Logics, Situations and Channels*

Greg Restall

Philosophy Department, University of Melbourne
restall@unimelb.edu.au

The notion of that information is relative to a context is important in many different ways. The idea that the context is small — that is, not necessarily a consistent and complete possible world — plays a role not only in situation theory, but it is also an enlightening perspective from which to view other areas, such as modal logics, relevant logics, categorial grammar and much more.

In this article we will consider these areas, and focus then on one further question: How can we account for information about one thing giving us information about something else? This is a question addressed by channel theory. We will look at channel theory and then see how the issues of information flow and conditionality play a role in each of the different domains we have examined.

Keywords: logic, situation theory, channel theory, information flow

1. Situations

Situation theory started as a very particular discipline, concentrating on the semantics of natural language.1 However, techniques from situation theory are similar to techniques used in many other areas of logical theory. In this article, I will look at situation theory abstractly. For us, a domain of situations will be some class of objects, which somehow ‘support information,’ and the information ‘carried’ by a situation has some kind of logical coherence. There

---

*Thanks to an audience at the 1996 ITALLC conference who gave me useful feedback on this material, and to Chungmin Lee for encouraging me to update these notes and present them to a current audience.

1 See Barwise and Perry’s book [7], and the literature which followed it.
are many different ways to make spell out these notions.

1. **Situation Theory.** The original applications of situation theory provide one obvious example. Consider claims of the form “Max saw Queensland win the Sheffield Shield”. How is this to be understood? For the Barwise and Perry of *Situations and Attitudes* [7], this was to be parsed as expressing a relationship between Max and a *situation*, where a situation is simply a restricted bit of the world. Situations are parts of the world and they support information. Max saw a situation and in this situation, Queensland won the Sheffield Shield. If, in this very situation, Queensland beat South Australia, then Max saw Queensland beat South Australia.

This shows why for this account situations have to be (in general) restricted parts of the world. The situation Max saw had better not be one in which Paul Keating lost the 1996 Federal Election, lest it follow from the fact that Max witnessed Queensland’s victory that he also witnessed Keating’s defeat, and surely that would be an untoward conclusion.

The situation theoretic analysis we’ve sketched shows the difference between ‘seeing’ and ‘seeing that’. It’s a very different claim you make by saying “Max saw that Queensland had won the Sheffield Shield”. You might want to say that this is a relationship between Max and a proposition, or Max and a fact. Whatever it is, it’s not the same kind of relationship expressed by the ‘seeing’ expression, as it doesn’t follow that Max saw that Queensland beat South Australia.

The details of the situation theoretic analysis of English expressions is not our brief here — that is *situation semantics* — our purpose is to note that here, situations are pieces of the world which support information. The situation we’ve talked about supports the information that Queensland won the Sheffield Shield, that Queensland beat South Australia, and it doesn’t support the information that Paul Keating lost the 1996 Election. Let’s denote this relationship as follows. We’ll abbreviate the claim that the situation $s$ supports

---

2 The Sheffield Shield was the annual cricketing competition earnestly fought out between the six Australian states. Queensland won for the first time (in over sixty years of trying) in 1995.

3 Or he saw a *scene* — the difference between situations and scenes developed in Barwise and Perry’s early work need not concern us here.
the information that $A$ by writing ‘$s \models A$’, and we’ll write its negation, that $s$ doesn’t support the information that $A$ by writing ‘$s \not\models A$’. This is standard in the situation theoretic literature. We won’t get too hung up on what sorts of things get substituted for the $A$ in these claims. In the Barwise and Perry of *Situations and Attitudes*, $A$ is a *state of affairs*. For later works by Barwise and Etchemendy [5] and Devlin [13], $A$ is an *infon*. We’ll stick to the more current usage by calling these expressions infons when discussing traditional situation theory.

The information carried by these situations has, according to Barwise and Perry, a kind of logical coherence. For them, infons are closed under conjunction and disjunction, and $s \models A \land B$ if and only if $s \models A$ and $s \models B$, and $s \not\models A \lor B$ if and only if $s \not\models A$ or $s \models B$. However, negation is a different story — clearly situations do not support the traditional equivalence between $s \not\models A$ and $s \not\models A \land \not\models A$. This is standard in the situation theoretic literature. We won’t get too hung up on what sorts of things get substituted for the $A$ in these claims. In the Barwise and Perry of *Situations and Attitudes*, $A$ is a *state of affairs*. For later works by Barwise and Etchemendy [5] and Devlin [13], $A$ is an *infon*. We’ll stick to the more current usage by calling these expressions infons when discussing traditional situation theory.

The information carried by these situations has, according to Barwise and Perry, a kind of logical coherence. For them, infons are closed under conjunction and disjunction, and $s \models A \land B$ if and only if $s \models A$ and $s \models B$, and $s \not\models A \lor B$ if and only if $s \not\models A$ or $s \models B$. However, negation is a different story — clearly situations do not support the traditional equivalence between $s \not\models A$ and $s \not\models A \land \not\models A$. This is standard in the situation theoretic literature. We won’t get too hung up on what sorts of things get substituted for the $A$ in these claims. In the Barwise and Perry of *Situations and Attitudes*, $A$ is a *state of affairs*. For later works by Barwise and Etchemendy [5] and Devlin [13], $A$ is an *infon*. We’ll stick to the more current usage by calling these expressions infons when discussing traditional situation theory.

What to do? Well, Barwise and Perry suggest that negation interacts with conjunction and disjunction in the familiar ways — $\not\models (A \lor B)$ is (equivalent to) $\not\models A \land \not\models B$, and $\not\models (A \land B)$ is (equivalent to) $\not\models A \lor \not\models B$. And similarly, $\not\not\models A$ is (equivalent to) $\not\models A$. This much gives us a logic of sorts of negation — it’s the *first degree entailment* of Anderson and Belnap [1]. Barwise and Perry don’t stop there. Actual situations must be consistent. That is, you don’t get $s \models A \land \not\models A$ for any $A$. So, this gives you a stronger logic of negation — you get $A \land \not\models A \vdash B$, where you define $A \vdash B$ to be satisfied just whenever for every situation $s$, if $s \models A$ then $s \not\models B$. The resulting logic of ‘actual’ situations is Kleene’s strong three valued logic, or the conjunction, disjunction and negation fragment of Lukasiewicz’s three valued logic.

