classical connectives, and classical logic for that matter, but make sure that you only use it in the manner you want to in a particular circumstance?"[18] A response along these lines requires a rejoinder larger than can be given here. However, we can still usefully identify the difference in stance that leads to it. The stance concerns the priority of model theory/semantics vs. proof theory. The traditional and still commonplace view of logic is that the proof-theory has primacy, and semantics is in its service (for completeness etc.). Proof-theory is often taught in the absence of any semantics whatsoever, with the model-theory only brought in after the proof-theory has been laid down. However, there is no *constraint* on our doing things this way. Instead, we are looking at a particular information type, *inferential information*, and looking for the semantic structure, in this case the information frame, that fits best. In this case the logical language will *follow*, as a system of information processing operators that are best suited to the task at hand.[19] One operator that easy to incorporate is *fusion*.

We already have *fusion* (⊗) on board: Fusion is as easy to accept, as is material implication to reject. Fusion simply mirrors the behaviour of the semicolon at the level of formulae. Hence, instead of the information resulting from the combination of the information contained in structures, fusion gives us the information resulting from the combination of the information contained in individual formulae. The operation of information combination, which in the case of inferential information has been refined to information application, is the basic operation on dynamic information processing. The philosophical toil is in the specification of the properties of the fusion

[18] Tim Williamson has expressed just such a concern via personal communication.

[19] See van Benthem (*forthcoming*) for a more detailed development of this methodological stance with respect to constructive information systems and intuitionist logic.

operator given the intended context of application. We know from the toil in §2 that in the context of inferential information the fusion operator is non-associative. As we also saw in §2, the move from non-commutativity to non-associativity corresponds to a move from pure sequences to pairs. We will shortly see how this is a reflection of robustly dynamic semantic operation when we examine double implication and the model-theoretic operation underpinning both fusion and double implication. Briefly however, things are slightly more involved when we consider fusion's close relative, *fission*.

Fission ($\oplus$) is the dual of fusion: In classical linear logic, fission is definable as follows: $A \oplus B = (A^{\perp} \otimes B^{\perp})^{\perp}$, where '$^{\perp}$' is linear negation. Linear negation is involutive, that is $A^{\perp\perp} = A$, and hence a duality operator. Were we not to be restricting ourselves to the positive fragment of the logic, and were allowing linear negation and hence duality, then we would have fission straightforwardly. We would also have the following De Morgan duality: $(A \otimes B)^{\perp} = (A^{\perp} \oplus B^{\perp})$, $(A \oplus B)^{\perp} = (A^{\perp} \otimes B^{\perp})$. Interpreting duality/linear negation in the context of inferential information would likely be a rewarding task. However, it is a task that awaits future research on negative information. Whilst restricted to the positive fragment, there is no way to make sense of the dual of fusion *qua* dual. It is of course entirely permissible that we take fission as a primitive. However, in this case it is difficult to arrive at any natural interpretation of fission in the context of inferential information.[20] What is not difficult however, is accepting *intensional implication*.

Intensional implication ($\rightarrow$) is obviously relevant in the context of inferential information. This is because we can give intensional implication an explicitly

[20] van Benthem (1995, §VI) makes use of the notion of *informational addition*, but extending this to inferential information is not easy.

procedural definition in terms of information processing. To see how this is done, we need to introduce the binary combination operator on information states, $\bullet$. $\bullet$ stands between states $x, y, \ldots \in S$ in the same way as does fusion between formulae and the semicolon between structures. Hence '$x \bullet y$' is read as *the result of applying the information in state x applied to the information in state y*, with all the caveats that go along with this as for the semicolon etc. We can combine $\bullet$ with $\sqsubseteq$ in order to give the frame condition for intensional implication:

$$x \Vdash A \rightarrow B \text{ iff for each } y, z \text{ such that } x \bullet y \sqsubseteq z, \text{ if } y \Vdash A \text{ then } z \Vdash B.$$

We read $x \bullet y \sqsubseteq z$ as something like *the result of the application of the information in x to the information in y develops into the information in z.* $\bullet$ exhibits the same properties on a frame as are expressed by the semicolon. This is just to say that if the frame is, for example, contractive, then so is the composition operator and vice-versa. Our information frame is (among other things) non-commutative. Hence $x \bullet y \neq y \bullet x$. In this case, we can give the frame condition for another implication, this time one where $y \bullet x \sqsubseteq z$:

$$x \Vdash B \leftarrow A \text{ iff for each } y, z \text{ such that } y \bullet x \sqsubseteq z, \text{ if } y \Vdash A \text{ then } z \Vdash B.$$

Hence, non-commutation means that the regular intensional implication splits into a double implication: *right implication* ($\rightarrow$) and *left implication* ($\leftarrow$). Right and left implications simply tell us in what order the procedure must be executed in order to generate the information state carrying the information conditional on the antecedent.

