General Information in Relevant Logic

Edwin D. Mares
Victoria University of Wellington

October 24, 2008

Abstract

This paper sets out a philosophical interpretation of the model theory in Mares and Goldblatt "An Alternative Semantics for Quantified Relevant Logic" (*The Journal of Symbolic Logic* 71 (2006)). This interpretation distinguishes between truth conditions and information conditions. Whereas the usual Tarskian truth condition holds for universally quantified statements, their information condition is quite different. The information condition utilizes *general propositions*. The present paper gives a philosophical explanation of general propositions and argues that these are needed to give an adequate theory of general information.

1 Introduction

This paper is a sequel to (19). In that article, we set out a formal semantics for quantified relevant logic.[1] Here I provide a philosophical interpretation for that semantics. This interpretation is undertaken in the framework of an informational semantics. An informational semantics, on my view, does not set out truth conditions for the connectives. Rather, it gives *information conditions* for the various connectives. For most of the standard connectives the usual classical truth conditions hold. An information condition is a condition that holds in states of partial information – situations – that indicates that the truth condition for the statement in question obtains.

My particular purpose is to set out an information condition for the universal quantifier. The nature of general information is a puzzle that has been familiar to philosophers at least since Russell discussed it in (21). We can know of each mammal in my lounge, for example, that it is a dog, without thereby having the information that all the mammals in my lounge are dogs. Some extra information is needed. The problem, then, is to formulate what this extra

[1] The formal problem that was solved by our semantics was that the standard quantified relevant logics (which we call 'QR' and 'RQ') were incomplete over a standard Tarskian semantics (as Kit Fine showed (12)). Fine also constructed a semantics over which RQ is complete, but it is very complicated and not very intuitive (although extremely ingenious) ((11)).

information is and to show how our formal semantics adequately represents our informal understanding.

I begin the paper by presenting my version of informational semantics. This is the theory of situated inference developed in (17). I then make some modifications and additions to the theory of situated inference both because they are philosophically warranted and they are required by my theory of general information. One particular addition is the use of a class of propositions. I argue that the use of propositions in the semantics is philosophically innocuous. Once the framework is set down, I proceed to develop a theory of general information. I give a brief and relatively informal presentation of the formal semantics, motivating it by appeal to the natural deduction system for the relevant logic R. I then examine three examples of general information: an ordinary case of visual information; a case of legal information; and a case in which there is mathematical information. I then make the theory more realistic by adding an existence predicate to the logic. This allows us to vary the domains from situation to situation. The paper ends with an appendix that contains the model theory.

2 The Need for Situations

I use a theory of information as a semantics for a given logic. This is a very different project to that of constructing a logic of information flow.[2] The language of a logic of information flow needs to represent the nodes and relations involved in information flow, such as agents, changes of states of agents, informational channels, and so on. To give a semantics for a logic, in contrast, is to give a theory of what the various particles of that language represent. On my view, a statement represents both its truth conditions and its information conditions. This bifurcated semantics will be presented later. First we need to present a framework for this semantics. This framework is the theory of situations.

I take the basis for the theory of situations from Barwise and Perry (8). A situation is an abstract structure. It is constituted in part by a set of fact like entities called “states of affairs”. There are various ontologies of states of affairs in the literature (see, e.g., (4), (7)). I do not assume any particular such theory; any will do for my purposes. Situations carry information. How they do so will be dealt with in the next section below. A situation may accurately describe a possible world or it may fail to do so. In accurately describing a world, a situation need not describe everything that is true of the world, but rather all the information carried by the situation must be true of that world. A situation which accurately describes a possible world is said to be a possible situation and a situation which does not accurately represent any possible world is said to be impossible.

Why use situations? The reason comes from the proof theory of relevant logic. Relevant logic was created to avoid the so-called paradoxes of material

[2] Most of the other contributions to this special issue are attempts to do the latter. I think that the two projects are compatible. Exactly what they have to do with one another, though, is far from clear. It would be an excellent project to look at the relationship between them.

and strict implication. In terms of the natural deduction system for relevant logic[3] the paradoxes are avoided by imposing a condition that all premises or hypotheses of a proof really be used in the proof. To understand how this works and what it has to do with the use of situations, let's consider an example.

The paradox that we will look at is: $A \to (B \vee \neg B)$. It is a paradox because A appears to have nothing to do with $B \vee \neg B$. Standard classical proofs for this have the following form:

$$\begin{array}{ll} \quad A & hyp \\ \quad \text{some proof of} & \\ \quad B \vee \neg B & \\ A \to (B \vee \neg B) & \to I \end{array}$$

Here the hypothesis that A is merely tacked on the start and then discharged in the last step, even though it may have had nothing to do with the proof of $B \vee \neg B$.

In order to reject such proofs, Anderson and Belnap add subscripts to each line of a proof to indicate which hypotheses or premises were used in the derivation of that line. For example, here is a relevant proof of $A \to ((A \to B) \to B)$:

$$\begin{array}{lll} 1. & \quad A_{\{1\}} & hyp \\ 2. & \quad\quad A \to B_{\{2\}} & hyp \\ 3. & \quad\quad A_{\{1\}} & 1,\ reit \\ 4. & \quad\quad B_{\{1,2\}} & 2,3,\ \to E \\ 5. & \quad (A \to B) \to B_{\{1\}} & 2-4,\ \to I \\ 6. & A \to ((A \to B) \to B)_{\emptyset} & 1-5,\ \to I \end{array}$$

When a hypothesis is introduced, it is given a number. When a hypothesis is used in rules such as $\to E$, its number is added to the subscripted set of the conclusion of that rule. When a hypothesis is discharged as it is in an instance of $\to I$, its number is removed from the subscripted set of the conclusion of the rule. An empty set subscripted to a line indicates that the formula on that line is a theorem of the logic.

We won't go into the rules of proof for all the connectives here. There are plenty of sources in which one can find them (e.g., (2), (19)). What is of interest to us here is the interpretation of the subscripts. I interpret a hypothesis or premiss in a proof, $A_{\{j\}}$, as meaning that the information that A is carried by a situation s_j. When one makes this hypothesis in a proof, what she is saying is "assume that there is a world in which a situation j obtains and j carries the information that A". When further hypotheses are made, say, $B_{\{k\}}$, the person doing the proof is also assuming that there is a situation s_k in the same world which carries the information that B. When we have multiple numbers in the subscript, as in $C_{\{1,...,n\}}$, in the same proof, what is shown is that there is a situation in the same world as $s_1, ..., s_n$ which carries the information that C

[3] Here and throughout the paper, by "relevant logic" I mean the logic R of relevant implication. This is somewhat unfair. There are many other (infinitely other) relevant logics and what I say cannot be applied to all or even most of them.

and the information in all of $s_1, ..., s_n$ is really used to derive that this is so. This is the core of the theory of *situated inference* developed in (17).