But this is not the end of the story of negation in situation semantics. There is an issue of whether the negation of an arbitrary infon is defined. In particular, if we have an infon of the form $\exists x F(x)$, where $s \models \exists x F(x)$ iff for some object $b$ in the situation $s$, $s \models F(b)$, then what is it for $s \not\models \exists x F(x)$ to be true? This is quite a subtle question, and it’s one in which situation theorists do not agree [3].

---

4 It is unsuprising that this is a difficult question: solving it means taking a stand on
2. INTUITIONISTIC LOGIC. Consider the Kripke-style semantics for intuitionistic logic [14, 22, 24, 48]. Under the received interpretation, you have a set of points, each of which represents the state of knowledge of a mathematical reasoner. These points are ordered under the relationship of possible extension. A point is a possible extension of another point if you can get from the first to the second by some (possibly extended) process of mathematical reasoning. Each point has a domain, of mathematical objects which have been constructed in the process of getting to that point. Now, what is it for a mathematical statement \( A \) to be true in a point? It is for there to be a demonstration of that statement using the resources you have at that point. Now consider the usual logical connectives. Using the same expression ‘\( \models \)’, this time relating points and mathematical claims, we have (according to the received view)

- If \( s \models p \) (\( p \) atomic) and \( t \geq s \) then \( t \models p \).
- \( s \models A \land B \) iff \( s \models A \) and \( s \models B \).
- \( s \models A \lor B \) iff \( s \models A \) or \( s \models B \).
- \( s \models \neg A \) iff for all \( t \geq s \), \( t \models \neg A \).
- \( s \models A \supset B \) iff for all \( t \geq s \), if \( t \models A \) then \( t \models B \).
- \( s \models \exists x A(x) \) iff \( s \models A(b) \) for some \( b \in D(s) \).
- \( s \models \forall x A(x) \) iff for all \( t \geq s \), and for all \( b \in D(t) \), \( t \models A(b) \).

It’s pretty clear that the points in the semantics of intuitionistic logic aren’t the sorts of things Barwise and Perry had in mind when they discussed situations. However, they too are objects which support information, and they can be thought of as “restricted bits of the world” (or at least things which describe restricted bits of the world) in much the same way as Barwise-Perry situations.

One difference between the logic of this account and that of the Barwise-Perry story is in the treatment of negation. Here we have a clear account of

‘universal facts’, if one thinks that \( \neg \exists x \neg F(x) \) is supported when and only when \( \forall x F(x) \) is supported. Then deciding when an infon supports \( \neg \exists x F(x) \) is tantamount to deciding when an infon supports \( \forall x F(x) \). That is, you need to decide when it is for a situation to provide a universal fact.
what it is for \(~A\) to be true at a point \(s\). It is for there to be no extension of \(s\) at which \(A\) is true. Here, as with Barwise and Perry, points can be incomplete. If \(s\) has an extension at which \(A\) is true, but \(A\) isn’t true at \(s\), then \(s \not\models A\) and \(s \not\models \sim A\). However, we do not have \(\sim \sim A\) as equivalent to \(A\). If every extension of \(s\) has some extension which supports \(A\), then \(s \models \sim \sim A\). But it need not follow that \(s \models A\). For Barwise and Perry, \(s \models \sim \sim A\) just when \(s \models A\).

Of course, one response is to say that Barwise and Perry are not utilising the same notion as intuitionists — they’re formalising different notions. If this is in fact the case, then there are further interesting issues — what sense can be made of intuitionistic negation in the context of Barwise-Perry situation semantics, and further, what sense can be made of the negation of Barwise and Perry for intuitionists? The former has not received any attention in the literature as far as I can find. The latter question takes us down the route of Nelson’s constructive negation [35], which has been fruitfully explored recently by David Pearce [36] and Heinrich Wansing [49] among others.

Another issue raised by these structures is their reliance on the relation \(\preceq\) between points. Propositions are persistent along \(\preceq\) in that you can show (by induction on the construction of formulae) that if \(s \models A\) and \(s \preceq t\) then \(t \models A\) too. Barwise-Perry situation theory also has an ordering \(\preceq\) on situations interpreted as ‘part of’, but there is some dispute as to whether a persistence condition ought to hold [3]. This is related to the semantics of the existential quantifier, and we will come to this issue again later.

3. TRUTHMAKING. Another domain for the sorts of semantic theories we’re looking for is not so often discussed. It’s the philosophical notion of a ‘truthmaker’ [21, 34]. A truthmaker for a statement \(p\) is an object \(x\) such that the existence of \(x\) entails \(p\). If we let the making true relation be denoted by \(\models\), then presumably you have a sort of logical structure similar to those given by situation theory or the semantics of intuitionistic logic. However, the logic of truthmakers is subtle and presents its own difficulties [38, 41].

4. MODAL LOGICS. The Kripke-style semantics for modal and temporal logics provide another example of a formal theory in which points support information, and this information has some kind of logical coherence. Here, the points are all complete, in that for each world (or time, or whatever) \(w\), \(w \models \sim A\) iff \(w \not\models A\). These logics enrich the standard propositional or predicate languages with other operators — most often unary operators \(\Box\) and \(\Diamond\), and it is only
these which actually use the machinery of possible worlds in any essential way. We have the familiar clauses

- \( w \models \Box A \) iff \( v \models A \) for each \( v \) where \( wRv \).
- \( w \models \Diamond A \) iff \( v \models A \) for some \( v \) where \( wRv \).

where \( R \) is a binary relation of accessibility between worlds. (Good introductory works are those by Hughes and Cresswell [25, 26], Chellas [11] and Goldblatt [23]. The definitive textbook on a mathematical approach to modal logic is Blackburn, de Rijke and Venema [10].)