Fusion and double implication are not independent. The manner in which they interact is as follows:

$$A \otimes B \vdash C \textit{ iff } B \vdash A \rightarrow C.$$

$$A \otimes B \vdash C \textit{ iff } A \vdash C \leftarrow B.$$

$\rightarrow$ and $\leftarrow$ are *residuals* of $\otimes$. As should hopefully be obvious by this stage, these two conditions are order-sensitive dynamic information processing versions of the Deduction Theorem. They make sense. If we have the information that C as a result of applying the information in A to the information that B, then from the information that B we have the information that C conditional on the information that A, and from the information that A we have the information that C conditional on the information that B, and vice versa.

We now know where we stand. The inferential information structure $\mathbf{I} \coloneqq \langle S, \sqsubseteq, \bullet, \otimes, \rightarrow, \leftarrow\rangle$ is a double implicational non-associative information frame. Our model $\mathbf{M} \coloneqq \langle \mathbf{I}, \Vdash\rangle$ such that $\Vdash$ is the evaluation relation obeying the aforementioned heredity condition:

$$\text{for all } A, \text{ if } x \Vdash A \text{ and } x \sqsubseteq y, \text{ then } y \Vdash A,$$

and also obeying the following conditions for each of our connectives:

$$x \Vdash A \otimes B \textit{ iff } \text{for some } y, z \in \mathbf{I} \text{ s.t. } y \bullet z \sqsubseteq x, y \Vdash A \text{ and } z \Vdash B$$

$$x \Vdash A \rightarrow B \textit{ iff } \text{for all } y, z \in \mathbf{I} \text{ s.t. } x \bullet y \sqsubseteq z \text{ , if } y \Vdash A \text{ then } z \Vdash B$$

$$x \Vdash B \leftarrow A \textit{ iff } \text{for all } y, z \in \mathbf{I} \text{ s.t. } y \bullet x \sqsubseteq z, \text{ if } y \Vdash A \text{ then } z \Vdash B.$$

We can now see how it is that the operation of information-application underpins our connective family. The binary composition operator $\bullet$ along with the binary partial order of informational development $\sqsubseteq$, comprise an explicitly information-based dynamic procedural operation on information states. This operation involves developing a state of information via the application of other states of information to each other in particular ways. We should not think of this model-theoretic operation is "propping up" the operations involving the connectives, but rather think of the connectives as reflecting, or indicating, particular dynamic operations and conditions on information states. The operational semantics, under their informational interpretation, take prime position. The connectives serve to indicate which of these operations, or strictly, which version of the application operation, is active. The semantic hue evident throughout our discussion of structural rules and connectives has now been "traced to its source". This fact also explains in greater detail why we have accepted and rejected the connectives that we have. All along, we have been concerned in one way or another with dynamic procedural operations on an information structure. The evaluation conditions for fusion and double implication are constitutively dynamic. By contrast, extensional conjunction and extensional disjunction do not have dynamic evaluation conditions. Instead they have the usual static conditions as follows:

$$x \Vdash A \wedge B \textit{ iff } x \Vdash A \text{ and } x \Vdash B$$

$$x \Vdash A \vee B \textit{ iff } x \Vdash A \text{ or } x \Vdash B$$

Similarly with material implication:

$$x \Vdash A \supset B \textit{ iff } x \nVdash A \text{ or } x \Vdash B$$

These connectives are simply irrelevant with respect to our particular dynamic/procedural concerns regarding inferential information. The rejection of fission is based on different considerations, primarily to do with our restricting ourselves to positive information as expounded above. We can now see *why* as opposed to merely *how* it is that we have **I**. Interestingly, **I** is simply the non-associative Lambek calculus **NL** under an information-based procedural interpretation. What exactly, if anything, this reveals about the relationship between the processes underlying inferential information on the one hand and the structural aspects of natural language on the other it is too early to say. For our purposes it is significant enough that **NL** allows for a procedural re-interpretation that, when interpreted informationally, gives us the information-application operation conditions for inferential information.