Note that this semantical analysis appeals to the notion of real use; it is not defined in the semantical theory but rather it is assumed by the semantical theory (see (17)).

The fact that situations provide partial representations of worlds is crucial to their use in the semantics of relevant logic. If all situations were complete, in the sense that they all carried the information that every instance of the law of excluded middle held, then on this semantical analysis it would be very difficult to invalidate the formula $A \rightarrow (B \vee \neg B)$. In fact, for each formula of our language, we need at least one situation in some model at which it fails.[4]

3 Situations and Worlds

In (17), I held that a world w made true a statement A if and only if there was a situation s in w such that $s \models A$. I think that this biconditional still holds at least of the actual world, but I no longer think that a statement is true at a world *because* the information that it is true is carried at some situation in that world. It seems to me that this inverts the real direction of explanation.

I suggest that the classical truth conditions for conjunction, disjunction, negation, and the quantifiers are correct. These conditions, relativized to possible worlds, of course are the following:

- $A \wedge B$ is true at w if and only if A is true at w and B is true at w
- $A \vee B$ is true at w if and only if A is true at w or B is true at w
- $\neg A$ is true at w if and only if A is not true at w
- $\exists xA$ is true at w if and only if A is true at w on some interpretation of the variable x
- $\forall xA$ is true at w if and only if A is true at w for all interpretations of the variable x

The standard truth conditions for the connectives have the nice property that they are *homomorphic*, that is, they are interpreted very directly in terms of the cognate connectives of the metalanguage. So, according to this theory, ‘and’ means ‘and’ (in the metalanguage), ‘or’ means ‘or’, and so on. Thus, it would seem to be a virtue of a semantic theory if it could retain the classical treatment of these connectives.

But there is a problem. As I said in the introductory section above, the semantics that I use for relevant logic is informational. That is, the logical

[4] At least this is true for standard versions of propostional and first order relevant logics. If we add the so-called Church constants T and F, we make T hold at every situation and F fail at every situation. We then we also add all instances of the schemes theorems $A \rightarrow T$ and $F \rightarrow A$. Also, if we have propositional quantifiers, we make $\exists pp$ hold at every situation and $\forall pp$ fail at every situation.

properties of a statement are to be understood primarily in terms of the information that is conveyed by it. Information and truth are closely linked, but they are not identical. While I am writing this sentence I am travelling in a train carriage. I have certain information available in my surroundings. I can see that the person sitting across from me is reading a magazine in French, I can hear that others are eating, talking too loudly on cellphones, and so on. Other statements are true of this carriage – such as those that stipulate where it was made, how old it is, and so on – but the truth of these is not among the information available to my senses. Not all information need be experientially available to one, but all of it is *situated*. In a given world, situations carry some information and others carry different information. Truth is not like this. In a given possible world, a proposition either true or false.

Thus I distinguish between truth conditions and *information conditions*. The information condition for a connective generalizes the conditions under which we have information of a given type available in actual contexts. The most obvious condition is for conjunction, which merely repeats the truth condition, viz, $s \models A \wedge B$ if and only if $s \models A$ and $s \models B$. (Here '$s \models A$' means 's carries the information that A'.) In section 4 below, we will look briefly at information conditions for negation and implication and after that we will examine the information condition for the universal quantifier. We will set aside the information conditions for disjunction and the existential quantifier, since they are rather difficult and deserve to be discussed in a paper dedicated just to them.

As a simplifying assumption, we treat all possible worlds and all situations as having the same domain of individuals, which we call I. Clearly, in our final view, we want to have situations have different domains of individuals. By and large, situations are supposed to carry only partial information about worlds. Situations should not all, therefore, tell us exactly which individuals exist. In section 11 below, I remedy this by showing how to alter the semantics to treat variable domains.

4 Information Conditions

In this semantical theory, each connective has associated with it an information conditions. The information conditions generalize the conditions under which information of the salient sorts is available to us in concrete contexts. Before we turn to the information conditions for the universal quantifier, let's look briefly at two other connectives – negation and implication – so that we can see two rather different ways of determining information conditions.

First, we will consider negation. A situation, as we have said, may be partial. But it does represent worlds as making true particular statements or propositions. For example, as I write this sentence, it is sunny in Wellington. Thus, the situation that captures the information to which I have visual access represents Wellington as being not cloudy. It does so by containing information that is incompatible with its being cloudy (i.e. the information that it is sunny). We generalize this sort of example by placing a binary relation (the incompatibility relation) on situations and then formulating a general information condition using it. The information that $\neg A$ is carried by a situation s if and only if s is incompatible with any situation which carries the information that A.[5]

My information condition for implication is motivated by the implication properties of that connective. According to the theory of situational inference, the primary sort of inference that we make with implicational information is a form of modus ponens. Suppose that we have the information in a situation s that $A \rightarrow B$. This tells us that we may draw certain conclusions from certain sorts of hypotheses. In particular, let us hypothesize that there is a situation t in the same world as s such that t carries the information that A. Our implicational information allows us to infer that there is also a situation u in the same world as s and t such that u carries the information that B.

The information condition for implication, then, is just this: $s \models A \rightarrow B$ if and only if, for any situation t that carries the information that A, if t is hypothesized to obtain in the same world as s, then (really using the information in both s and t) we can derive that there is also situation in that world which carries the information that B.

There are a few points that need to be made about this information condition. First, the notion of the real use of hypotheses appears in the information condition for the connective (as well as in the description of the natural deduction system). One might have expected that the semantics for the natural deduction system gave an explanation of real use, instead of utilizing that very same notion. Instead, the notion of real use has an intuitive meaning (that I think is clear) and a more technical contextual definition in terms of the natural deduction system and in the semantic theory. The completeness theorem, in my opinion, shows that these two definitions succeed in defining what is in effect the same notion.

