5 Situations, Constraints and Channels

(Update of Chapter 4)

Edwin Mares[*], *Jeremy Seligman*[†], *Greg Restall*[‡]

[*]Department of Philosophy, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand, E-mail: Edwin.Mares@vuw.ac.nz
[†]Department of Philosophy, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand, E-mail: j.seligman@auckland.ac.nz
[‡]Department of Philosophy, School of Philosophy, Anthropology and Social Inquiry, The University of Melbourne, Old Quad, Parkville, Victoria 3010, Australia, E-mail: restall@unimelb.edu.au

5.1 From Situation Semantics to Situation Theory

Central to the project of situation semantics was the goal of a 'relational' theory of meaning, which would explain meaning in terms of the relationship between situations containing meaningful entities or actions such as utterances and situations they are about (Barwise and Perry, 1983). The contrast is primarily with the once dominant view of Davidson, Montague, and many others following Tarski's seminal ideas about semantics, according to which the meaning of a declarative statement (at least) is to be understood in terms of the conditions under which it is true. There are some difficulties in making this contrast clear. After all, truth-conditional theories of meaning typically also involve a theory of reference, which is concerned specifically with the relationship between words and things, and the relation of reference is present in all attempts to produce a situation-based semantics. Likewise, truth has been studied within a situation theoretic framework, most notably by Barwise and Etchemendy (1987), in their treatment of the Liar paradox. Dressing up a truth-conditional account of meaning in the notation of situations and infons is a futile exercise.

The important difference is in the theoretical status of semantic vocabulary. For Tarski, semantics involves a clear separation between syntax and semantics, and this separation has been honoured by most of his followers. By contrast, situation semantics aimed to do without this separation, taking reference, for example, to be a relation like all others, and with no special theoretical status. That 'Jon' refers to Jon is just a fact to be modelled as $\langle\langle \mathsf{refers}, \text{'}\mathsf{Jon}\text{'}, \mathsf{Jon}; 1\rangle\rangle$. A consequence of representing semantic facts in the object language is that there is no need for a hierarchical theory of meaning, on which the meaning of an expression is some unitary theoretical entity, such as a truth-condition, derived from its more basic semantic properties. Instead, the many facets of meaning can be left unassembled and ready to be used for whatever purpose is required.

Handbook of Logic and Language. DOI: 10.1016/B978-0-444-53726-3.00005-0

As a consequence, many aspects of semantic theory that are usually treated safely in the metalanguage need to be made explicit. For example, the relation between a conjunction $\sigma \wedge \tau$ and its conjunct σ can be represented as one of 'involvement', so that $\langle\langle \mathsf{involves}, \sigma \wedge \tau, \sigma; 1\rangle\rangle$ is a fact of the same genus if not the same species as $\langle\langle \mathsf{involves}, \mathsf{smoke}, \mathsf{fire}; 1\rangle\rangle$. Moreover, a theory is needed to explain the way in which information about 'involvement' works. Some account must be given of what it is for a situation to support infons of this kind. In the case of natural regularities, such as the relationship between smoke and fire, an obvious thought is that the localisation to a situation can accommodate the defeasible nature of the regularity. In some situations, smoke really does indicate the presence of fire, in others it does not. For logical laws, such as Conjunction Elimination, no such localisation is necessary, and yet to say that every situation supports the infon $\langle\langle \mathsf{involves}, \sigma \wedge \tau, \sigma; 1\rangle\rangle$ conflicts with the idea that situations are informationally limited. A situation concerning a cricket match in New Zealand may not contain any information about the weather in Bangkok, so it would be unfortunate if it supported the infon $\langle\langle \mathsf{involves}, \langle\langle \mathsf{hot\ and\ humid}, \mathsf{Bangkok}; 1\rangle\rangle, \langle\langle \mathsf{hot}, \mathsf{Bangkok}; 1\rangle\rangle; 1\rangle\rangle$.

It is to deal with issues such as these that a general theory of constraints is required. Constraints, such as 'involvement', are relations between infons and possibly other parts of the situation theoretic universe, that make explicit the regularities on which the flow of information depends. The replacement of truth by 'support' as the fundamental theoretical concept of situation theory succeeds in localising information but at the expense of opening an explanatory gap: how is it that information in one situation is related to information in another situation?

5.2 Early Channel Theory

One approach to the flow of information stands to the 'constraint' view of the last section as an Austinian account of propositions stands to a Russellian account. It is one thing to consider constraints as true propositions: generalisations or relations between situation types: one can go quite some way given that approach. The development of channel theory in the 1990s marked a new approach.

Consider this example: a student looks at her marked assignment, and the fact that this perceptual situation carries some information (the 'A' written in red ink in the top corner of the first page) gives her the information that her tutor thought her assignment was a good one, that she's passed her class, that she's completed her degree, and that she's likely to get a job. The information carried in *this* situation gives her information about *other* situations: the marking situation for one, and her future prospects, for another. How are we to take account of this? Barwise (1993), in his paper "Constraints, Channels, and the Flow of Information" marks out a few desiderata concerning information carrying *across* situations.

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 $s_1 : A \wedge A'$ carries the information that $s_2 : B \wedge B'$.

cases: If $s_1 : A$ carries the information that $s_2 : B \vee B'$ 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 $s_3 : C$.

In these principles, we flag the manner in which information about a situation can carry information *about some other situation*. 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 first-order information. These links may be present or absent, distinguishing veridical information flow from cases of misinformation.

Channels are means of connection between situations. That a channel c links s and t is denoted '$s \overset{c}{\mapsto} t$'. An example of a channel given by Barwise and Seligman (1994) in "The Rights and Wrongs of Natural Regularity" is the Rey Channel, linking thermometer-reading situations with patient situations. The fact that the thermometer's mercury level has a particular height usually indicates something about the temperature of a patient. The channel grounds the regularity connecting 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 $s \vDash A$, and the channel c supports a regularity of the form $A \to B$ (if the height is x then the temperature is y) then given that the situation t is connected to s by the channel c ($s \overset{c}{\mapsto} t$) we can infer $t \vDash B$. In $s \overset{c}{\mapsto} t$, s is a *signal* for the channel c and t is a target. A channel c supports the *constraint* $A \to B$ if and only if for each signal-target pair s and t where $s \overset{c}{\mapsto} t$, if $s \vDash A$ then $t \vDash B$.

An important feature of channel theory is the presence of multiple channels, in just the same way as multiple situations feature in situation theory. Information flows in more than one way—it is not just a matter of physical law, or convention, or logical entailment. The Rey channel is partly a matter of physical law, but it is also a matter of convention. 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). 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.

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 say that the thermometer situation s 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.[1]

Now we may consider a channel-theoretic elaboration of principles of information flow.

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 $c_1; c_2$ which composes c_1 and c_2, satisfying $s \xmapsto{c_1;c_2} t$ iff there's a u where $s \overset{c_1}{\mapsto} u$ and $u \overset{c_2}{\mapsto} t$. Then it is simple to show that if $s \overset{c_1}{\mapsto} u \models A \to B$ and $u \overset{c_2}{\mapsto} t \models B \to C$ then $s \xmapsto{c_1;c_2} u \models A \to C$.

Here, $c_1; c_2$ 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 $A \vdash B$ is cashed out as "for each s, if $s \models A$ then $s \models B$", then A entails B iff $1 \models A \to B$.

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 $s_1 : A \wedge A'$ carries the information that $s_2 : B \wedge B'$.

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 $s \xmapsto{c_1 \parallel c_2} t$ iff $s \overset{c_1}{\mapsto} t$ and $s \overset{c_2}{\mapsto} t$. Then it is clear that if $s_1 \overset{c_1}{\mapsto} s_2 \models A \to B$ and $s_1 \overset{c_2}{\mapsto} s_2 \models A' \to B'$ then $s_1 \xmapsto{c_1 \parallel c_2} s_2 \models A \wedge A' \to B \wedge B'$.

[1] Restall (1995) discusses the behaviour of counterfactuals in a channel-theoretic setting in "Information Flow and Relevant Logics".

cases: If $s_1 : A$ carries the information that $s_2 : B \vee B'$ 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 $s_1 \overset{c_1}{\mapsto} s_2 \vDash A \rightarrow B \vee B'$, $s_2 \overset{c_2}{\mapsto} s_3 \vDash B \rightarrow C$ and $s_2 \overset{c_3}{\mapsto} s_3 \vDash B' \rightarrow C$, then $s_1 \xmapsto{c_1;(c_2\|c_3)} s_3 \vDash A \rightarrow C$.

Models in which constraints stand to channels as infons stand to situations immediately brought to light new connections between situation theory and other areas. Restall (1995) showed that if we identify channels with situations, then any model of the conditions on channels is a model of Routley and Meyer's ternary relational semantics for relevant logics. The channel theoretic reading for $y \overset{x}{\mapsto} z$ is another way to conceive of the three-place relation $Rxyz$ on points in a ternary relational model, and the clause for the relevant conditional:

$$x \vDash A \rightarrow B \text{ iff for each } y, z \text{ where } Rxyz \text{ if } y \vDash A \text{ then } z \vDash B$$

is merely a *flattenning* of the channel-theoretic account, in which channels are brought down to the level of situations. In this way, one is free to think of points in a model for relevant logic as situations.[2] Situations not only are a site for the carrying of information purely about what is contained in them: they may also constrain or maintain connections between *other* situations. Given this perspective, different conditions on the three-place relation R correspond to different ways in which the topography of those connections are to be understood.

In a similar fashion, Barwise et al. (1996) showed that a generalisation of the Lambek calculus can be conceived of in a channel theoretic manner (and given a nice two-level cut free sequent system), in which the traditional Lambek calculus is recovered if the picture is flattened, and channels and situations are identified in the same manner.[3]

This work on channel theory through to the mid-1990s was, it must be said, a transitional phase. A greater level of generality was reached with the publication of Barwise and Seligman's *Information Flow*: *the logic of distributed systems* (1997).

5.3 Situated Inference

A somewhat different way of relating relevant implication to channels is developed in the theory of situated inference in Mares (2004). This theory is a descendent of David Israel and John Perry's theory of information (1990). In the theory of situated inference, relevant implication represents constraints, combinations of constraints, and the logical manipulation of these constraints.

[2] This is already congenial to the relevantist, for points in Routley–Meyer models may be incomplete with respect to negation, just as situations are. Relevant models typically also allow for *inconsistent* points, which are perhaps a little more difficult to motivate from purely situation-theoretic considerations.

[3] For more ways to interpret traditional logical structures, such as accessibility relations in a multimodal frame, see Restall (2005).

The key distinction in this theory is between worlds and situations and the central notion underlying the theory of inference is that of a constraint, in the sense we have already seen. Constraints themselves can be present as information in situations. Here is a rather blatant example of a situated inference. Suppose that an agent is in a situation in which there is the constraint that all massive bodies, heavier than air, released near the surface of the earth fall towards the earth. From the hypothesis of the existence of an actual situation (i.e. a situation in her world) in which a massive body is released near the surface of the earth, she can infer that there is an actual situation in which that body falls towards the earth. But, like channel theory, the theory of situated inference does not require that constraints be general in this way. They may concern particular objects or circumstances.

The theory of situated inference deals with misinformation in a different way from early channel theory. In early channel theory, a situation in which there is a certain sort of information and one in which there fails to be that sort of information (but appears the same as the first situation) is explained by the presence or absence of a second situation—a channel. In the theory of situated inference, the difference is between two situations that have different sorts of information actually available in them. This difference is not caused by a deep philosophical disagreement over the nature of information. Rather, the two theories have different purposes. Early channel theory is meant to explain information flow, whereas the theory of situated inference is meant to give a theory of deductive warrant for inferences made with partial information. But the idea that constraints be available as information to agents in situations means that the notion of availability used here has to be sufficiently general. We clearly do not have perceptually available to us many of the constraints that we use in inferring. Other sorts of reliable causal processes must be allowed to be counted as making information available to us if this theory is to be viable.

In Mares (2004), the theory of situated inference is used to provide an interpretation of Alan Anderson and Nuel Belnap's natural deduction system for relevant logic and we use this natural deduction system here to make the theory clearer.[4]

At a line in a derivation in this system, we not only have a formula, but a formula subscripted with a set of numbers. These sets, on the theory of situated inference, refer to situations. Thus, for example, if we have $A_{\{1\}}$ at a line of a proof, this is to be read as 'A is satisfied by a situation s_1' or '$s_1 \models A$'. So that we can see how the subscripting mechanism works, let's look at a proof of $A \vdash (A \to B) \to B$:

1.	$A_{\{1\}}$	*hyp*
2.	$\quad A \to B_{\{2\}}$	*hyp*
3.	$\quad A_{\{1\}}$	1, *reit*
4.	$\quad B_{\{1,2\}}$	2, 3, $\to E$
5.	$(A \to B) \to B_{\{1\}}$	$2-4$, $\to I$

[4] Situated inference is also used in Mares (2004) to give a reading of a model theory like the one discussed in the section on early channel theory in which implication is modelled using a three-place relation. But this interpretation is too involved to be included in the current chapter.

When we make a hypothesis in a proof, it takes the form of the formula hypothesized subscripted with a singleton of a number not used before in the proof. When we use the rule of implication elimination (as in line 4 of this proof), the subscript of the conclusion of the rule is the union of the subscripts of the premises. When we discharge a hypothesis (as in line 5 of this proof) we remove the number of the hypothesis from the subscript of the conclusion. The hypothesis' number must be present in the last line before the discharge in order to allow the discharge to take place. In this way we ensure that the hypothesis is really used in the subproof, and this ensures that the resulting logic is a *relevant* logic. In this way we represent informational connections and we read $A \rightarrow B$ as "A carries the information that B".

A formula subscripted by a set with more than one number in it, e.g. $B_{\{1,2\}}$ from line 4 of our derivation, is to be read as saying that on the basis of the hypotheses about s_1 and s_2, we can derive that there is a situation in the same world which satisfies B.

The treatment of conjunction in the theory is also interesting. In relevant logic, there are two types of conjunction: extension conjunction ($\wedge$) and intensional conjunction or "fusion" ($\circ$). The introduction rules for the two connectives make the difference clear. From A_α and B_β we may infer $A \circ B_{\alpha\cup\beta}$. With regard to extensional conjunction, on the other hand, the subscripts on two formulas must be the same before we can conjoin them: from A_α and B_α we may infer $A \wedge B_\alpha$. In situated terms we can explain clearly the difference in meaning between these two connectives. If we are in a situation in which we have the information that $(A \wedge B) \rightarrow C$, then we have warrant to infer, on the basis of the hypothesis of an actual situation which satisfies *both A and B*, that there is an actual situation which satisfies C. If, on the other hand, we have the information that $(A \circ B) \rightarrow C$, then we have warrant to infer, on the basis of the hypothesis of an actual situation which satisfies A *and* the hypothesis of an actual situation that satisfies B, that there is an actual situation that satisfies C.

The reason for introducing fusion here is to make a point about relevant (and situated) inference. The premises of a relevant inference are bound not by extensional conjunction, but by fusion. In a situation s, the postulation of situations $s_1 \models A_1, \ldots, s_n \models A_n$ deductively warrants the postulation of $s_{n+1} \models B$ if and only if $s \models (A_1 \circ \cdots \circ A_n) \rightarrow B$. The natural deduction rules for fusion bear out this close relationship to situated inference. The introduction rule tells us that we are allowed to infer from A_α and B_β to $A \circ B_{\alpha\cup\beta}$ and the elimination rule tells us that we may infer from $A \circ B_\alpha$ and $A \rightarrow (B \rightarrow C)_\beta$ to $C_{\alpha\cup\beta}$. These rules together with the rules for implication allow us to prove the equivalence of $A_1 \rightarrow (A_2 \rightarrow \cdots (A_n \rightarrow B) \cdots)_\alpha$ and $(A_1 \circ A_2 \circ \cdots \circ A_n) \rightarrow B_\alpha$, which is a situated version of the deduction theorem.

This ends our discussion of situated inference. In the next section we examine a more dynamic view of the relationship between constraints and information.

5.4 Modern Channel Theory

A further development of the theory of channels introduces the distinction between tokens and types into the model. Logical relations of entailment and contradiction are formal: they depend only on the form of the propositions they relate. This applies even to those entailments that depend on the meaning of non-logical words: the inference from 'Jon is taller than Jerry' to 'Jerry is shorter than Jon' is a formal inference from 'X is taller than y' to 'y is taller than x'. As such, the distinction between token and type, between a specific use of these sentences and the sentence types, is irrelevant. Informational dependencies between many events do not share this indifference to tokens. Whether or not an observed plume of smoke indicates the presence of a nearby fire depends a great deal on the circumstances, on the particularities of the occasion. This dependence on the specific circumstances can be modelled by a type. If B (for 'background condition') is the type of situation in which smoke really does mean fire, then we can take the relation between smoky events and fiery events to be mediated by a channel of type B. We have seen how this idea can be developed in the previous sections. Another way is to get serious about tokens, bringing them into the model explicitly. This continues the situation theoretic methodology of internalising metalinguistic concepts to advance the theory. Before we can describe the resulting model of channels, it is therefore necessary to say something about the results of applying a type-token distinction throughout the underlying theory.

In fact, the type-token distinction is already part of situation theory in the relationship between situations and the set of infons they support. In early work on situation theory, this was referred to as the distinction between real situations and abstract situations. It had little effect on the development of the theory because of the principle of extensionality, which identifies situations that support the same infons, and so ensures that there is only one situation of each basic type.[5]

5.4.1 Classifications and Local Logics

The way in which token entities (situations, events, objects or whatever) are categorized into types depends on the method of classification used. Different languages, for example, give us different ways of grouping objects together. In early work on situation theory, the basic divisions into objects, properties and relations, was called a *scheme of individuation*. This can be regarded as the result of putting the signature of the language into the model, with the possibility that there can be more than one. The multiplicity of schemes allows not only for radically different ways of conceiving of the world by agents from different cultures (or species) but for the more mundane fact that we adapt our conceptual repertoire according to context, employing only those distinctions that are salient and/or useful. Seligman (1990) took the typing relations between specific situation tokens (called *sites*) and their types as the primary objects of study, called *perspectives*, allowing multiple typing relations to account for the possibility of different schemes of individuation. In addition to classifying tokens into

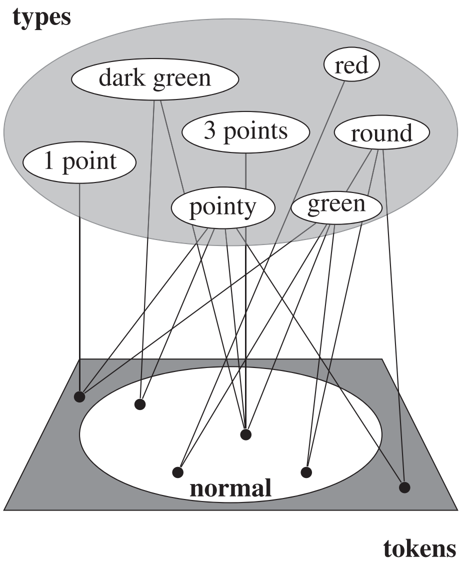
[5] The one situation one type rule is violated by the development of a hierarchy of complex situation types, as described in Seligman and Moss (2010), pp. 171–244, but the set of types of a situation is still fully determined by this one type, its principal type.

types, the perspectives we adopt typically impose a logical structure. If we classify according to colour, for example, we allow two different objects to be classified as being of the same colour but disallow one object to be classified as being of two different colours. Yet, if we use names to classify, only one object may be classified by a name but two names may be used to classify the same object. This was modelled by means of an *involves* relation and a *precludes* relation between types.

These ideas were developed by Barwise and Seligman (1997) into the theory of local logics. A classification A is simply a binary relation $\models_A$ between a set $\mathsf{tok}(A)$ of *tokens* and a set $\mathsf{typ}(A)$ of *types*. A *theory* T is a set $\mathsf{typ}(T)$ of types together with a binary relation $\vdash_T$ between subsets of T. When the relation $\vdash_T$ models some constraints on classification in A, we can understand $\Gamma \vdash_T \Delta$ as a prohibition on classifying tokens to have all the types in Γ and none of the types in Δ. For example, we can represent a mutually exclusive and exhaustive classification using colour terms as a theory with types $C = \{\text{red, yellow, green, blue}\}$ such that $\emptyset \vdash C$ and $\alpha, \beta \vdash \emptyset$ for each pair α, β of distinct colour terms. Allowing the possibility that some tokens fail to be constrained, a *local logic* L is a classification $\mathsf{cla}(L)$ together with a theory $\langle \mathsf{typ}(L), \vdash_L \rangle$ and a set N_L of *normal* tokens such that if $\Gamma \vdash_L \Delta$ then there is no normal token of all the types in Γ and none of the types in Δ. In other words, the classification of normal tokens is required to respect the constraints implicit in the theory.

The paradigm example of a local logic is obtained from a theory T expressed in a language L and a class M of models of T. The tokens are the models of L, which are classified into types, the sentences of L, by the relation of satisfaction.[6] The theory is the consequence relation $\Gamma \vdash_T \Delta$ iff $\Gamma, T \vdash \Delta$ and the set of normal tokens is M. The example is atypical in having a token that is a counterexample to every inference not licensed by T. Local logics used to model classification by humans or other finite agents are unlikely to have this property. For example, suppose that some primary school children have been asked to classify some leaves. Each child is required to think of his or her own way of classifying. Most children use ‘colour’ as a classifying attribute but they select different values (green, dark green, brown, yellow, etc.). Some classify by the number of ‘points’, others use ‘rounded’ or ‘pointy’. In producing their classifications, the children obey various constraints, either explicitly or implicitly, some self-imposed, some deriving from language. The exclusivity of ‘pointed’ and ‘rounded’ may be partly linguistic, partly a matter or conscious choice. We can model these constraints as a theory on the set of classificatory types they use. Occasionally, a leaf is discovered that violates some of these constraints—one that is both pointy and rounded, for example. This is classified but flagged as strange, and we model it as a token that lies outside the set of normal tokens. Moreover, unlike the local logic obtained from a formal theory, there may be invalid inferences, such as the inference from being red and pointy to having three points, but without the child having found

[6] Strictly speaking, either the definition of classification should permit proper classes of tokens or we should choose a representative set of models, rather than the class of all models, as tokens in this classification. Nothing of present concern hangs on the difference, so we equivocate.



1 point $\vdash$ pointy
3 points $\vdash$ pointy
pointy, round $\vdash$
red, green $\vdash$
red, dark green $\vdash$

Figure 5.1 A local logic for classifying leaves.

a red pointy leaf. Even if there are red pointy leaves in the token set and they all have three points, this may be a regularity of the sample, not one that is part of the child's classificatory method.

A third example of a local logic, also paradigmatic, is the observed behaviour of a dynamic system which at any instance occupies one of a number of distinct states. We model each instance of the system as a token in the token set, which may not include tokens of all possible states but only the ones that have been observed. Our observations are rarely capable of discriminating individual states, so observation types correspond to a limited selection of subsets of the state space.[7]

5.4.2 Channels Defined Algebraically

In a setting in which there are many local logics, it is important to see how they are related to each other, how they can be combined and modified in systematic ways. As is usual in algebra, this can be done by determining a class of transformations

[7] To be a little more precise, we can derive a local logic from a probabilistic model of the behaviour of a system with outcome set Ω, event set Σ a σ-algebra, a probability measure $\mu\colon \Omega \to [0, 1]$, a set T of times at which the system is observed and a function $s\colon T \to \Omega$ specifying the state of the system at each time. Instances of the system are modelled by the set T, so we can take this as the set of tokens. The set of types is Σ, with $t \models e$ iff $s(t) \in e$. The entailment relation $\vdash$ can then be defined probabilistically, as $\Gamma \vdash \Delta$ iff all counterexamples have zero probability, i.e. there is no event e with $\mu(e) > 0$ such that both $e \subseteq \bigcap\Gamma$ and $\bar{e} \subseteq \bigcap\Delta$. Aberrant behaviour of the system, perhaps due to outside influences or initial conditions can be marked as non-normal. This example does not quite conform to the framework of Barwise and Seligman (1997) because of the probabilistic consequence relation, which requires the additional work of Seligman (2009).

from one local logic to another. In Barwise and Seligman (1997), an *infomorphism* f from classification A to classification B is defined as a pair of contra-variant functions $f^\wedge\colon \mathtt{typ}(A) \to \mathtt{typ}(B)$ and $f^\vee\colon \mathtt{tok}(B) \to \mathtt{tok}(A)$ which preserves the classification relation, in that $f^\vee(b) \models_A \alpha$ iff $b \models_B f^\wedge(\alpha)$ for each token b of B and each type α of A. Some examples of transformations of classifications within this class are the restriction to a subset of tokens, the addition of new types, the splitting of a token into two tokens and the identification of two types which have the same tokens. In all these cases, the infomorphism records a correlation between the two classifications that survives these modifications. Infomorphisms form a Cartesian closed category and so a variety of standard algebraic operations, such as quotients, products and sums can all be defined, giving a rich theory of what can be done when comparing different classifications and constructing new classifications from old.[8] All of this can be easily extended to local logics by requiring the transformations also to preserve the theory and set of normal tokens; such transformations are called *logic-infomorphisms*. The example of local logics constructed from theories and models provides a good illustration of the general idea: an infomorphism from the local logics of $\langle T_1, M_1\rangle$ to that of $\langle T_2, M_2\rangle$ is an interpretation of theory T_1 in theory T_2 together with a transformation of models in M_2 into models in M_1 that preserves the satisfaction relation in the obvious way.

In particular, the algebra of classifications provides a way of modelling relations between classifications, which is exactly what is required for a theory of channels. In any category, the concept of a binary relation between two objects, X and Y, can be represented as a pair of transformations from a third object R to X and to Y. In the category of sets, R is just the set of pairs of related elements and the transformations are the projection functions, which when a is related to b, take the pair $\langle a, b\rangle$ to a and to b, respectively. In the category of classifications, transformations are infomorphisms and so we model a channel $\mathcal{C}$ between classification A (the *source*) and classification B (the *receiver*) as a pair of infomorphisms: $s_{\mathcal{C}}$ from C to A and $r_{\mathcal{C}}$ from C to B. The classification C is called the *core* of the channel. We think of a token c of C as the specific connection within the channel that relates source token $s^\vee_{\mathcal{C}}(c)$ to receiver token $r^\vee_{\mathcal{C}}(c)$.

To take an example from communication theory, if messages about one system (the source) are sent along a communication channel such as a telegraph wire to influence the behaviour of another system (the receiver), then we can model the channel core as a classification of token-states of the wire with infomorphisms mapping each wire token to the corresponding token-states of the source and receiver, as shown in Figure 5.2. The source types and receiver types are also both mapped to corresponding types of the channel, which serves to model the way in which information flows from source to receiver.

[8] Basic properties of the categories of classifications and local logics, together with various applications, are explored in Barwise and Seligman (1997). But Goguen (to appear) shows these to be special cases of a much broader class of categories, which have been extensively studied in computer science.

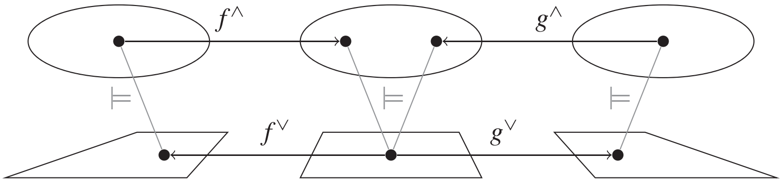


Figure 5.2 The telegraph channel.

Information flow along channels is modelled using local logics. Given an infomorphism f from classification A to classification B and any local logic L 'living' on A (one that has A as its underlying classification), we can define the image fL of L under f, which is a local logic living on B. Likewise, any logic L' living on B has an inverse image $f^{-1}L'$ on A.[9] Local logics living on the same classification are ordered the obvious (contra-variant) inclusions: $L_1 \leq L_2$ iff $\vdash_{L_1} \subseteq \vdash_{L_2}$ and $N_{L_2} \subseteq N_{L_1}$. The ordering is a complete lattice, so logics may be combined by meets and joins. In this way information distributed around a network of channels may be moved from one classification to another (by imaging and inverse-imaging) and combined using joins and meets.[10]

Naturally, channels compose sequentially: if there is a channel $\mathcal{C}_1$ from A to B and another channel $\mathcal{C}_2$ from B to D, they can be combined to give a channel $\mathcal{C}_1; \mathcal{C}_2$ from A to D. This is just an application of the fact that the category of infomorphisms is Cartesian closed. There is also a parallel composition $\mathcal{C}_1 \bullet \mathcal{C}_2$ of channels $\mathcal{C}_1$ and $\mathcal{C}_2$ having the same source and receiver. To characterise these constructions precisely, we need the concept of a *refinement* of one channel by another—this is just an infomorphism between the channel cores that commutes with the source and receiver infomorphisms of the two channels, ensuring that any information that flows in one channel also flows in the refined channel. The two compositions $\mathcal{C}_1; \mathcal{C}_2$ and $\mathcal{C}_1 \bullet \mathcal{C}_2$ each provide the least refined channel that 'agrees' with the component channels, in the sense of commuting with the source and receiver infomorphisms.[11]

These constructions can be extended to whole networks of channels, if we generalise the concept of a channel to allow multiple sources and targets. In fact, there is no need to make a distinction between source and target, as information flows in both directions. In general, then, a channel is modelled as a set of infomorphisms

[9] More precisely, fL is the least logic on B that makes f a logic-infomorphism from L to fL, while $f^{-1}L'$ is the greatest logic on A that makes f a logic-infomorphism from $f^{-1}L'$ to L'.

[10] The additional structure of local logics is essential for modelling information flow across channels in a flexible way. Earlier attempts, such as Barwise (1993) and Barwise and Seligman (1994), focussed too closely on the relation between types. But it is logical structure—entailment, contradiction, etc.—that is transformed by infomorphisms, not propositional content. The calculus of information flow is therefore a calculus of logical structure not a calculus of propositional content.

[11] One can also regard the composite channels as channels between the cores of the component channels, modelling the information flow between them.

with a common domain (the channel core). Each set of channels has a common refinement, a generalised channel with projections to all the classifications in the set, which models the common information flow in the system. Similar constructions can be performed within the the models of constrained classification using local logics and logic-preserving infomorphisms.

5.4.3 Connections to Modal Logic

The logic of constraints has been investigated by van Benthem in a number of different guises (van Benthem, to appear, 2000). A simple language for reasoning about constraints has atomic propositions $Ts_1, \ldots, s_n$, where $s_1, \ldots, s_n$ are situations, from some set *Sit*, and T is an n-ary situation type, from some set $Type_n$. To this, we add Boolean operations and a modal operator $[S]$ for each set of situations S, with $[S]\varphi$ meaning that φ is determined by the information shared by situations in S.[12] The language allows one to state that situation s_1 being of type T_1 carries the information that situation s_2 is of type T_2, using the formula

$$[Sit](T_1 s_1 \supset T_2 s_2)$$

In other words, the implication from s_1 being of type T_1 to s_2 being of type T_2 holds in every situation. When we are restricted to a set B of situations, this is modified to

$$[B](T_1 s_1 \supset T_2 s_2)$$

A *constraint model* $M = \langle State, C, V \rangle$ for this language consists of a set *State* (of *local states*), a set C of *global states*, which are functions from *Sit* to *State*, and a function V assigning an n-ary relation on *State* to each type in $Type_n$. Formulas are evaluated in a global state w as follows:

$$\begin{array}{lll} M, w \models Ts_1, \ldots, s_n & \text{iff} & \langle w(s_1), \ldots, w(s_n) \rangle \in V(T) \\ M, w \models [S]\varphi & \text{iff} & M, v \models \varphi \text{ for all } v \in C \text{ such that} \\ & & v(s) = w(s) \text{ for all } s \in S \end{array}$$

with the standard clauses for the Boolean operators. The resulting logic is decidable and is axiomatised in van Benthem (to appear) as a polymodal S5, with additional axioms $([S_1]\varphi \supset [S_2]\varphi)$ for each $S_2 \subseteq S_1$.[13]

[12] We can ignore the distinction between situations/situation types and their names because the language has no need for quantifiers over situations. We also place no restrictions on the size of the set of situations, although in practice it must be finite if we are to obtain a recursively enumerable language.

[13] van Benthem has an additional operator U and axioms $(U\varphi \supset [S]\varphi)$ for each S, but U can be defined to be $[Sit]$. As an interesting technical aside, he also notes that constraint logic is equivalent (mutually embeddable) in the polyadic logic of dependency, see van Benthem (2000), in which the standard semantics for first-order logic is modified only by restricting the set of assignment functions to a limited set G, and then introducing a quantified formula $\exists \boldsymbol{x}\, \varphi$ for each sequence $\boldsymbol{x}$ of variables, which is satisfied by an assigment g if there is an assignment $g \in G$ that is identical to g except possibly on the variables in $\boldsymbol{x}$.

To this simple language, the apparatus of standard epistemic logic and dynamic epistemic logic can easily be added. Given a set I of agents, we can add a modal operator K_i representing the knowledge of agent i, for each $i \in I$, operators $[e]$ for each event e in some set E, and public announcement operators $[!\varphi]$ for each formula φ. The operators K_i and $[e]$ each require a binary relation on the set C of global states and public announcement is defined by model restriction: $M, g \models [!\varphi]\psi$ iff $M, g \models \varphi$ and $M|_\varphi, g \models \psi$, where $M|_\varphi$ is the restriction of M to those global states that satisfy φ in M. The resulting logic is still decidable, as shown in van Benthem (to appear).

Logic can also be used to characterise information flow along an infomorphism. Consider the two-sorted first-order language for describing the structure of a classification with binary predicate $s \models t$ where a is of sort $\mathtt{tok}$ and t is of sort $\mathtt{typ}$. Now say that formula $\varphi(a, t)$ implies formula $\psi(a, t)$ *along an infomorphism* f from A to B if for all $a \in \mathtt{tok}(A)$ and $t \in \mathtt{typ}(B)$,

$$\varphi(a, f^\wedge t) \text{ is true in } A \text{ iff } \psi(f^\vee a, t) \text{ is true in } B$$

Then say that φ *infomorphically entails* ψ, if φ implies ψ along any infomorphism. van Benthem (2000) shows that informorphic entailment is characterised by the existence of an interpolant of a special kind. A *flow formula* is one that has only atomic negations, existential quantification over tokens and universal quantification over types, i.e. it is constructed from $a \models t$, $\sim a \models t$, &, $\vee$, $\exists a$ and $\forall t$. Flow formulas are all preserved over infomorphisms and morever the following are equivalent:

1. φ infomorphically entails ψ
2. there is a flow formula α such that $\varphi \vdash \alpha \vdash \psi$[14]

This result can be extended to special classes of classifications and restrictions on infomorphisms, and even to infinitary languages. We refer the reader to van Benthem (2000) for further details.

Finally, we note that channels can be used to model some of the operations of information flow in modal logic. A model for a modal language can be regarded as a classification M of points by formulas, and the accessibility relation r_a for each modal operator $[a]$ determines a local logic L_a on $M+M$, such that $\langle u, v\rangle$ is normal iff $r_a(u, v)$ and $\langle 1, [a]\varphi\rangle \vdash_a \langle 2, \varphi\rangle$. Given two such models M and N and their corresponding local logics L_a (on $M + M$) and P_a (on $N + N$), a channel B between M and N is a *bisimulation channel* if B; L_a and P_a; B are logical refinements of each other. The relation that B determines between tokens of M and N, namely $\langle \pi_1 b, \pi_2 b\rangle$ for each token b of B, is a bisimulation iff B is a bisimulation channel. The usual definition of bisimulation can thus be seen as a consequence of a kind of equivalence between the two models when representation as channels.

It is sometimes complained that the concept of information flow modelled by channels is too static, failing to account for changes in information that result from

[14] In van Benthem (2000), these results are stated in terms of Chu-transformations, but the difference is only one of terminology.

certain events, such as public announcement, which are captured very neatly in other frameworks, such as that of dynamic epistemic logic. But the operations on models that are characteristic of dynamics can also be modelled as channels. For example, the result of public announcement of φ is the restriction $M|_\varphi$ of the model M to the extension of φ. This determines a local logic on $M + M|_\varphi$ in which $\langle u, v\rangle$ is normal iff $u = v$, $\langle i, p\rangle \vdash \langle j, p\rangle$ for $i = 1, 2$, and $\langle 1, [!\varphi]\psi\rangle \vdash \langle 2, \psi\rangle$. In this case, and in the basic case of the representation of modal logics as channels, it would be nice to have results that characterise the constructions in purely categorical terms but that work has not yet been done.

References

Barwise, J. 1993. Constraints, channels and the flow of information, in: Aczel, P., Israel, D., Katagiri, Y., Peters, S. (Eds.), Situation Theory and Its Applications, 3, number 37 in CSLI Lecture Notes. CSLI Publications, Stanford, CA, pp. 3–27.

Barwise, J., Etchemendy, J., 1987. The Liar: An Essay on Truth and Circularity. Oxford University Press, New York.

Barwise, J., Gabbay, D., Hartonas, C., 1996. Information flow and the Lambek calculus, in: Seligman, J., Westerståhl, D. (Eds.), Logic, Language and Computation, Proceedings Information-Oriented Approaches to Logic, Language and Computation, vol. 58. CSLI Lecture Notes, CSLI Press, Stanford, CA, pp. 47–62.

Barwise, J., Perry, J., 1983. Situations and Attitudes. MIT Press, Cambridge, MA.

Barwise, J., Etchemendy, J., 1987. The Liar: An Essay on Truth and Circularity. Oxford University Press, New York.

Barwise, J., Seligman, J., 1994. The rights and wrongs of natural regularity. Philos. Perspect. 8, 331–364.

Barwise, J., Seligman, J., 1997. Information Flow: The Logic of Distributed Systems. Cambridge University Press, Cambridge, UK.

Goguen, J., 2006. Information Integration in Institutions, in: Moss, L. (Ed.), to appear in the Jon Barwise Memorial Volume, Indiana University Press, Bloomington, IN. cseweb.ucsd.edu/~goguen/pps/ifi04.pdf.

Israel, D., Perry, J., 1990. What is information?, in: Hanson, P. (Ed.), Information, Language, and Cognition. University of British Columbia Press, Vancouver, pp. 1–19.

Mares, E.D., 2004. Relevant Logic: A Philosophical Interpretation. Cambridge University Press, Cambridge, MA.

Restall, G., 1995. Information flow and relevant logics, in: Seligman, J., Westerståhl, D. (Eds.), Logic, Language and Computation: The 1994 Moraga Proceedings. CSLI Publications, Stanford, CA, pp. 463–477.

Restall, G., 2005. Logics, situations and channels. J. Cogn. Sci. 6, 125–150. Available from: http://consequently.org/writing/channels/.

Seligman, J., 1990. Perspectives in situation theory, in: Mukai, K., Cooper, R., Perry, J. (Eds.), Situation Theory and its Applications, vol. 1. CSLI Lecture Notes, CSLI Press, Stanford, CA, pp. 147–191.

Seligman, J., 2009. Channels: from logic to probability, in: Sommaruga, G. (Ed.), Formal Theories of Information: From Shannon to Semantic Information Theory. Springer-Verlag, Berlin and Heidelberg, LNCS 5363, pp. 193–233.

Seligman, J., Moss, L., 2010. Situation theory, in: van Benthem, J., ter Meulen, A. (Eds.), Handbook of Logic and Language. Elsevier BV, pp. 171–244.

van Benthem, J., 2006. Information as correlation vs. information as range, in: Moss, L. (Ed.), to appear in the Jon Barwise Memorial Volume, Indiana University Press, Bloomington, MA. www.illc.uva.nl/Publications/Research Reports/pp. 2006-07.text.pdf

van Benthem, J., 2000. Information transfer across chu spaces. Logic J. IGPL 8 (6), 719–731.