We are now in a position to briefly examine our inferential information structure in terms of its corresponding algebraic structure. This is a precondition on our being able to shed some extra light on the issue surrounding a further connective, the truth constant t.

4. Inferential Information and Algebraic Structures

In the most general terms, an algebra is simply a set, together with some operations on this set. Algebras are suitably abstract, and therein lie their appeal: we may concentrate on structural properties without getting bogged down in syntactic details. The structure that we are concerned with is a propositional structure, consisting of points a, $b\ldots \in P$. These points will correspond to the information states $s \in S$. The partial order relation $\leq$ on P may be read as *entailment*, or *deducibility*, or *consequence*. The entailment relation on a propositional structure is a partial order: $a \leq a$ for all $a \in P$, if $a \leq b$ and $b \leq c$ then $a \leq c$, and if $a \leq b$ and $b \leq a$ then $a = b$. The entailment ordering $\leq$ is different to the information ordering $\sqsubseteq$. They are in fact *duals*. If $x \sqsubseteq y$, then x carries less (or at least no more) information than y, and x will support *less* (or at least no more) formulae than does y. However, in a *propositional* structure $a \leq b$, a will entail more propositions than b. As Restall (2000, 238) puts it, "[i]f $x \sqsubseteq y$, then x is "less defined" than y. If $a \leq b$, then a is true "less often" than b." They are duals in the sense that the further *up* the information order you go, the greater the number of propositions that are *supported*, and the further *down* the entailment ordering you go, the greater the number of propositions that are *entailed.*

An important algebraic structure for our purposes is a *groupoid*. A groupoid is a set with a binary operation, $\circ$, on the set. A groupoid with a partial ordering $\leq$ that is well-behaved with respect to the groupoid operation (i.e., if $a \leq a'$ and $b \leq b'$, then $a \circ b \leq a' \circ b'$) is a *partially ordered groupoid* or *pogroupoid* for short. In pogroupoids, $\circ$ is a match for fusion. We can now get a match for the right conditional and the left conditional via the *right residual* and *left residual* respectively:

Right residual: $a \circ b \leq c$ iff $b \leq a \rightarrow c$

Left residual: $a \circ b \leq c$ iff $a \leq c \leftarrow b$

A groupoid with both left and right residuals is a *residuated groupoid*. The right residual specifies the deduction theorem and its converse, see Dunn and Hardegree (*op. cit.*, 109). We can now specify the algebraic structure that matches our information frame. **I** is a *residuated pogroupoid* $(P, \leq, \circ, \rightarrow, \leftarrow)$

At this point, we return briefly to the issues of connectives, in order to discuss, and reject, the truth-constant t. In algebraic terms, t corresponds to the identity operator 1. By adding 1 to **I**, we move beyond the family of structures of groupoids to the family of structures of *monoids*. How might we go about interpreting 1 in the context of inferential information? The behaviour of 1 is specified by the following: For all $a \in P$:

$$a \circ 1 = 1 \circ a = a$$

In terms of information, 1 is the information warranted by logic, or the body of logical information. A related way of thinking about 1 is that it is the conjunction of all *theorems*. In this case, we have it that 1 may act as a *license* for valid inferences, such that $1 \leq A \rightarrow A$. This has obvious utility in the context of information flow in general, as in this case we want to be able mark the difference between the information flowing from some contingent fact or other on the one hand, and that flowing from truths of logic on the other. However, in the context of inferential information, we are already "inside" the flow of information from logic. In the context of inferential information then, although the identity operation may be omnipresent, it is redundant insofar as marking a distinction is concerned. Similarly with its semantic analogue, 0, the "logic situation".[21]

Then, the algebraic structure that matches our information frame **I** is a *residuated pogroupoid*, $(P, \leq, \circ, \rightarrow, \leftarrow)$. There is no argument for this *per se*, it is simply a function of residuated pogroupoids being the algebraic match for double implicational non-associative frames. At this point is worth reiterating the pluralist stance: it does not follow form this that residuated pogroupoids are the only algebraic structures relevant to inferential information flow, and especially not to information flow in general. Rather, they are the basic algebraic structures given the match to the

[21] Similarly also with the trivial truth constant $\top$, or "top": $\top$ is such that for every $a \in P$, $a \leq \top$.

relevant information frame **I**, and **I** is the relevant information frame for the very reasons which have by this stage, hopefully, been explained.