Second, it might seem disappointing to some[6] that the inferential properties are used to motivate a semantical analysis rather than having the semantics justify the inferential properties of the connective. I think that this is inevitable with an informational semantics. The way in which information is made available to us must closely reflect the way in which we obtain and manipulate it.

Third, one might wonder under what conditions there is sufficient information available to make inferences of this sort. Information that allows us to make situated inferences is of a sort that (17) calls “informational links”. An informational link is a perfectly reliable connection, such as a law of nature or a convention. Thus, for example, if we have the information that it is a law that every physical object attracts every other physical object, then given a situation (real or hypothetical) which carries the information that i and j are physical, we can infer that there is also a situation in the same world in which i attracts j.

[5] This analysis of negation in relevant logic by mean of a compatibility relation is due to Dunn (10).

[6] Johann Van Benthem expressed this very disappointment. I am indebted to him for making me face up to this issue.

Not only are informational links responsible for situated inference, but we are also allowed to manipulate hypotheses by logical means. For example, we can rearrange hypotheses in whatever order we wish. Suppose that $s \models A \rightarrow (B \rightarrow C)$. Let us hypothesize that there is a situation in the same world as s which carries the information that A and a situation in that same world which carries the information that B. Then we can infer that there is a situation that carries the information that C. There is nothing in the order of these hypotheses that seems to affect the inference. That is, we can make the hypothesis first that there is a situation that carries the information that B then the hypothesis that there is one that carries the information that A. We would still be entitled to infer that there is a situation that carries the information that C. Thus, by the information condition for implication, it would seem that $s \models B \rightarrow (A \rightarrow C)$. This shows that allowing the rearrangement of hypotheses makes valid certain inferences.

Note that the concept of an informational link that I utilize here is very different from the one in channel theory, which is also used as a basis for a semantics for relevant logic (Barwise (1993), Restall (1996)). A channel is an entity that connects pairs of situations. It can be considered to be a situation as well. Consider an example from Restall (1996). Consider the statement: for any number n, if n is even so is $n+2$. Restall represents this as saying that from any situation which carries the information that a particular number n is even, there is also a situation that carries the information that $n+2$ is even. And connecting these two situations is another situation that carries the rules of arithmetic that yield a proof that $n+2$ is even given that n is even (ibid. 471f). The philosophical difference between my semantics and the Barwise-Restall semantics is that they consider links to be situations (or situation like entities), whereas for me they are information that is contained in situations. The formal difference is that channel theory is most naturally used as a semantics for weaker relevant logics that, in particular, do not contain this version of modus ponens: $((A \rightarrow B) \wedge B) \rightarrow B$. For in order to make valid that thesis, all models would have to be constrained so that every channel is a channel from itself to itself, which is not an easy postulate to motivate.

5 Propositions

Before we can go on to examine the information condition for universally quantified statements, we will need to understand another key notion in our semantical theory. This notion is that of a proposition. A proposition here is to be taken in its usual sense of what is believed, denied, desired, wondered about, and so on. That is, a proposition is a possible object of propositional attitudes. Our propositions are unstructured, like UCLA propositions, but are taken to be sets of situations rather than sets of worlds. I don't deny that there may be structured propositions in addition to unstructured ones, but as the bearers of logical relations (especially entailments), I think that unstructured propositions are a necessary part of semantics. The key relation that we have in mind is entailment. On the unstructured proposition view, entailment is just the subset relation, that is, a proposition π entails a proposition σ if and only if $\pi \subseteq \sigma$.

The Routley-Meyer semantics for relevant logic – which is the formal basis for my philosophical semantics below) – places a partial order, $\sqsubseteq$, on situations.[7] On my reading, taken from Barwise and Perry's understanding of a very similar relation, $\sqsubseteq$ is a relation of *informational containment*. This means that for situations s and s', $s \sqsubseteq s'$ if and only if all the information carried by s is also carried by s'. Thus, it would seem that since they are to be taken to represent information, propositions must be closed upwards under $\sqsubseteq$. In more mathematical terminology, propositions are "upsets". We will use this fact about propositions again in section 10.1 below.

Not just any upset of situations should count as a proposition in our semantics. Our semantics, as we have said, is informational; in particular it is supposed to capture the way in which *we* understand the world. Thus, the limits of what counts as a situation depends on the discriminatory capacities of human beings. A proposition is a set of situations between all of which people could find important similarities. Arbitrary sets of situations may not satisfy this requirement. It falls out of our view of propositions that every statement expresses a proposition. That is, for any formula A, the set of situations that carry the information that A is a proposition.

Note that the entailment relation and the informational containment relation are not identical (they are, in a certain sense, dual to one another). Propositions – which are sets of situations – entail one another. The containment relation relates situations. If $a \sqsubseteq b$, then the total information in b entails all the information in a. In that sense they are dual to one another.

6 General Information

The universal quantifier is like implication. It is easiest to begin with its inferential properties, since its information condition is not immediately obvious (if it were, I wouldn't need to write this paper). To make matters easier, let us suppose that we have a model in which every individual has a name. We will see a more rigorous and formal treatment in the appendix below. Clearly, (in this model) the central inferential purpose of the universal quantifier is that it allows one to infer from $\forall xA$ to $A[c/x]$ for all names c. To licence this inference it must be that if s carries the information that $\forall xA$, then it also carries the information that $A[c/x]$ for every c.

In order to explain the information condition for the universal quantifier, we now bring in propositions. As we have said, every statement expresses a proposition. In particular, the statements $\forall xA$ and $A[c/x]$ express propositions – the sets of situations $|\forall xF(x)|$ and $|F(c)|$ respectively. Putting this together with what we said in the previous paragraph, it must be the case that $|\forall xA|$ is

[7] A formal presentation of the Routley-Meyer semantics and our extension of it to treat first order relevant logic is given in an appendix to the present paper.

a subset of $|A[c/x]|$, that is, that $|\forall x A|$ entails $|A[c/x]|$. Generalizing, we obtain

> (MC) If s carries the information that $\forall x A$
>
> then s is in some proposition that entails $|A[c/x]|$ for every name c.

I call this the "minimal constraint" on general information. It is minimal because it is derived directly from the central inferential role of the universal quantifier and the definitions of 'proposition' and 'entailment'.