Here, to get interesting results, you are forced to construe most of the worlds as “non-actual” in some sense or other. If the actual world is represented as one point \( w \), then any world which disagrees with \( w \) in the evaluation of formulae is some alternate way things could have been, but in fact isn’t. This is not necessary in any of the previous accounts. We could well think of each situation as part of the one world (and this is the view of the Barwise and Perry of \textit{Situations and Attitudes}, and of Israel and Perry [27]) and so, that there are no ‘non-actual’ situations. However, one could just as easily allow non-actual situations into our semantics. If we do, there would be some hope that we get a semantics for necessity and possibility, analogous to that of traditional modal logic. If not, you need to have a radically different account of what it is for a situation to support the infon \( \Box A \) (if indeed, that is an infon).

5. RELEVANT LOGICS. We have already mentioned Anderson and Belnap’s first degree entailment in the context of the kind of logical coherence exhibited by the class of infons supported by a situation. The Routley-Meyer semantics for relevant logics is another case where we have points supporting some kind of information, which has a degree of logical coherence [2, 15, 19, 32, 44, 45]. The semantic structures associated with relevant logics are interesting in a number of ways. Firstly, as relevant logics reject the inferences \( A \land \neg A \vdash B \), and from \( A \vdash B \lor \neg B \), but they keep the intuition that entailment on a frame is defined by taking \( A \vdash B \) to be valid just when for every point \( x \) either \( x \not\models A \) or \( x \models B \), they must insist that there are points \( x \) such that \( x \not\models A \land \neg A \), and other points \( x \) such that \( x \not\models A \lor \neg A \). The second rejection we’ve seen before, in Barwise-Perry situation theory, and in the frames for intuitionistic logic. The first condition—in explicitly accepting points which are ‘inconsistent’ is more
novel. It is necessary for the frames of relevant logics to get the desired results.

What are we to make of something like \( x \models A \land \neg A \)? Much ink has been spilled on this issue,\(^5\) but I will be too dogmatic on the point here. Save to say that for anyone finding the semantics of relevant logics useful or interesting, the points in frames will not be interpreted as possible worlds. One way to think of them is that some are ways the world (or part of the world) could be, and some are ways the world couldn’t be. Much more will need to be done with this to make it philosophically acceptable (or even defensible) but this should be enough to give you a way of thinking about it.

The semantics of relevant logics assign treat conjunction and disjunction in the usual way, but negation and (relevant) implication are given a novel treatment.

\[
\begin{align*}
\bullet \ x \models \neg A & \iff x^* \not\models A. \\
\bullet \ x \models A \rightarrow B & \iff \text{for each } y, z \text{ where } Rxyz, \text{ if } y \models A \text{ then } z \models B.
\end{align*}
\]

These two clauses have been the cause of quite a bit of discussion.\(^6\) Let’s look at negation first. If \( x^* \) is not always the same point as \( x \), we can see that you’ll be able to allow \( x \models A \land \neg A \) or \( x \not\models A \lor \neg A \) by a judicious choice of evaluation. But what is the function \( ^* \)? Given a point \( x \), what is \( x^* \)? If \( x \) is a way the world could be or a way the world couldn’t be, then what is \( x^* \)? I don’t think much sense can be made of this by itself. But, J. Michael Dunn’s work on the semantics of negation is a real help [16, 17, 18]. Ignore \( ^* \) for the moment, and consider a binary relation \( C \) of compatibility between points. If we have such a relation of compatibility to hand, then it makes sense to evaluate negation as follows

\[
\bullet \ x \models \neg A \iff \text{for each } y \text{ where } xCy, y \not\models A.
\]

Then with this evaluation to hand, and provided that there is a relation \( \leq \) on points with respect to which propositions are persistent, then \( x^* \) is simply the

---

\(^5\) See the literature around the Copeland and Routley-Meyer debate [12] and others, and in more recent work by J. Michael Dunn [18], Ed Mares [32] and me [42].

\(^6\) See the discussion cited in the previous footnote.
unique ≤-maximal point y such that xCy. (Check that this assumption makes the original clause for negation an instance of the ‘compatibility’ clause.)

Now, there may or may not be a reason to think that there is a unique ≤-maximal point y compatible with x. That would be one condition you could impose upon a compatibility relation, just as imposing different conditions on accessibility relations in modal logics gives you different logics. There’s a theory about logics of negation just as rich as that of traditional modal logics. This is a way of understanding the original “Routley star” which makes some kind of sense of it.

Now for the ternary relation R. Ternary relations are often difficult to understand, but some sense can be made of R [32, 40]. The reason why relevant logics have a ternary accessibility relation is again related to the view of entailment. Specifically, since entailment is relevant, we reject the inference A ⊬ B → B, since the succedent need not have anything to do with the antecedent. But then, traditional treatments of implication operators (which take A → B to be true at a point iff for associated points y where y ⊨ A then y ⊨ B) won’t work, because these treatments make B → B true at every point. So A ⊬ B → B follows immediately. What we need is a way to ensure that B → B is not true everywhere, and the ternary relation fills this need. Just have points y and z where Rxyz, y ⊨ B, and y ⊭ A, and you’re done.

But what does this mean? One interpretation is simple to give. Rxyz holds just when whenever you apply the information given by x (as rules for inference) to the information in y (premises) you get conclusions which are in z. You can think of the points as theories, or more concrete sorts of objects. Whatever, something like this will motivate the general structure of the ternary relation. Just as is familiar with modal logics, different conditions on the accessibility relation will give you different logics of implication.

Note that the situation here is similar to that with modal logics in that to get a wide range of logics, you need ‘non-actual points’ of some kind. In fact, as we’ve seen, to get the kinds of logical theories relevant logicians look for, you need not just ‘merely possible’ points, but outright impossible ones.

6. CATEGORICAL GRAMMAR. Joachim Lambek’s categorical grammar is at first sight a completely different kettle of fish [29, 30]. Here, the points in the semantic structure are pieces of syntax, and the ‘infons’ are syntactic classifications of various kinds. For example, the classifications into noun
phrases, verbs, sentences, and so on. The interest comes with the way in which these classifications can be combined. For example, a form of conjunction $A \circ B$ can be defined, where we say $x \models A \circ B$ iff $x$ is a concatenation of two strings $y$ and $z$, where $y \models A$ and $z \models B$. We can also define ‘slicing’ operations, setting $x \models A \setminus B$ iff for each $y$ where $y \models A$, $yx \models B$; and $x \models B / A$ iff for each $y$ where $y \models A$, $xy \models B$.