5. Conclusion

The central claim has been that the model-theoretic semantics underpinning **NL**, when given an information-based procedural re-interpretation, give an accurate account of inferential information. The running rationale for this revolved around the operation of information-application. The dynamic model-theoretic operation of information combination $x \bullet y \sqsubseteq z$ states that that combination of the information in states x and y develops into the information in z. The combination operator $\bullet$ is the directional application operation via non-commutation. The information-application operation allows us to generate successive information states given that the application is carries out in the appropriate manner. The appropriate manner is laid out in §2. The successive rejection of Contraction, Commutation, and Association correspond to the move from multisets, to lists/ordered-sequences, and finally to pairs respectively. The move from sequences to pairs is important, and has a perspicuous exposition in terms of inferential information:

We can think of an inferential sequence as a list of mechanisms that will allow information to flow through. For these mechanisms to allow information to flow, they have to be paired with one of the mechanisms that they are next to. However, they have to be paired up in the right way. This is because each pairing results in a "gate" that changes the shape, or pattern, of the information that flows through. If the mechanisms in the list are combined or paired up in the wrong way, then the pattern of information that results form being pushed through the first gate may be such that it is incompatible with the shape of the gate created by the next pairing. Hence, combination needs to be restricted to the *appropriate* pairs so that the information can flow through all the way to the end. Associativity on the frame destroys our ability to make this restriction; hence we require a non-associative frame.

But do we require frames at all? Or more specifically, do we require the model-theoretic semantics that has been outlined above? The Lambek calculi are commonly given category-theoretic or Curry-Howard lambda-term semantics. So then, do we *really* need, or want, the model-theoretical perspective on inferential information? There is good reason to think that we do. In an inferential procedure, all the inference rule executions have something in common, irrespectively of which rules and classical connectives they involve. They all involve the combination of information. This information-application operation is captured by the semicolon and by fusion at the structural and formula level respectively. Via the binary combination operator on information states, the model-theoretic semantics wears this operation on its sleeve perspicuously. This is crucial, as it means that our semantics explicitly manifests the very operation of information application that underpins inferential information. Although other applications of the structures captured by the Lambek calculi may be better served by different semantics, the model-theoretic perspective on inferential information has an undeniably rich pay off.

References

Abramsky, S. (2007): Information, Processes and Games, *forthcoming* in van Benthem, J. and Adriaans, P. (eds.), 2007.

Allo, P. (2007): Logical Pluralism and Semantic Information, *Journal of Philosophical Logic*, 36: 6, 659-694.

Allo, P. (2006): Local Information and Adaptive Consequence, *Logique et Analyse*, 149, 461-88.

Baltag, A, and Sadrzadeh, M. (2006): The Algebra of Multi-Agent Dynamic Belief Revision, *Electronic Notes in Theoretical Computer Science*, 157(4), 37-56.

Barwise, J. (1993): Constraints, Channels, and the Flow of Information, in *Situation Theory and its Applications*, 3., Aczel (*et. al.*) (*eds.*), CSLI Lecture Notes 37, CSLI Publications, Stanford.

Barwise, J. and Perry, J. (1983): *Situations and Attitudes*, CSLI Publications, Stanford.

Barwise, J. and Seligman, J. (1997): *Information Flow: The Logic of Distributed Systems*, Cambridge Tracts in Theoretical Computer Science # 44, Cambridge University Press, Cambridge.

Blackburn, P., de Rijke, M., Venema, Y. (2001): *Modal Logic*, Cambridge Tracts in Theoretical Computer Science # 53, Cambridge University Press, Cambridge.

Benthem, J. van (*forthcoming*): The Stories of Logic and Information, *Synthese*, Special *Knowledge*, *Rationality*, *and Action* issue: Proceedings of the First Workshop in the Philosophy of Information and Logic, University of Oxford, November 3-5, 2007.

Benthem, J. van and Martinez, M. (2007): The Stories of Logic and Information, *forthcoming* in van Benthem, J., and Adriaans (eds.) 2007.

Benthem, J. van and Adriaans, P. (2007): *Handbook on the Philosophy of Information*, Handbook of the Philosophy of Science Series, Elsevier.

Benthem, J. van (1995): *Languages in Action: Categories, Lambdas, and Dynamic Logic*, MIT Press, Cambridge.

Buszkowski, W. (1989): Remarks n Autoepistemic Logic, typescript.