On my view, information conditions are biconditionals. So we should ask: what needs to obtain other than the consequent of MC in order for s to carry the information that $\forall x A$? My answer is 'nothing'. My reason for this comes from looking at the introduction rule for the universal quantifier. The standard introduction rule (with the addition of subscripts) is:

$$\begin{array}{l|l} y & \vdots \\ & A(y)_\alpha \\ \hline \forall x A_\alpha & \qquad \forall I \end{array}$$

In words: from a subproof of $A(y)_\alpha$, where y does not occur free in any proof on which this subproof depends, we can infer that $\forall x A(x)_\alpha$. If we accept this introduction rule, then we can derive that the consequent of MC is also sufficient for $s \models \forall x A$.

Suppose that the consequent of MC obtains, that is, we have a situation s which is in a proposition π that entails $|A[c/x]|$ for every name c. Let us add a new propositional constant p to our language and set $|p| = \pi$. It is legitimate to claim that, at least in principle, we can add a sentence that expresses π, since our notion of a proposition is that of a possible content of thought. Thus, $|p|$ entails $|A[c/x]|$ for every c. We can express this entailment fact in the syntax of the natural deduction system by the formula $p \rightarrow A[c/x]_\emptyset$. If we have a formula with the empty set as a subscript, we can add it at any step in a proof (see (2)).

Now we can construct the following proof:

1.	$p_{\{s\}}$		*Premise*
2.	x	$p \rightarrow A_\emptyset$	*assumption*
3.		$p_{\{s\}}$	1, reit
4.		$A_{\{s\}}$	$2, 3, \rightarrow E$
5.	$\forall x A_{\{s\}}$		$2-4, \forall I$

It seems, then, that the consequent of MC is also a sufficient condition for s to carry the information that $\forall x F(x)$. So, we have the following biconditional:

s carries the information that $\forall x A$

if and only if

s is in a proposition that entails $|A[c/x]|$ for every i.

I suggest that this biconditional be taken to be the information condition for the universal quantifier. Generalizing this condition, we get:

$$s \models \forall x A \text{ iff there is some proposition } \pi \text{ such that}$$
$$\text{(i) } \pi \subseteq |A[c/x]| \text{ for all names } c \text{ and (ii) } s \in \pi.$$

This condition has the virtue also of making valid the elimination rule for the universal quantifier. For suppose that we have a situation s such that $s \models \forall x A$. Our information condition tells us that s is in a proposition π such that $\pi \subseteq |A[c/x]|$ for all names c. Now consider a term τ. We take a name c that has the same referent as τ and (after proving a few lemmas (see (19))), we know that $|A[c/x]| = |A[\tau/x]|$. Thus, $s \in \pi \subseteq |A[\tau/x]|$, so $s \in |A[\tau/x]|$, i.e. $s \models A[\tau/x]$.

Our information condition for $\forall$ is minimal in a philosophical sense. Any theory that holds that propositions are sets of situations and that accepts the above introduction and elimination rules for the universal quantifier must also accept the above biconditional. Although our information condition is minimal in a philosophical sense, it is not minimal in terms of the formal logic. In order to formulize this condition, we have added to the basic model theory of situations and relations on situations a set of propositions (i.e. a designated set of sets of situations). Taking care of this set of propositions can be formally cumbersome, but this model theory has been shown has various formal virtues – it has been used to characterize a wide range of logical systems (modal and relevant) and used it to prove some independence results (see (13), (14), and (19)). Thus, the condition can be motivated on a formal as well as philosophical grounds.

7 Russell's Argument

What we have said so far about the informational nature of 'all' mirrors to a great extent an argument due to Bertrand Russell (21). Russell examines a case of "exhaustive induction" – a case in which one comes to a general conclusion by enumerating all the salient particular facts.

For example, Zermela is in my lounge, she is a mammal and she is a dog. Lola is in my lounge and she is a mammal and a dog. Thus, we conclude, every mammal in my lounge is a dog. But this is not a valid inference. We need another piece of information: it would be valid if we knew, for example, that the only mammals in my lounge are Zermela and Lola. But this extra piece of information is general – it tells us about *all* the mammals in my lounge (that they are identical either to Lola or Zermela). Russell thinks this point is universal:

> You can never arrive at a general proposition by inference from particular propositions alone. You will always have to have at least one general proposition in your premisses. ((21) 206)

Thus, there is a sort of circularity about inferences of this sort to generalities. They presuppose that other general propositions hold.

Now Russell, and following him, Armstrong ((3) and (4)) hold that this argument shows that there must be general facts of some sort. They think that for every true statement (or perhaps proposition) there must be a thing that makes it true (its "truthmaker"). Moreover, Armstrong explicitly and perhaps Russell implicitly believe that a truthmaker makes true a proposition by the truthmaker's entailing that proposition. Thus, the failure of an entailment from the particular facts about the mammals in my lounge to the proposition that every mammal in my lounge is a dog shows that there must be an extra fact of some sort.

The thesis that truthmakers entail the things that they make true has been criticized in the literature (e.g. (15)). I do not want to enter into this debate here. We are interested not in the nature of facts but of information. I do think that Russell's argument shows this: in addition to the particular facts about the denizens of my lounge there must be some sort of general information if we are to validly infer that every mammal in my lounge is a dog. On my view, entailment – like other logical relations – has to do primarily with information. The failure of entailment shows that we need more information. This extra information, in my view, is a general proposition of some sort.

Thus, it seems that from the proof theory and Russell's argument that we have good motivation for the inclusion of general propositions in our semantics. In what follows we look at the nature of general propositions.

8 A Legal Example

Here we will look at three sorts of cases in which we have general information. I do not pretend that these are exhaustive. I wish I had a complete taxonomy of types of general information, but I don't. For the present, I give these three rather different sorts of cases to indicate how we come to have general information and how this process is captured by our semantics.

I start with a case that fits in straightforwardly what we have said so far about the nature of situations and information.

The seventeenth amendment to the Constitution of the United States says "the Senate of the United States shall be composed of two senators from each state". Thus, it is a law in America that each state can have only two senators. I don't think that there is currently any judgment on what would happen if a state tried to appoint more than two senators, but for the sake of our example let us assume that the appointments would not be legal and no more than two of the appointments (if any) would be considered actual appointments. Allowing this supposition, we can say that it is a legal fact that there can (in an alethic sense) be at most two senators from any state. This fact, and the proposition that represents it, constitutes an informational link in the sense of section 4 above.

Here are some facts about American law and the senators from the state of Wisconsin:

1. It is a fact that Russell Feingold is a senator from Wisconsin.

2. It is a fact that Feingold is a Democrat.
3. It is a fact that Herb Kohl is a senator from Wisconsin.
4. It is a fact that Kohl is a Democrat.
5. It is a fact that Russell Feingold and Herb Kohl are numerically distinct.
6. It is a law that there can be at most two senators from any state.
7. The state of Wisconsin has not broken any law to do with the appointment of Senators.

These facts together with what we have supposed about American law entail that every senator from Wisconsin is a Democrat. As we have said, the legal proposition that there can be at most two senators from any state is an informational link. When conjoined with the other facts listed here entails that all senators from Wisconsin are Democrats. Facts 1, 3, 5 and 6 entail that for any individual, if it is a senator from Wisconsin, it is identical either to Feingold or Kohl. This is a general proposition. Together with the facts that Kohl and Feingold are both Democrats this general proposition entails that every Senator from Wisconsin is a Democrat.

9 A Mathematical Example

Now we turn from legal to mathematical information. In proving a universally quantified statement in mathematics, if we are performing a direct proof (as opposed to a reductio), we prove that a statement with a free variable holds, that is, we show that a statement holds of any arbitrarily chosen object of the appropriate kind. We then infer that the statement holds of all objects from that domain.

Let's consider a familiar example. In a standard proof that the set of prime numbers is infinite, we show that any prime number we can construct a larger prime number. Here is a sketch of the proof. Suppose that n is an arbitrary prime number. Then we take all of the prime numbers less than or equal to n, multiply them together and add 1. Then, since this new number m is at least one greater than n, it is strictly greater than n. Moreover, we can show that m is prime (since its only factors are itself and 1). Thus, for *every* prime number, there is at least one prime number strictly greater than it. Therefore, there are infinitely many prime numbers.

The point that is of interest to us is that we can infer the universally quantified statement 'for every prime number there is at least one prime number strictly greater than it' from the proof that for an arbitrarily chosen prime n that there is a prime number strictly greater than n. This of course is just an informal use of the rule of $\forall I$ that we give in section 6 above. We are not, however, concerned here with the exact form of the proof, but with the fact that the existence of the proof is the general information that allows the inference of the universally quantified statement.

Two questions arise here, one metaphysical and the other logical. The metaphysical question is easy to brush aside in the current context. The question is what we mean by a proof here – are we committed to Platonism about proofs? Although my view is compatible with Platonism, it is compatible with almost every philosophy of mathematics. Any reasonable philosophy of maths will either treat the expression 'there is a proof' literally or adopts some paraphrase that makes at least some statements of the form 'there is a proof of p' true. Any such paraphrase should be acceptable in the current context. That is, it should determine a general proposition of the salient sort.[8],

The logical question is rather more difficult. This is the question of what sort of proofs count as general information of the requisite sort. Only one answer is possible in this context: relevantly valid proofs. This raises a further problem. A lot of proofs that mathematicians actually give are not (as stated) relevantly valid. For example, they often use the rule of disjunctive syllogism. Luckily, we can recast many of these proofs as relevantly valid proofs. For example, suppose that we are considering a proof accepted among mathematicians of the form of a disjunctive syllogism,

$$\frac{\begin{array}{c} p \vee q \\ \neg p \end{array}}{\therefore q}.$$

Disjunctive syllogism is not a relevantly valid form of inference. But it is commonly used by mathematicians and almost everyone else.

In order to deal with this problem, in (17) I suggested that inferences of this form can be thought of as enthymemes. The fully displayed inference is of the form

$$\frac{\begin{array}{c} ((p \vee q) \wedge \neg p) \Rightarrow q \\ (p \vee q) \wedge \neg p \end{array}}{\therefore q}.$$

The double arrow in the major premise is a relevant indicative conditional (described in chapter 7 of (17)). Now we have a valid inference – it is just a case of modus ponens. Moreover, I argue in (17) that if the premises of the original inference are true then the fully displayed inference is sound. In this way we can recapture a good number of standard mathematical proofs. Whether it is sufficient to explain all of what we think of as good mathematical practice is not clear, but not a topic for the present paper.

Before we leave the topic of mathematical information, I want to point out that I am not claiming that all general mathematical information comes from proofs. We do not prove all mathematical generalities. Many widely accepted axioms (such as Peano's axioms for arithmetic) are generalities. Thus, there

[8] I am grateful to Stephen Read for this view. Read, at one time at least, thought that the truthmakers for necessities were proofs. So he (at least at that time) went much further than my claim that the certain situation carries certain mathematical information because in those situations there exist proofs.

must be some other sort of mathematical information. What that is, however, I will leave to those working in the epistemology of mathematics.[9]

10 A Visual Example

My last example is about common empirical experience.

Suppose that we see in front of us a glass bowl with three peaches in it. Since the bowl is glass, if we can picks the bowl up and look it from all sides (or even remove and replace the peaches) we can see that *all the pieces of fruit in the bowl are peaches.* It doesn't seem much time for us to realize that in cases like this we can make universal generalizations about these peaches and certain of their properties and that we can justify these generalizations using exhaustive induction. Let's call this sort of case "locally Tarskian" because it seems that we are given the entire domain of pieces of fruit in the bowl and, for certain formulas A, that the general information that $\forall xA$ is available in this situation if and only if A holds of each of the peaches we can see.

Russell's argument that we examined in section 7 above, however, shows that from a logical point of view, there is something interesting about locally Tarskian circumstances. We need not only information about the individual pieces of fruit in order to make generalizations about them, but that these individual pieces of fruit are the only ones in the bowl. We need now to discuss the nature of this further (general) information.

10.1 The Closed World Assumption

Our starting point for our discussion of locally Tarskian information is the idea of *closed worlds* that I borrow from default logic and artificial intelligence. The term "closed world assumption", I think, was introduced into the literature by Ray Reiter. As it is used in artificial intelligence programming, the closed world assumption is the assumption that a database or program has available to it all the salient facts. For example, suppose that an insurance database lists a number of valuable items belonging to a particular policy holder. If the program searching the database employs the closed world assumption, it will say if asked that the policy holder does not own any other valuables. Thus, the program assumes that the facts in the database, for the purposes of making inferences, constitute a "closed world".

The technical machinery of the relevant semantics gives us a choice in interpreting the closed world assumption in our framework. In order to understand

[9] An alternative view of mathematical generality can be found in the classic, but nowadays often overlooked paper Alice Ambrose (1). Ambrose says that a matematical sentence of the form 'All fs are g' is true because of "an empirical truth about language" (p 117). On her view, all necessary truths are true because of the grammar of our language in the Wittgensteinian sense of being appropriately connected by the rules that govern the use of the relevant predicates in the language. If this view is right, I can fit it easily into my framework. Conventional rules of this sort create informational links and can be the bases of general propositions.

the distinction between the two alternative interpretations we need to understand the distinction between asserting a negative proposition and denying a positive proposition. In the context of logics with non-classical negations like relevant logic,[10] we must distinguish between the assertion of $\neg A$ and the denial of A. A proposition A is *accurately denied* in a situation s if and only if A fails to obtain in s.[11] The information that the negation $\neg A$ obtains in s if and only there is some feature of s that is incompatible with the truth of A. The difference here can be understood in terms of the partial order on situations (see section 5 above). If the denial of A is accurate in s, then it is still possible that there are $\sqsubseteq$-extensions of s in which the information that A obtains. Negations, on the other hand, are persistent. If $\neg A$ obtains in s, then it obtains in all $\sqsubseteq$-extensions of s. We will call these two interpretations of the closed world assumption, the *denial interpretation* and the *negation interpretation*.

The denial interpretation is the one that is commonly used in artificial intelligence, in particular in connection with the interpretation of negation in Prolog. Here is a quotation from Ivan Bratko's textbook on Prolog:

> According to this assumption *the world is closed* in the sense that everything that exists is in the program or can be derived from the program. Accordingly then, if something is not in the program (or cannot be derived from it) then it is not true and consequently its negation is true. This deserves special care because we do not normally assume that 'the world is closed' .... ((9) 135)

Bratko suggests that $\neg A$ according to Prolog really means 'there is not enough information in the program to prove that A' (ibid. 134-5). This form of negation does not match the meaning of 'not' in English or the meaning of any other natural language negation.

With regard to negation and, more importantly for our present purposes, the universal quantifier, the denial interpretation is not what we need. As we have seen in section 5 above, propositions are persistent; thus, every general proposition must also be persistent. This eliminates the denial interpretation. So, it seems we must accept the negation interpretation of the closed world assumption for our present purposes.[12] Thus, in the cases above, we assume the proposition that there are no more relevant facts concerning the particular subject matter.

[10] The case of intuitionist logic is interesting. Intuitionists deny the law of excluded middle, but they cannot assert the negation of any particular example of excluded middle, since $\neg(A \vee \neg A)$ is a contradiction in intuitionistic logic. David McCarty has convinced me that this and other similar cases can be understood by using intuitionist logic in an intuitionist metalanguage. Thus, the denial of excluded middle is formalized as $\neg \vdash (A \vee \neg A)$. But my point still holds: if we abandon classical negation, then denials cannot be *merely* assertions of negations.

[11] I have given a semantics for denials and statements that assert that some proposition is denied in (18).

[12] One might object that I have just used disjunctive syllogism – which notoriously is invalid in relevant logic – to argue for a thesis. This is true, but I interpret disjunctive syllogism as an enthymeme, which has a valid form. See (17).

Note that there is a more important difference here between our locally Tarskian circumstances and the closed worlds of artificial intelligence. We do not always assume that our situations are closed in every way, but only in particular ways. Because of sensory clues, background information, and so on, we sometimes assume that there are no more facts of some salient sort. But as Bratko says in the above quote, we usually do not assume that the information available to us is all the information in the universe on a given topic. Thus, our assumptions are much more targeted than the "closed world assumption" to which artificial intelligence appeals.

The idea is that when we come to think that a circumstance is locally Tarskian, that there is a range of predicates that we can make generalizations about if the salient exhaustive induction is successful. In our example of the peaches in the bowl, we can infer after looking at smelling each peach that all of the peaches in the bowl appear ripe, but we cannot merely on the basis of this deduce that all of them came from the same tree. We are not deductively warranted in making generalizations using predicates that explicitly or tacitly refer to entities that are not present in the current situation. Our world is closed only with regard to those entities that are present and the properties and relations that they have to one another. We are not deductively warranted in making generalizations about things that are not present.

11 Varying Domains

In the previous section, we appeal to the idea that certain entities can be present in a situation and others not present. In order to make sense of this distinction in our semantics, we need to have the domains of individuals vary from situation to situation. In (19), we have a single domain of individuals, but Goldblatt and I have modified this semantics to allow for variable domains.

First, we add an existence predicate 'E' into our language. We take our quantifiers to have existential import. That is, when we use $\forall$ we mean 'every existing thing'. The same holds for the existential quantifier, but as we have said we are avoiding mention of it in this paper. We change our natural deduction rules for the universal quantifier as well. Our elimination rule becomes:

$$\begin{array}{l} \forall x A_\alpha \\ \mathrm{E}\tau_\beta \\ \vdots \\ A[\tau/x]_{\alpha\cup\beta} \quad \forall E_e \end{array}$$

where x is free for τ in $A(x)$. This rule tells us that when we instantiate a universally quantified variable we assume that the term we are using refers to something that exists. The introduction rule is similar:

$$\begin{array}{c|l} y & \mathrm{E}x_{\{k\}} \\ & \quad\vdots \\ & A(y)_\alpha \\ \hline \multicolumn{2}{c}{\forall x A(x)_{\alpha-\{k\}} \quad \forall Ie} \end{array}$$

This version of the introduction rule in some ways makes more sense than the original version. In the original version, we are told to assume the existence of an object just called 'x' and then show that it is A, but the assumption isn't explicitly stated in the proof (at least not in the object language). Here the existence assumption is explicitly stated in the proof in the object language.

In this version of the semantics, having a piece of general information, $\forall x A(x)$, is just to have the licence to infer from the existence of something to its being A. Thus, if we have that licence we are in possession of that general information and vice versa. Thus, now our information condition for the universal quantifier says:

$$s \models \forall x A \text{ if and only if there is a proposition } \pi \text{ such that}$$
$$\text{(i) } \pi \text{ entails } |\mathrm{E}x \rightarrow A[c/x]| \text{ for every name } c \text{ and (ii) } s \text{ is in } \pi.$$

Thus, a general piece of information has become a licence to infer from the existence of an entity to its having a property of some sort.

12 Summary

In this paper, I set out a theory of the universal quantifier in an informational semantics. According to this theory, the information that $\forall x A$ is carried by a situation s if and only if a proposition obtains at s which entails every instance of A. This proposition is called a "general proposition". This information condition for 'all' is more reasonable then the standard Tarskian condition in an informational semantics, since informational semantics contain indices with partial information about the world. If we do not know that a list of facts is exhaustive in the salient ways, then we are not entitled to draw general conclusions from it. More information is needed. The use of general propositions, thus, fills the informational gap between the particular facts contained in a situation and generalizations that we draw from that situation. I argue that this use of general propositions is mirrored by the way that we actually make generalizations in very different sorts of circumstances. At the end of the paper, I briefly present a modification of the view to allow the domain of individuals to vary from situation to situation.

13 Appendix: The Formal Semantics

The model theory I present here is essentially the same as (19), but there is an important change. The model does not contain a set of propositional functions as in (19), but instead distinguishes between premodels and models, following Goldblatt and Hodkinson (13). If the reader compares the current approach to that of (19), he or she will see that the current approach is significantly simpler and makes the soundness proof much easier.

A constant domain model structure is a sextuple $\mathcal{S} =< S, 0, R, C, I, Prop>$ such that the following hold. S is the set of situations of the model. 0 is a non-empty set known as the "logical situations". It is at the situations in 0 that the theorems of the logic are verified. R is a ternary relation on S. R is used to treat implication – it is used to formalize the theory of situated inference (see chapter 3 of (17)). C is the compatibility relation discussed in section 4 above – it is a binary relation on S. I is the set of individuals of the model. We need only specify that I is a non-empty set (of individuals). *Prop* is the set propositions – it is a set of subsets of S.

In order that this sextuple be a model structure for the logic **QR** it also must obey the semantic postulates given below. Where $s \sqsubseteq t$ iff $\exists x(x \in 0 \ \& \ Rxst)$ and s^* is a $\sqsubseteq$-maximal situation compatible with s,

SP1 $\sqsubseteq$ is transitive and reflexive (not necessarily anti-symmetric).

SP2 $Rsss$ (full reflexivity, to guarantee $((A \to B) \wedge A) \to B$).

SP3 If $Rstu$, then $Rsut$ (to guarantee $A \to ((A \to B) \to B)$.

SP4 If $\exists x(Rstx \ \& \ Rxuv)$, then $\exists x(Rsux \ \& \ Rxtv)$ (this is the so-called Pasch postulate to guarantee the transitivity of implication).

SP5 If $Rstu$ and $s' \sqsubseteq s$, then $Rs'tu$ (this guarantees the persistence of implicational information, i.e. if $a \models A \to B$ and $a \sqsubseteq b$, then $b \models A \to B$).

SP6 For all $s \in S$, s^* exists.

SP7 If $Rstu$, then Rsu^*t^*(this guarantees that contraposition holds).

SP8 $s^{**} = s$ (this guarantees that double negation elimination holds).

Before we can define models for our language, we need to set out the language itself. It is a standard first order language with relation symbols of various arity, individual constants, countably many individual variables $(x_1, x_2, ...)$, the unary connective $\neg$, the binary connectives $\wedge$ and $\to$, the quantifier $\forall$, and parentheses. The logic **QR**, which is here being modelled, is in a Hilbert-syle axiomatization Anderson and Belnap's logic **R** (see (2)) together with the axiom, $\forall x A \to A[\tau/x]$, where x is free for τ in A, and the rule $\vdash A \to B \Longrightarrow \vdash A \to \forall x B$, where x does not occur free in A.

As I said above, I follow (13) in distinguishing between premodels and models. Here is their definition of a premodel (adapted from modal logic to relevant logic). A *premodel for* **QR** is a pair $\mathcal{M} =< \mathcal{S}, |\cdot|^{\mathcal{M}} >$ where S is a model structure and $|\cdot|^{\mathcal{M}}$ is an interpretation function on the language. $|\cdot|^{\mathcal{M}}$that assigns

- to n-place relation symbol of our language P, a function $|P|^{\mathcal{M}} : I^n \longrightarrow$*Prop;*
- to each individual constant c, an individual $|c|^{\mathcal{M}} \in I$.

The interpretation is then used to give information conditions for all the sentences of the language, by means of a fairly standard inductive definition. Before we give the rest of the definition, we need the notion of a variable assignment function. An assignment function f is a function from the natural numbers ω into the set of individuals I. If $f(n) = i$, then f assigns i to x_n. We use the symbol '$f[i/n]$' to mean a variable assignment function just like f except that it assigns i to x_n. Moreover, for any term (individual variable or constant) τ, $\tau^{\mathcal{M}}f$ is $|\tau|^{\mathcal{M}}$ if τ is a constant and $f(n)$ if τ is x_n. Now we can give the inductive clauses defining $|\cdot|^{\mathcal{M}}f$:

- $|P(\tau_1 ... \tau_n)|^{\mathcal{M}}f \;=\; |P|^{\mathcal{M}}(\tau_1^{\mathcal{M}}f, ..., \tau_n^{\mathcal{M}}f) \in Prop$;
- $|A \wedge B|^{\mathcal{M}}f \;=\; |A|^{\mathcal{M}}f \cap |B|^{\mathcal{M}}f$;
- $|\neg A|^{\mathcal{M}}f \;=\; \{s \in S : \forall x(Csx \;\supset\; x \notin |A|^{\mathcal{M}}f)\}$;
- $|A \rightarrow B|^{\mathcal{M}}f \;=\; \{s \in S : \forall x \forall y((Rsxy \;\&\; x \in |A|^{\mathcal{M}}f) \;\supset\; y \in |B|^{\mathcal{M}}f)\}$;
- $|\forall x_n A|^{\mathcal{M}}f \;=\; \sqcap_{i \in I}(|A|^{\mathcal{M}}f[i/n])$.

Where X is a set of sets of situations, $\sqcap_{i \in I} X = \cup\{\pi \in Prop\colon\; \pi \subseteq \cap X\}$.

A premodel $\mathcal{M}$ is a *model* if for all assignments to variables f and all formulas of the language A, $|A|^{\mathcal{M}}f \in$*Prop.* A formula A is valid on the class of all models if for each model $\mathcal{M}$, $0 \subseteq |A|^{\mathcal{M}}f$. In order to define a valid deduction, we define a generalization of our ternary accessibility relation:

$$R^0 st \text{ iff } s \sqsubseteq t$$

$$R^1 stu \text{ iff } Rstu$$

$$R^{n+1} s_1 ... s_{n+1} s_{n+2} t \text{ iff } \exists x(R^n s_1 ... s_{n+1} x \;\&\; Rxs_{n+2}t).$$

Since it is understood that these relations are just products of the original ternary relation, we drop the superscripts on the symbol R. A deduction of B from $A^1_{\{1\}}, ..., A^n_{\{n\}}$ is valid on a model if and only if, for every assignment f and for all situations $s_1, ..., s_n$ and t, if $Rs_1 ... s_n t$ and $s_k \in |A^k|^{\mathcal{M}}f$ (for $1 \leq k \leq n$), then $t \in |B|^{\mathcal{M}}f$. It can be seen with a little work that the rules of $\forall I$ and $\forall E$ preserve validity in models (as do the other rules discussed above).

In order to accommodate variable domains, we add to our language a unary predicate 'E'. 'Eτ' means 'τ exists'. We also add to our definition of a model structure a function D from situations to subsets of I. Then we have $s \in |\mathrm{E}\tau|^{\mathcal{M}}f$ if and only if $\tau^{\mathcal{M}}f \in Ds$. This simply says that an entity exists at a situation if it is in the domain of that situation. We modify our information condition for the universal quantifier so that it now says:

$$|\forall x_n A|^{\mathcal{M}}f \;= \sqcap_{i \in I} |\mathrm{E}x_n \rightarrow A|^{\mathcal{M}}f[i/n]$$

We now have a semantics that validates $\forall Ie$ and $\forall E_e$. The logic that results from adding these rules to the rules for Anderson and Belnap's **R** is both sound and complete over this semantics. In order to axiomatize this logic, we add to the logic **R** the axiom $\forall xA \rightarrow (\mathrm{E}\tau \rightarrow A[\tau/x])$, where x is free for τ in A, and the rule $\vdash A \rightarrow (\mathrm{E}x \rightarrow B) \implies \vdash A \rightarrow \forall xB$, where x does not occur free in A.[13],[14]

References

[1] A. Ambrose "Mathematical Generality" in M. Lazerowitz and A. Ambrose, *Necessity and Language*, London: Croom Helm, 1985, 101-120

[2] A.R. Anderson, N.D. Belnap, and J.M. Dunn, *Entailment: Logic of Relevance and Necessity, Volume II*, Princeton: Princeton University Press, 1992

[3] D.M. Armstrong, *A Combinatorial Theory of Possibility*, Cambridge: Cambridge University Press, 1989

[4] D.M. Armstrong, *A World of States of Affairs*, Cambridge: Cambridge University Press, 1997

[5] D.M. Armstrong, *Truth and Truthmakers*, Cambridge: Cambridge University Press, 2004

[6] J. Barwise "Constraints, Channels, and the Flow of Information" in P. Aczel, D. Israel, Y. Katagiri, and S. Peters (ed.), *Situation Theory and its Applications*, volume 3, Stanford: CSLI, 1993, 3-27

[7] J. Barwise and J. Etchemendy, *The Liar*, Oxford: Oxford University Press, 1987

[8] J. Barwise and J. Perry, *Situations and Attitudes*, Cambridge, MA: MIT Press, 1983

[9] Ivan Bratko, *Prolog Programming for Artificial Intelligence*, Workingham, UK: Addison-Wesley, 1986

[10] J.M. Dunn "Star and Perp" *Philosophical Perspectives* 7 (1993): 331-57

[11] K. Fine "Semantics for Quantified Relevant Logic" Journal of Philosophical Logic 17 (1988): 27-59; reprinted in (2) §53.

[13] We allow statements about things that do not exist at a situation to be true at that situation (in my opinion, we want to allow statements such as 'if Sherlock Holmes is a famous detective, then he is famous' to be true in some actual situations).

[14] Acknowledgements: I am especially grateful to Rob Goldblatt, with whom I have developed the theory of quantification discussed in this paper and to Sebastian Sequoiah-Grayson who commented extensively on an early version. Early drafts of this paper were read at the Australasian Association of Philosophy meeting in Wellington, Korea National University, and the Workshop on Philosophy of Information and Logic in Oxford. This paper was also given as the 2007 Alice Ambrose Lazerowitz/Tomas Tomoczko Memorial Lecture at Smith College. I am grateful to those audiences, especially Phil Bricker, Luciano Floridi, Jay Garfield, Carrie Jenkins, Sehwa Kim, Dan Nolan, Josh Parsons, Johann Van Benthem, and Mark Jago.

[12] K. Fine "Incompleteness of Quantified Relevance Logics" in J. Norman and R. Sylvan (eds.), *Directions in Relevant Logic*, Dordrecht: Kluwer, 1989

[13] R. Goldblatt and I. Hodkinson "Commutativity of Quantifiers in Varying-Domain Kripke Models" in D. Makinson, J. Malinowksi, and H. Wansing (eds.), *Trends in Logic: Towards Mathematical Philosophy*, Heidelberg: Springer, forthcoming

[14] R. Goldblatt and E.D. Mares "A General Semantics for Quantified Modal Logic" in G. Gouvenatori, I. Hodkinson, and Y. Venema (eds.), *Advances in Modal Logic, Volume 6*, London: College Publications, 2006

[15] J. Heil, *From an Ontological Point of View*, Oxford: Oxford University Press, 2003

[16] E.D. Mares "A Star-Free Semantics for R" *The Journal of Symbolic Logic* 71 (2006) pp 163-187

[17] E.D. Mares, *Relevant Logic: A Philosophical Interpretation*, Cambridge: Cambridge University Press, 2004

[18] E.D. Mares "The Semantics of Denial" in B. Brown and F. Lepage (eds.), *Truth and Probability: Essays in Honour of Hughes Leblanc*, London: College Publications, 2006, 49-58

[19] E.D. Mares and R. Goldblatt "An Alternative Semantics for Quantified Relevant Logic" *The Journal of Symbolic Logic* (2006)

[20] G. Restall "Information Flow and Relevant Logics" in J. Seligman and D. Westerståhl (eds.), *Logic, Language and Computation*, Stanford: CSLI, 1996, 463-477

[21] B.A.W. Russell, *The Philosophy of Logical Atomism* in Russell, *Collected Papers*, Volume 8, London: George Allen and Unwin, 1986, 157-244