What’s interesting here is the fact that the points in the structure — syntactic strings — can be seen in just the same way as the points in our other structures. They support information (here their syntactic types) and this information has a form of logical closure. If $x \models A \circ (A \setminus B)$, then $x \models B$. We have a form of modus ponens for example. If you like, you can enrich the logic of strings with conjunction and disjunction, and if you do it in the obvious way (using the same clauses as in the other logics) you get a formal logic quite a bit like ones which are independently motivated, in completely different domains [39]. These logics are studied in depth by logicians and theoretical linguists [28, 33].

7. MAPS AND DIAGRAMS. Another example of a semantics in which we have objects supporting information is given by maps and diagrams. If you have a map, then any region in the map supports information. (Of course, this could be misinformation.) Presumably, this notion of information supporting is susceptible to the same sorts of treatments as the other notions we’ve seen. Each region in the map supports information, and the body of information carried by a region has some sort of logical closure. Perhaps it will turn out that an appropriate semantics for maps and diagrams will be very similar to the structures we’ve already seen, and perhaps it will be different. Of course, there is the issue of whether linguistic representation, with the syntax of well formed formulae and the associated paraphenalia, will capture all of the important information carried by maps and diagrams. On a more mundane note, there is the issue of disjunctive information. We might agree that a situation supports $A \lor B$ just when it supports one of $A$ and $B$, but is this true of maps? Can a map region tell me that this (depicted by a wiggly line) is either a road or a river, without telling me which? If disjunction is an appropriate notion in this domain,

---

You won’t be the first to notice the similarity to the clause for $\rightarrow$ in the semantics for relevant logics. Replace $Rxyz$ with $z = xy$ and you’re done. $A \rightarrow B$ is $B/A$. 
then it will presumably be treated differently, if maps are not ‘precise.’

We will not consider the details of maps and diagrams here. They are beginning to receive quite a bit of attention from different quarters, and it will be interesting to see just what comes of these investigations [6, 37, 47].

2. A Worked Example

To apply what we’ve looked at, we’ll try working through a ‘baby example’ which will show us some of the choices you have to make in building a formal semantics based on these ideas.

Consider a plane, marked off into pixels in a regular grid. A pixel can be inhabited or uninhabited. Here is an inhabited pixel and an uninhabited pixel.

So, here is what a part of our world might look like

Now, let’s consider how to construct the semantics of a language describing such a world. The language might be the traditional language of first order logic, with conjunction, disjunction and negation, quantifiers and identity, with other special predicates. The predicates appropriate for our language depend on what the ‘objects’ are. For this discussion we’ll let the objects be the inhabited pixels and the merelogical fusions of inhabited pixels. We have to decide whether we want “disconnected objects”. That is, is the fusion of the two atoms below a single object?

---

8 For examples of formal semantic structures in which disjunction is not treated with the usual clause, consider the Beth semantics for intuitionistic logic, and the Fine semantics for relevant logics [20, 14, 48].

9 To see this example worked out in more detail, see ‘Modelling Truthmaking’ [43].
Let's decide for the moment that it isn't. Quantifiers range only over “connected entities” only.

Predicates appropriate for our world might be \textbf{Is\_Square}(x), \textbf{Contains}(x, y), \textbf{Larger\_than}(x, y), \textbf{Disjoint}(x, y), each with obvious interpretations. Now, given a world, we could interpret the language in the obvious way, using the standard clauses of first order logic. \textbf{Is\_Square}(a) is true iff \( a \) is a square in the world, \( A \land B \) is true iff \( A \) and \( B \) are true, \( \forall x A(x) \) is true iff \( A(a) \) is true for each object \( a \) in the world, and so on. But to do that would be to ignore how these things are made true by regions of the world — it would be erase the distinctions we are able to make. We can do better, and introduce a semantic theory which utilises the techniques we’ve seen in the previous section.

Let’s allow that there are regions of the world which will function as the points in our semantic structure. To do this, we have to be formal. The world can be represented as a function \( w: \mathbb{Z} \times \mathbb{Z} \rightarrow \{□, ■\} \). A region then is a \textit{partial function} \( r: \mathbb{Z} \times \mathbb{Z} \rightarrow \{□, ■\} \), and we will allow disconnected regions (for a technical reason we’ll come to later) so the domain \( L(r) \) of the function \( r \) (which we might call the region’s \textit{location}) will be a subset of \( \mathbb{Z} \times \mathbb{Z} \). Then a region \( r \) is a part of the world \( w \) iff \( r \subseteq w \). We can graphically represent regions by picking out the coordinates of one point of the region, as follows.

The coordinate indicates the location of the part in ‘world’. Note here that this region has a hole in it. That is, \( r(4, 5) = □ \), and \( r(5, 4) = ■ \), but \( r(6, 3) \) is undefined.

Now every region \( r \) has a domain associated with it. \( D(r) \) is the collection of objects in \( r \). These objects are the connected partial functions \( o: \mathbb{Z} \times \mathbb{Z} \rightarrow \{■\}.\)
which are subsets of $r$, and which always return $\square$. Clearly the domain of a world is the collection of all objects in the world, and if $r_1 \leq r_2$, then $D(r_1) \subseteq D(r_2)$ — domains increase as regions increase.

Now we can put this account to work. For simplicity of presentation, we will assume that every object has a name which features in our language. That is, there’s an interpretation function $I$ which maps names to objects. Now consider a predicate like $\text{Is\_Square}$. What is it for $\text{Is\_Square}(a)$ to be true in a region $r$? One way to cash this out is as follows:

$$r \models \text{Is\_Square}(a) \text{ iff } I(a) \in D(r) \text{ and } I(a) \text{ is square}$$

It’s true in a region that $a$ is a square just when $a$ is in that region and $a$ is a square. There’s nothing more to $a$’s squareness than what is in the region $r$. This is very different to the predicate $\text{Larger\_than}$. If $a$ is in $r$ but $b$ isn’t, then we will not have $r \models \text{Larger\_than}(a, b)$, even though $a$ might be larger than $b$. We should have instead, the clause

$$r \models \text{Larger\_than}(a, b) \text{ iff } I(a), I(b) \in D(r) \text{ and } I(a) \text{ is larger than } I(b)$$

This semantic story enables us to reflect the distinction between ‘local’ and ‘non-local’ predicates in our formalism. A predicate $F$ is locally true of $a$ just when any for any region $r$ containing $a$, we have $r \models Fa$. So, $\text{Is\_Square}$ will be locally true of an object if it is true of that object, but $\lambda x. \text{Larger\_than}(x, b)$ need not be locally true of an object.

We can give our primitive language the resources of conjunction and disjunction without any difficulty at all using the familiar clauses. More interesting are negation and quantification. Here, we have issues to face. Let’s do negation first. One way to handle it is to follow the treatment of Barwise and Perry style situation theory, and to just define the negative extensions of predicates on regions, just as we have defined the positive extensions of predicates. For example, we would have

$$r \models \sim \text{Is\_Square}(a) \text{ iff } I(a) \in D(r) \text{ and } I(a) \text{ is not square}$$

$$r \models \sim \text{Larger\_than}(a, b) \text{ iff } I(a), I(b) \in D(r) \text{ and } I(a) \text{ is not larger than } I(b)$$
This will give us intuitively plausible results, but it comes at a cost. We have to define $\neg A$ for every complex expression $A$. Defining negated conjunctions and disjunctions is easy, but defining negated quantifiers and other expressions is more difficult. Instead of going down that route, let’s explore the kind of negation we get when we exploit a relation of compatibility between regions.

What is it for $r_1$ and $r_2$ to be compatible? Pretty clearly, $r_1 Cr_2$ iff there is no point $(n, m)$ where $r_1(n, m)$ and $r_2(n, m)$ are defined and different. If there are no points of disagreement in this sense, the two regions ‘fit together’. What happens if we define negation as follows?

$$ r \models \neg A \text{ iff for all } r' \text{ where } rCr' r' \not\models A $$

Then, for example, we get exactly the same results as with the two clauses we quoted before. For example, $r \models \neg \text{is_square}(a)$ iff $I(a) \in D(r)$ and $I(a)$ is not square. For if $I(a)$ is not square and $I(a) \in D(r)$, then $rCr'$ ensures that $r' \not\models \text{is_square}(a)$ as any regions in which $I(a)$ is square will not be compatible with $r$. And if $I(a)$ is not in $D(r)$ then we can find a compatible state in which $I(a)$ isn’t square, by setting the interpretation $I(a)$ of $a$ to be any non-square we like. The converse reasoning is similar.

Negation defined in that way acts well on our primitive predicates. However, it is perhaps not so well behaved in general. Suppose $r \models \neg \neg A$. This means that for all $r'$ where $rCr'$, $r' \not\models A$. And that in turn means that there’s some $r''$ where $r' Cr''$ and $r'' \models A$. Does this mean that $r \models A$? Well, it can’t in general. Here is why: Suppose that there is some general rule which is true in all worlds. That is, it’s true in all total regions. Examples of such a rule is “for all $x$ and $y$, either $x$ and $y$ are disjoint, or there is a $z$ which is a part of both $x$ and $y$”. Exactly how this is formalised is irrelevant to the point we’re making. The only important thing is that it’s plausible that it’s true in each world, and it’s implausible that it’s true at each and every region. And this is certain. If small regions record merely local facts, then it would be odd for such a general claim as that to hold at each region. So, if there are such propositions, let one be $A$. Given this, we must have $r \models \neg \neg A$, for any region $r$. Why is this? Take any $r'$ compatible with $r$. There is some world $w$ compatible with $r'$ and $w \models A$, so $r' \not\models \neg A$. But this just means that $r \models \neg \neg A$ as we desired.

So, under some plausible conditions, $\neg \neg A$ differs from $A$ in that it is true at
different regions. Note however, that negation isn’t completely non-classical, if
the evaluation \( \models \) is persistent along \( \subseteq \). For worlds \( w, w \models \neg A \iff w \not\models A \), since
for any world \( w, wCr \iff r \subseteq w \), and then \( w \models \neg A \iff \text{for any } r \text{ where } r \supseteq w, r \not\models A \iff w \not\models A \). Now this may or may not be a problem to you. If it is, then a way
out is possible—it is to take inconsistent regions seriously.

According to this approach, a region is a total function from \( \mathbb{Z} \times \mathbb{Z} \) to \( \{0, \{\blacksquare\}, \{\square\}, \{\blacksquare, \square\}\} \), where \( \blacksquare \in r(m, n) \) indicates that according to \( r \), \( (m, n) \) is
inhabited, and \( \square \in r(m, n) \) indicates that according to \( r \), \( (m, n) \) is uninhabited.
We are simply allowing \( r \) to be confused about points (thinking them to be both
inhabited and uninhabited) as well as not carrying any information about
points.

Objects are still partial functions from \( \mathbb{Z} \times \mathbb{Z} \) to \( \{\blacksquare, \square\} \) which are constant
\( \blacksquare \) where defined, but we need to extend our notion of what it is for an object
to occur in a region. A natural definition is for \( o \) to be in \( D(r) \) just when for
every \( \blacksquare = o(m, n), \blacksquare \in r(m, n) \). That is, that \( r \) records the presence every
‘atom’ in the object. This allows for an object to be present in a region, even
when the region has inconsistent information about part of that object. A natural
definition of containment of regions is the pointwise definition: \( r \subseteq r' \iff \text{for each } (m, n), \text{ we have } r(m, n) \subseteq r'(m, n) \). The only other major definition to
tackle is that of compatibility. Here, we can set \( rCr' \iff \text{there is no } (m, n) \text{ where } \blacksquare \in r(m, n) \text{ and } \square \in r'(m, n) \text{ or } \square \in r(m, n) \text{ and } \blacksquare \in r'(m, n) \). That is, there is
no point at which information from \( r \) conflicts with information from \( r' \).

Given this definition, it is reasonably easy to show that \( r \models \neg \neg A \iff r \models A \). For
we can define for any region \( r \) a corresponding region \( r^* \) which agrees with \( r \)
wherever \( r \) takes either \( \{\blacksquare\} \) or \( \{\square\} \), but which takes the value \( 0 \) where \( r \) takes
the inconsistent value \( \{\blacksquare, \square\} \), and vice versa. Then it is simple to see that \( rCx \iff x \subseteq r^* \), and as a consequence, \( r \not\models \neg A \iff r^* \not\models A \), (if propositions are
persistent).

We also have to deal with quantifiers. This is quite a tricky situation. Existential quantification is easier. It is quite plausible to assume the following:

\[ r \models \exists xA(x) \iff \text{for some } a \in D(r), r \models A(a) \]

This is a persistent proposition, and it has a clear semantic content.

Universal quantification is quite a different story. What is it for \( \forall xA(x) \) to be
true in a region \( r \)? There are at least three different ways to go. The first is to consider regional quantification, by setting

\[
r \models \forall x \in r'. A(x) \text{ iff } r' \subseteq r \text{ and for all } a \in D(r'), \ r \models A(a).
\]

The second is to abandon persistence, and to set

\[
r \models \forall x A(x) \text{ iff for all } a \in D(r), \ r \models A(a).
\]

The third is to ‘wire in’ persistence and keep the quantification universal by setting

\[
r \models \forall x A(x) \text{ iff for each } r' \supseteq r, \text{ and for each } a \in D(r'), \ r' \models A(a).
\]

This last account is that of the intuitionistic universal quantifier. If all of one’s points are consistent, then worlds, which are maximal points, will have a completely classical semantics — \( \forall x A(x) \) is true in a world just when \( A \) is true of all objects in the domain of that world.\(^{10}\)

3. Channels

It has been a problem in situation theory to understand how information carried by one situation can give you information about another. Why is it that \( s \models A \) can give you information that \( s' \models B \)?

Let’s have a look at a few examples. A student looks at her marked assignment, and the fact that this perceptual situation carries some information (it has an ‘A’ written on it) gives her the information that she’s passed her subject, that she’s completed her degree, and that she’s likely to get a job. Not only does it carry all of this information in some way or other, but she can legitimately infer things about other situations. She can infer that her tutor thought her assignment was a good one. So, the information carried by this

---

\(^{10}\) For more detail on this approach, including an interpretation of a relevant entailment connective, see “Modelling Truthmaking” [43].
situation gives her information about other situations — the marking situation for one, and her future job prospects, for another.

How are we to take account of this phenomenon? Barwise, in his paper “Constraints, Channels, and the Flow of Information” [4] marks out a few features he takes to be desirable for any account of information flow. We’ll use these as our starting point.

**XEROX PRINCIPLE:** If $s_1 : A$ carries the information that $s_2 : B$ and $s_2 : B$ carries the information that $s_3 : C$ then $s_1 : A$ carries the information that $s_3 : C$.

**LOGIC AS INFORMATION FLOW:** If $A$ entails $B$ (in some sense to be determined) then $s : A$ carries the information that $s : B$.

**INFORMATION ADDITION:** If $s_1 : A$ carries the information that $s_2 : B$ and $s_1 : A'$ carries the information that $s_2 : B'$, then $s_1 : A \land A'$ carries the information that $s_3 : B \land B'$.

**CASES:** If $s_1 : A$ carries the information that $s_2 : B \lor B'$ and $s_2 : B$ carries the information that $s_3 : C$ and $s_2 : B'$ carries the information that $s_3 : C$ then $s_1 : A$ carries the information that $s_3 : C$.

The other desirable feature is that we have some robust account of how information flow can be *fallible*. The student seeing her ‘A’ correctly gathers the information that she has passed. However, in a perceptually indistinguishable situation, she is reading a forgery. What are we to make of this case?

The characterising feature of channel theory is that there are objective ‘links’ between situations which support the flow of information, just as there are objective situations, which support information itself.

Channels give rise to connections between situations. That a channel $c$ links $s$ and $t$ is denoted ‘$s \leftarrow c \rightarrow t$’. An example of a channel given by Barwise and Seligman in ‘The Rights and Wrongs of Natural Regularity’ [8] is the Rey Channel, which links thermometer situations with patient situations. In particular, the fact that the thermometer’s mercury level has a particular height indicates something about the temperature of a patient — usually. The idea, is that this channel grounds the regularity out there in the world which grounds

---

11 Barwise also suggests a contraposition clause involving negation, but there are problems with his account of contraposition, so we won’t look at it here [40].
the connections between thermometer readings and patient temperatures.

So, if we have a situation $s$ which includes a thermometer, and we have a thermometer reading $A$, so we have $s \models A$, and the channel $c$ supports a regularity of the form $A \rightarrow B$ (if the height is $x$ then the temperature is $y$) then given that the situation $t$ is connected to $s$ by the channel $c$ ($s \xrightarrow{c} t$) we can infer $t \models B$. And this story makes some kind of sense. It appeals to our sense that there are ‘objective connections’ in the world, and these ground our ability to infer information in heterogeneous systems.

We need some small pieces of terminology. In $s \xrightarrow{c} t$, $s$ is a signal for the channel $c$ and $t$ is a target. A channel $c$ supports the constraint $A \rightarrow B$ just when for each signal-target pair $s$ and $t$ where $s \xrightarrow{c} t$, if $s \models A$ then $t \models B$.

An important feature of channel theory is the fact that we allow more than one channel. Information flows in more than one way — it is not just a matter of physical law, or convention, or logical entailment. The Rey channel mentioned above is but one of many. This channel is partly a matter of physical law, but it is also a matter of human convention — it is we who make thermometers and who use them in particular ways. Another obvious family of channels which is a mix of physical law and convention is the doorbell. Someone pushes a button, rings the doorbell, and indicates to us that someone is at the door. This can be analysed as a chain of channels. One from the bell situation to the doorbell button situation, another from the button situation to the situation out on the verandah. That is, information about the state of the bell (that it’s ringing) gives us information about the state of the button (that it’s been pressed). This is largely due to the wiring of the electrical network. Then information that the button has been pressed gives us the information that there’s someone on the verandah waiting to get in. These channels can be thought of as ‘chaining together’ to form one larger channel — and if you like, you can analyse the situation more closely and see that the smaller channels are composed of even smaller ones, which account for the way information flows between these situations.

We can use these distinctions to give a taxonomy of what goes on in information flow. And one thing which channel theory is useful for is in giving us a way to see how different things can go wrong in our inferring about situations.

For example, suppose that the thermometer has not been near any patient,
but the nurse takes a reading. If anyone infers anything about a patient’s
temperature from the thermometer reading, they are making a mistake. In this
case, the channel does not connect any patient situation with the thermometer
situation. We can say that the thermometer situation is a *pseudo signal* for the
channel $c$.

That this kind of error can be accommodated can help us analyse things like
the problems of counterfactuals. The conditional “If I drink a cup of tea, I feel
better” is grounded by a complex (physiological, psychological and no doubt,
sociological) channel which links tea drinking situations to better feeling
situations. The conditional is true. However, it is not true that if I drink a cup of
tea with poison in it, I feel better. But isn’t this a counterexample to the
regularity we thought we saw? It doesn’t have to be, for a situation in which I
drink tea with poison is a pseudo signal of the channel I discussed. The channel
does not link all tea drinking situations with matching better feeling ones. It
merely links “appropriate” ones.\(^\text{12}\)

However, not all channel errors are like this. Consider the following
situation, again discussed by Barwise and Seligman in “Rights and Wrongs”.
An air traffic controller is looking at a radar screen, which a complex channel
of electrical wiring, computational machinery and radar equipment connects to
a situation in the sky. Typically, a blip at a particular point on the screen
‘corresponds’ to an aircraft at a particular position in the sky. This channel
supports constraints. Consider the situation in which the aircraft is in ‘simulator
mode’. Instead of taking its images from external input, it simply plays a
selection of prerecorded images. In this case, the screen situation is pretty
clearly a pseudo signal for the channel connecting screen situations to sky
situations. But consider the case where there is sunspot activity, and this causes
a blip to appear on the screen. What should we say then? Is the screen a pseudo
signal in this case? The screen situation is clearly an exception to the regularity
$\text{Blip} @ p \rightarrow \text{Plane} @ p'$, but does this make the whole screen situation a pseudo
signal to the channel? Well, firstly, we can be sure that if the operator infers that

\(^{12}\)You should not take this as a sign that there is some fully developed theory of the
semantics of counterfactuals in terms of channels. There isn’t. This is merely a sign
as to how one might be formed. I discussed this problem a little in my paper linking
channel theory and the semantics of relevant logics [40].
there is a plane at position \( p' \), then they are making a mistake, even if there is no plane there, for the fact that the screen has a blip at \( p \) isn’t appropriately connected to the fact that there is a plane at \( p' \). However, it seems clear that if the rest of the screen is functioning correctly, and is not being interfered with by sunspot activity, then inferences made about other planes on the basis of other blips on the screen are warranted. Barwise and Seligman wish to argue that this case is one of a stronger kind of exception regularities. In this case the fact \( s \models \text{Blip}@p \) is an exception to the regularity \( \text{Blip}@p \rightarrow \text{Plane}@p' \), but the situation itself is not a pseudo signal to the channel. In this case, the channel is not exceptionless.

Another way to reason, of course, is to say that this does make the screen situation \( s \) a pseudo signal to the channel, which supports \( \text{Blip}@p \rightarrow \text{Plane}@p' \), but it is not an exception to a weaker channel which links more situations, including those which are affected by some kind of sunspot activity. I will not enter this debate about exceptions and channels here. Instead, I will look at how an account using channels can help us meet the desiderata we saw at the start of this section, and then we will see how channels can be incorporated into the kinds of point semantics we looked at in the first section.

**XEROX PRINCIPLE:** If \( S_1 : A \) carries the information that \( S_2 : B \) and \( S_2 : B \) carries the information that \( S_3 : C \) then \( S_1 : A \) carries the information that \( S_3 : C \).

This will be met if we require for every pair of channels \( c_1 \) and \( c_2 \), that there be a channel \( c_1 \cdot c_2 \) which composes \( c_1 \) and \( c_2 \), satisfying \( s \models c_1 \rightarrow t \) iff there’s a \( u \) where \( s \models c_1 \rightarrow u \) and \( u \models c_2 \rightarrow t \). Then it is simple to show that if \( s \models c_1 \rightarrow u \models A \rightarrow B \) and \( u \models c_2 \rightarrow B \rightarrow C \) then \( s \models c_1 \cdot c_2 \rightarrow A \rightarrow C \).

Here, \( c_1 \cdot c_2 \) is said to be the serial composition of \( c_1 \) and \( c_2 \).

**LOGIC AS INFORMATION FLOW:** If \( A \) entails \( B \) (in some sense to be determined) then \( s : A \) carries the information that \( s : B \).

Here, we need only an identity channel \( 1 \), which maps each situation onto itself. Then if \( A \models B \) is cashed out as “for each \( s \), if \( s \models A \) then \( s \models B \),” then \( A \) entails \( B \) iff \( 1 \models A \rightarrow B \).

**INFORMATION ADDITION:** If \( S_1 : A \) carries the information that \( S_2 : B \) and \( S_1 : A' \) carries the information that \( S_2 : B' \), then \( S_1 : A \land A' \) carries the information that \( S_2 : B \land B' \).

Here we need the parallel composition of channels. For two channels \( c_1 \) and \( c_2 \) we would like the parallel composition \( c_1 \parallel c_2 \) to satisfy \( s \models c_1 \parallel c_2 \rightarrow t \) iff \( s \models c_1 \rightarrow t \)
and \( s \xrightarrow{c_2} t \). Then it is clear that if \( s_1 \xrightarrow{c_1} \rightarrow A \rightarrow B \) and \( s_1 \xrightarrow{c_1} \rightarrow A' \rightarrow B' \) then \( s_1 \xrightarrow{c_1|c_2} \rightarrow A \land A' \rightarrow B \land B' \).

**CASES:** If \( s_1 : A \) carries the information that \( s_2 : B \lor B' \) and \( s_2 : B \) carries the information that \( s_3 : C \) and \( s_2 : B' \) carries the information that \( s_3 : C \) then \( s_1 : A \) carries the information that \( s_3 : C \).

Again using parallel composition, if \( s_1 \xrightarrow{c_1} \rightarrow A \rightarrow B \lor B' \), \( s_2 \xrightarrow{c_2} \rightarrow B \rightarrow C \) and \( s_2 \xrightarrow{c_2} \rightarrow B' \rightarrow C \), then \( s_1 \xrightarrow{c_1(c_2|c_2)} \rightarrow s_3 \equiv A \rightarrow C \).

Now let’s consider what might count as a channel in our applications. The first one, traditional situation theory, we’ve already discussed as the motivating examples of channel theory as considered by Barwise and Seligman. So consider intuitionistic logic. There, recall the clause for the conditional:

\[
\vdash A \supset B \text{ iff for all } t \geq s, \text{ if } t \models A \text{ then } t \models B.
\]

Here, we only have the fact that \( t \models A \) tells us that \( t \models B \). This is not particularly interesting. However, we can modify it somewhat as follows:

\[
\vdash A \supset B \text{ iff for all } t \models A, \text{ if } u \geq s, \text{ if } t \models A \text{ then } u \models B.
\]

So here we have a primitive channel relation: each point \( s \) is a channel, connecting points \( t \) to the common ‘descendants’ of \( s \) and \( t \). So, a pseudo signal to a channel \( s \) in this context will be another point \( t \) which has no common descendants with \( s \). In that case, even though we may have \( s \models A \supset B \), if \( t \models A \), this tells us nothing about any other point, for no points descended from \( t \) are also descended from \( s \).

An interesting fact here is that the channels are on the same level as the ‘situations’. The points in the model structure themselves give rise to connections among themselves. We will see more of this later. Another interesting fact is that these are exceptionless channels: there is no notion of channel failure, except for the possibility of pseudo signals. And in intuitionistic logic, there is no chance of avoiding pseudo signals without strengthening the logic. Were we to demand that any two points have a common descendant, then this extends the logic with the axiom \( \sim A \lor \sim \sim A \), an intuitionistic non-theorem.

In the semantics of modal logic, things are different. Given a binary
accessibility relation $R$ it is quite reasonable to take it to be a channel. We set $x \xrightarrow{R} y$ iff $xRy$. But then, what is it for this channel support the constraint $A \rightarrow B$?

You might at first think that this will have something to do with strict implication, $\Box (A \supset B)$, but this is not the case, for strict implication does not indicate a link between what goes on in different possible worlds. That $A$ strictly implies $B$ (at world $x$) shows us only that at every world accessible from $x$, if $A$ is true in that world, then so is $B$. Strict implication is a static notion indicates truth preservation at a class of worlds. To look at what is transferred from world to world by the relation $R$ we must use some other notion. Expanding the definition, we see tha $R \models A \rightarrow B$ iff for every $x$ where $x \models A$, if $xRy$, then $y \models B$. In other words, if $x \models A$, we have $x \models \Box B$. So $R \models A \rightarrow B$ is equivalent to $\forall x(x \models A \supset \Box B)$. Then we can compose this channel with itself, by setting $R^2$ to be defined in the usual way: $x \xrightarrow{R^2} y$ iff there’s a $z$ where $xRz$ and $zRy$. Then $R^2 \models A \rightarrow B$ iff $\forall x(x \models A \supset \Box \Box B)$, and obviously, if $R \models A \rightarrow B$ and $R \models B \rightarrow C$, then $R^2 \models A \rightarrow C$. Then if we are in a multimodal setting (with more than one accessibility relation) we can consider different channels given by composing these relations, and furthermore, there is scope for allowing the identity channel (which models entailment) and converse relations (which utilise the ‘backward looking’ modal operators).

The semantics of relevant logics are very similar to that of intuitionistic logic, except that here, the channeling relation is quite explicit. Here, a point $x$ is a channel between $y$ and $z$ just when $Rxzy$. Again, channels are exceptionless, but as before, they may have pseudo signals.

The frames for the Lambek calculus are even simpler. Here, each syntactic string $x$ gives rise to a channel, connecting strings $y$ to other strings $xy$ given by concatenating $x$ with $y$. But this account is novel. There is no pseudo signal for any channel. Each channel takes one input and gives one output. However, in the frames for the Lambek calculus, we have two different ‘conditionals’, \ and /.

We can see these as giving rise to two different sorts of channels. Firstly, $x$ maps $y$ to $xy$, but on the other hand, it maps $y$ to $yx$. How are we to think of this? One way is to revive a notion mentioned in one of Seligman’s early articles, “Physical Situations and Information Flow” [46], the notion of a conditional channel. Seligman gives an example of a conditional channel: the phone connection between his office phone and my mobile. Here, there is a connection between these situations, only when the environment is correct. So,
there is a connection between his office phone and my car phone in the context of a background environment situation. So, a channel of this kind gives rise to a ternary relation between situations: environments, signals and targets.

How does this help in the case of the Lambek calculus? Using the notion of a conditional channel see the frame as giving rise to just two conditional channels, Left and Right. Left takes \(x\) as environment and \(y\) as signal to output \(xy\). Right takes \(x\) as environment and \(y\) as signal to output \(yx\). The constraints supported by Left are those expressed by the conditionals of the form \(B/A\). Those supported by Right are those expressed by conditionals of the form \(A/B\). In this way we can distinguish the two sorts of conditionals of the Lambek calculus as those supported by different sorts of connections between points in the structure. But to do so, we need to use the notion of a conditional channel.

This notion of a conditional channel makes sense in the semantics of relevant logics too. For we can consider a frame of points with a whole family of ternary relations, each expressing a different kind of connection between signal and target points in some environment. Then, each ternary relation would have a corresponding conditional, which would express a particular sort of connection.

Finally, map semantics. Here, there is some interesting work being done already, so I will simply indicate its main thrust [31]. The central idea is that channels don’t do much good relating map regions to other map regions, as they aren’t particularly interestingly connected. The best connections to look at are those between map regions and the environment being mapped, over time. In this way, channel theory gives you a nice taxonomy to understand different sorts of mapping errors and the ways in which we can extract information from maps.

Much more work needs to be done on developing applications of channel theory. Some of these applications are found in Barwise and Seligman’s book Information Flow [9]. I hope I have given you enough get a taste of this area, to help you understand the literature, and more importantly, develop applications of your own.


[31] Oliver Lemon and Ian Pratt. Putting channels on the map: A formal semantics for graphical information systems. Computer Science, University of Manchester. Email lemonog@cs.man.ac.uk, 1996.