Coecke, B. and Martin, K. (2002): A Partial Order on Classical and Quantum States, *research Report* PRG-RR-02-07, web.comlab.ox.ac.uk/oucl/publications/tr/rr-02-07.html

Dunn, J. M. and Hardegree (2001): *Algebraic Methods in Philosophical Logic*, Oxford Science Publications, Oxford Logic Guides # 41, Oxford University Press, Oxford.

van Eijck, J. and Visser, A. (eds.) (1994): *Logic and Information Flow*, Foundations of Computing Series, MIT Press, Cambridge Massachusetts.

Fagin, R., Halpern, J. Y., Moses, Y., and Vardi, M. Y. (1995): *Reasoning about Knowledge*, MIT Press, Cambridge, MA.

Floridi, L. (2003): Outline of a Theory of Strongly Semantic Information, *Minds and Machines*, 14, 197-222.

Gabbay, Dov. M., (1996): *Labelled Deductive Systems vol. 1*, Oxford Logic Guides no. 33, Oxford University Press, Clarendon – Oxford.

Galatos, N., Jipsen, T., Kowalski, T., Ono, H. (2007): *Residuated Lattices: An Algebraic Glimpse at Substructural Logic*, Studies in Logic and the Foundations of Mathematics, Volume 151, Elsevier, The Netherlands.

Hintikka, J. (1970): Surface Information and Depth Information, in Hintikka, J. and Suppes, P. (eds.): *Information and Inference*, Synthese Library, Humanities Press, New York, 263-97, 1970.

Hintikka, J. (1973): *Logic, Language-Games, and Information*, Oxford, Clarendon Press.

Jago, M. (2006): Imagine the Possibilities (Information Without Overload), *Logique et Analyse*, 49, 345-70.

Lambek, J. (1961): On the Calculus of Syntactic Types, in R. Jakobsen (ed.), *Structure of Language and its Mathematical Aspects*, American Mathematical Society, Providence, 166-178.

Levesque, H. J. (1990): All I know: A Study in Autoepistemic Logic, *Artificial Intelligence*, 42, 263-309.

Mares, E. (2004): *Relevant Logic: A Philosophical Interpretation*, Cambridge University Press, Great Britain.

Mares, E. (1997): Relevant Logic and the Theory of Information, *Synthese*, 109, 345-60.

Paoli, F. (2002): *Substructural Logics: A Primer*, Trends in Logic: Studia Logica Library, Volume 13, Kluwer Academic Publishers, The Netherlands.

Priest, G. (2001): *An Introduction to Non-Classical Logic*, Cambridge University Press, Great Britain.

Restall, G. (2002): *An Introduction to Substructural Logics*, Routledge, London.

Restall, G. (1996): Information Flow and Relevant Logics, in Seligman, J. and Westerstahl, D. (eds.): *Logic, Language and Computation*, CSLI, Stanford California, 463-77, 1996.

Restall, G. (1994): A Useful Substructural Logic, *Bulletin of the IGPL*, 2(2), 137-48.

Sequoiah-Grayson, S. (2009): Negative Information and Split Negation, *forthcoming* in *The Logica Yearbook 2008*, proceedings of the 22st International Symposia Logica 2008, Filosophia, Prague, 2009.

Sequoiah-Grayson, S. (2008a): The Scandal of Deduction: Hintikka on the Information Yield of Deductive Inferences, *Journal of Philosophical Logic*, 37: 1, 67-94.

Sequoiah-Grayson, S. (2008a): Metasemantic Information and the Scandal of Deduction, *forthcoming* in *The Logica Yearbook 2007*, proceedings of the 21st

International Symposia Logica 2007, Filosophia, Prague, 2008.

Sequoiah-Grayson, S. (2007): The Metaphilosophy of Information, *Minds and Machines*, 17: 331-44.

Sequoiah-Grayson, S. (2006): Information Flow and Impossible Situations, *Logique et Analyse*, 49, 371-98.

Shannon, C. E. (1950): The Lattice Theory of Information, reprinted in N. J.A. Sloane & A. D. Wyner (eds.): *Collected Papers*, Los Alamos, Ca: IEEE Computer Society Press, 1993.

Troelstra, A. S. (1992): *Lecture Notes on Linear Logic*, CSLI Lecture Notes No. 29, CSLI Publications, Stanford.

Wansing, H. (1993): *The Logic of Information Structures*, Lecture Notes in Artificial Intelligence, No. 681: Subseries of Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg.