SRI International

CONCEPTS OF INFORMATION: COMPARATIVE AXIOMATICS

Technical Note 469

June 12, 1989

By: David Israel, Sr. Computer Scientist

Artificial Intelligence Center
Computer and Information Sciences Division
and
Center for the Study of Language and Information
Stanford University

APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITED

To appear in *Artificial Intelligence and Philosophical Logic*, edited by R. Thomason, Kluwer, 1989.

The research reported in this paper has been made possible by a gift from the System Development Foundation.

333 Ravenswood Ave. • Menlo Park, CA 94025
(415) 326-6200 • TWX: 910-373-2046 • Telex: 334-486

Concepts of Information: Comparative Axiomatics

David Israel
Artificial Intelligence Center and CSLI
SRI International

1 Introduction

There is much talk about this being the *Age of Information* and about a Post-Industrial Revolution centered on *information processing*. But what exactly is *information*?[1]

Jon Barwise has drawn a potentially unflattering analogy between the Bronze Age and the Information Age.[2] People in the Bronze Age were quite expert at working with bronze, but it was a long time after the end of the Bronze Age that scientists were able to determine the true nature of bronze. Indeed, by the time researchers got around to providing the requisite materials for an adequate theory of the nature of bronze, all the denizens of the Bronze Age were long dead. *Mutatis mutandis* for us and information? The fond hope is that what follows will go some distance toward allaying fears of history's repeating itself in this regard.

One difference, at any rate, between that age and this is that there are many different notions of information afoot. There are even many different notions within information theory: the (average) amount of information generated by a source (the entropy of the source), the (average) amount transmitted along a channel, and so forth. There are also the related, but distinct, notions from the theory of algorithmic complexity.

---

[1]The research reported in this paper has been made possible by a gift from the System Development Foundation. Very special thanks are due to John Perry and to Rich Thomason.

[2]"Information and Circumstance," *Notre Dame Journal of Formal Logic* 27, No. 3, (July 1986), pp. 307-323; reprinted in *The Situation in Logic*, CSLI Lecture Notes, No. 17.

These are all quantitative concepts in highly developed, mathematically rigorous theories. They have played almost no role in artificial intelligence (AI). On the other hand, notions of content have played an increasingly important role in AI. The point is especially clear once we note that such *semantic* concepts as meaning, truth, and reference concern content and not average quantity of information generated or transmitted. Kenneth Sayre put this contrast as follows:

> The difference between information theory and semantics is the difference between the study of the conditions for the communication of any message whatsoever and the study of the content of particular messages.[3]

Though the idea of the semantic content has become central, very little has been done to provide systematic analyses of the notion of information content. In most cases, the meaningfulness of the interpretable items, often expressions of a formal language, and their particular meanings are simply assumed as given. But one may aim, instead, at a language-independent account of information content for a wide range of interpretable events or objects.

There are various ways of answering the question, "What is Information?" One is conceptual analysis; another is by way of axiomatizing the concept, combined perhaps with analysis of models of the axiomatization. A relevant comparison case is that of *knowledge*. Philosophers tried, literally for millenia, to analyze the notion, arguably with little success. Beginning in the late fifties and early sixties, researchers instead began to axiomatize one or another notion of knowledge. Most notable among these attempts were those of Jaakko Hintikka.[4] The case of knowledge presents more than a useful methodological analogy, though. Current work in AI and in theoretical computer science can be seen as presenting axiomatizations of a notion of information based more or less directly on the earlier work in epistemic modal logics. In the remainder of this introduction, I shall remind the reader of two axiomatizations of knowledge. The first is the standard—that is, most widely accepted—formalization of knowledge as an $S4$ modal operator; the second, widely rejected as appropriate for *knowledge*, simply adds the characteristic $S5$ axiom. The reason for presenting this second axiomatization

[3] "Philosophy and Cybernetics," in *Philosophy and Cybernetics*, Kenneth Sayre and Frederick Crosson (eds.), NY, 1967, p.11.

[4] *Knowledge and Belief*, Cornell University Press, Ithaca, 1962.

is that the contributions from AI and computer science that we shall look at do propose $S5$ as the appropriate epistemic logic.

One may very well doubt that there is any intuitively acceptable notion of *knowledge* for which $S5$ is appropriate. We argue that it may be appropriate, however, for a certain notion of *information.*

The second section contains brief and simplified expositions, first of a theory of information due to John Perry and myself, and then of work of the computer scientists Joseph Halpern and Yoram Moses.[5] We include as well a very brief summary of the closely related work of Stan Rosenschein. A comparison follows, the main theme of which is to see which conceptions validate which axioms. The paper concludes with an attempt to explain these differences.

1.1 Epistemic modal logic

We shall limit ourselves to the propositional case. We assume a language $\mathcal{L}_0$ for the classical propositional calculus. We abstract from such details as the set of truth-functional connectives taken as primitive, so long as it is functionally complete. We extend $\mathcal{L}_0$ to a language $\mathcal{L}$ by adding one unary sentence operator K_α, to be read as α *knows that.* Here the subscripted α can be understood in a number of different ways. For present purposes, we assume that there we are axiomatizating the knowledge of an idealized subject, to whom all actual subjects are more or less pale approximations. This interpretation allows us to suppress the subscript, which for the remainder of this section we shall do.

$S4$ is the smallest set of wffs of $\mathcal{L}$ containing all classical tautologies, all instances of the following schemata:

1. $K(p \rightarrow q) \rightarrow (Kp \rightarrow Kq)$
2. $Kp \rightarrow p$
3. $Kp \rightarrow K(Kp)$

and closed under *Modus Ponens* and *Necessitation* (if $p \in S$, then $Kp \in S$).

The $S5$ schema is as follows:

4. $\neg Kp \rightarrow K(\neg Kp)$

We assume familiarity with the now standard Kripke-style semantics for S's 4 and 5. A crucial difference for what follows between the two is that $S5$, with the operator now read as "it is necessary that" (usually written as L or as $\Box$) is a likely candidate for the logic of absolute or metaphysical

[5]See also the paper by Halpern and Ron Fagin, in this volume.

necessity. Surely truth in all metaphysically possible worlds is just truth in *all* possible worlds. The operator is then understood as a quantifier whose domain is the collection of all possible worlds, and no notion of an accessibility relation between worlds is needed.[6] Indeed, in contexts where metaphysical necessity is at issue, $S5$ is still presented this way. It is only within the general context of semantics for modal logics that one notes that in $S5$ structures the accessibility relation is an equivalence relation, whereas in $S4$ structures, the accessibility relation is reflexive and transitive only.

Consider the property of being a necessary truth. Intuitively, this is a property of propositions and, indeed, in $S5$ it can be thought of as such. A proposition is necessary just in case it is true in all possible worlds. In $S4$, however, as in other modal systems weaker than $S5$, necessity is a relation between propositions and worlds. A proposition is necessary at a world w just in case it is true in all worlds accessible from w. Here, the relativity to w is ineliminable.

2 The Israel-Perry Conception

2.1 The basic principles

We present a summary version of what we take to be certain central principles governing the relation: *x carries the information that y*.[7] Some terminological preliminaries. We take as canonical information reports sentences like the following:

> (*) The fact that the x-ray has the pattern Φ indicates that Patches has a broken leg.

Here Φ is the characteristic pattern of discontinuity of the shaft, presenting itself as a whitish, slightly jagged hair, where there should be smoothness. Note that the subject of (*) designates a fact or state of affairs rather than an object. We call this the *indicating fact*. We shall limit ourselves to simple cases, where this fact will involve a single object's having some property at

[6] As in Kripke's original completeness theorem for $S5$, "A Completeness Theorem in Modal Logic," *JSL*, Vol. 24, No. 1, March 1959, pp. 1-14.

[7] What follows is a summary of parts of "What is Information," to be published in *Information, Language, and Cognition: Vancouver Studies in Cognitive Science, Vol. I*, P. Hanson (ed.) Vancouver, University of British Columbia Press, 1989 (in press). For the the overall tenor of our account, we owe a great debt to Fred Dretske's *Knowledge and the Flow of Information*, MIT Press, Cambridge, MA, 1981.

some time, in our case the x-ray's having a certain deposition pattern at t. We call this object—the x-ray sheet—the *bearer* of the information and the property, having the pattern Φ, the *indicating property.* We shall call the proposition designated by the embedded sentence the *informational content.*

In stating the basic principles of information, we use such terms as 'fact' and 'state of affairs' in something akin to one of their ordinary senses; in the next section we explicate them within situation theory.

(A) What carries information is a fact.

(B) The information that is carried by a fact is a *true* proposition.

We now know the types of the arguments of the relation. (B), of course, is simply our way of expressing the condition that the relation is *truth-involving.* People do often speak of *misinformation*, and speak of information in a way that is clearly neutral as between truth and falsity. However that may be, the relation we are focusing on is one between facts and *true* propositions.

What underlies the phenomenon of information is the fact that phenomena are lawlike; they are subject to laws, nomic regularities, principles or as we shall say, *constraints.* It is in virtue of the fact that reality is lawlike, that facts about what is going on in one part of reality can carry information about what is the case perhaps in some other, quite remote part.

We shall give an account of constraints as a special sort of fact, in which one type of situation *involves* another. We shall have little to say about the involvement relation, beyond saying that it is at least as strong as Hume's relation of constant conjunction between types of situation. In a world knitted together by constraints—whether these be constant conjunctions or some more metaphysically potent connections—situations, in virtue of the states of affairs they make factual, carry information. The fact that there is a situation of some type carries the information that there are situations of the types that are "involved in" that type.

This conception licenses the notion of the information carried by a fact relative to a constraint. And from that notion we can derive the notion of the information carried by a situation, *simpliciter*, as that information carried by the situation relative to some constraint or other. We can also derive the notion of the *total* information carried by a situation—that which it carries relative to all actual constraints. But it is not clear that either of the latter two is a useful notion. The useful notion, and the one that we think is implicit in our actual thinking about information, is of information *relative* to a constraint or to a set of constraints, fixed typically by the context—by our purposes and interests, theoretical and practical.

The treatment of the relation of *carrying information* that we present here abstracts from the phenomenon of the use of information by agents in the control of behavior—their own and that of other agents and devices. Despite that, crucial aspects of the treatment are motivated by the project of providing an account of that phenomenon. In particular, this is true of the relativity to constraints that we claim is characteristic of the useful notion of the information carried by a state. Agents may differ with respect to which constraints they or their designers can exploit or by *attunement* to which, their behavior can be governed. Some agents, ourselves for instance, actually set about discovering constraints and can learn to exploit many of those so discovered. But even we have limits in this regard, though it is hard to say what they are. With respect to man-made informational devices, such as x-ray machines and x-rays, or computational artifacts of various kinds, natural constraints or laws may be supplemented by various agent-dependent regularities, such as practices and conventions. Still, it is only certain such laws and regularities that are relevant to the design and use of these devices.[8]

> (C) The information a factual state of affairs carries is relative to a constraint.

This conception does not require that there be any intrinsic connections between facts and the propositions they indicate. If the fact that the x-ray has a certain pattern were embedded in a different sort of world, where different constraints held sway, it would carry quite different information than it actually does.

> (D) The information a fact carries is not an intrinsic property of it.

(D) is an immediate corollary of (C).

The informational content of a fact can involve objects quite remote from those involved in the fact. Patches is not a part or aspect of the x-ray mentioned in (*), nor is he a constituent of the fact that the x-ray has the

[8]We shall return to this point below, when we compare our conception of information with those of Halpern-Moses and Rosenschein.

pattern Φ, but something remote from it. The x-ray is not broken, and does not have bones. Information typically involves a fact indicating something about the way things are elsewhere and elsewhen, and this is what makes information useful and interesting.

> (E) The informational content of a fact can concern remote things and situations.

This conception of information can explain how an x-ray could carry the information that some dog had been x-rayed and had a broken leg. But it is not clear how it can account for the specific information the x-ray carries about Patches that is reported in (*). As we noted above, Patches is not a part of the x-ray, and it does not seem that his having a broken leg could be constantly conjoined with x-rays exhibiting the pattern that the vet recognizes. So how can the informational content of the x-ray have him as a constituent?

We shall call the sort of information involved, for example, in (*), *incremental* information. The conception is most easily grasped if we think of what the x-ray indicates *to* the vet. If she does not know which dog the x-ray is of, it simply indicates that some dog that has a broken leg has been x-rayed. We call this the *pure* information. But if she knows that Patches was x-rayed, then the pattern on the x-ray indicates to her the additional or incremental information that Patches has a broken leg. This is the incremental information carried by the x-ray, given the fact that the x-ray is of Patches. The fact that is given *connects* the indicating situation and the specific objects the information is about, so we shall call it the *connecting fact.* We must be careful that this example does not mislead as to our intentions, however. Incremental information is important in understanding the use humans and other agents make of information, but humans and mental states need not be brought into its analysis. Incremental information about specific objects is an objective feature of the world that is there for us and other agents to use.

> (F) Informational content can be specific as well as remote; the propositions that are informational contents can be about specifc objects that are not part of the indicating fact.
>
> (G) Indicating facts contain such information only relative to connecting facts; the information is incremental, given those facts.

We claim that the interesting informational relation is *doubly relative.* The relation we are usually interested in is that of *a fact's carrying the incremental information that P, relative to a constraint and a set of connecting facts.*

2.2 The framework

We have noted that carrying information is a relation between facts and propositions. We have also said that what underlies information are laws or constraints involving types of situations. So what are all these things? And what precisely are we saying about them?[9]

2.3 Situations

A basic idea of situation theory is that there is a concrete reality, which has concrete parts but not concrete alternatives. This reality can be thought about, perceived, studied and analyzed in a variety of different ways, from a variety of different perspectives, for a variety of different purposes. But ultimately everything that exists, everything that happens, everything that is true, has its status because of the nature of this reality. The parts of reality are what we call *situations.* Situation theory is committed to there being situations, but not to there being a largest total situation of which the rest are parts.

2.3.1 Relations, argument roles, locations, individuals, issues, positive and negative states of affairs

When we think or talk about reality, we need some way of analyzing it. This we call a *system of classification and individuation.* Such a system consists of domains of situations, relations, locations, and individuals. The commonplace that different schemes can be used to study the same reality is one to which situation theory subscribes. But this fact should not be thought of as showing that situations are structureless, with their properties projected onto them by language or thought. Rather, situations are rich in structure and support a variety of schemes, suited (or unsuited) to various needs.

[9] We present here only enough of situation theory to capture a simple case of information carrying, and enough as well to point out the contrast with other approaches based on modal model structures. See the bibliographical note at the end.

Each relation R comes with a set of argument roles. For example, the relation of *examining* comes with the roles of *examiner*, *examinee*, and *location* of examining. Objects of appropriate sorts play these roles. The examiner must be some sort of organism or device. The examinee must be a physical object or quantity of stuff. The location of examining must be a spatio-temporal location.

A relation, together with appropriate objects assigned to its roles, gives rise to an *issue*, namely, the issue of whether or not the objects stand in the relation. There are two possibilities, and each of these we call a *state of affairs.*

Example. If *examining* is the relation, Doctor Smith is the examiner, Patches is the examinee and Dr. Smith's office at a certain time is the location (call it l), then there are the following two states of affairs:

$\ll Examines,\ examiner : Dr. Smith,\ examinee : Patches,\ loc : l;\ 1 \gg$

$\ll Examines,\ examiner : Dr. Smith,\ examinee : Patches,\ loc : l;\ 0 \gg$

The first state of affairs resolves the issue positively, the second negatively. We say the first has a positive and the second a negative polarity.

The relation of *examining* is the major constituent of these states of affairs; Doctor Smith, the location l, and Patches are the minor constituents. The polarities should not be thought of as constituents at all.

We don't assume that the argument roles of a relation have a natural order—that is, an order independent of that in which they are expressed in a given language or in a given construction in a language. But in this paper we shall often use the order suggested by English to identify argument roles, without explicitly mentioning them.[10] For the first state of affairs above, we write:

$\ll Examines,\ Dr. Smith,\ Patches,\ l;\ 1 \gg$

2.3.2 Facts and other states of affairs,makes-factual, the partiality of situations

Situations determine whether a given state of affairs or its dual (the one with the same assignment of constituents to roles but the other polarity) is a fact. We do not assume that every state of affairs has a dual, but we do assume this for every basic state of affairs. These last have basic relations as their major constituent. This primitive relation we call *making factual* or *supporting.*

[10]The argument role for spatio-temporal locations will always be displayed last. Occasionally, we shall leave it out altogether.

$s \models \sigma$ means that s *makes* σ *factual*

We will also make use of the property of *being factual.* A state of affairs is factual iff some real situation supports it.

$\models \sigma$ means that σ *is factual*

The following are uncontroversial theses about the $\models$ relation.
Given a state of affairs and its dual,

- Some situation will make one of them factual.
- No situation will make the other one factual.
- Some situations will leave the issue unresolved, that is, they will make neither of them factual.

The following is a controversial thesis about this important relation:

- Some situation resolves all issues.

This, of course, is the thesis that there is a largest total situation.

The third thesis tells us that situations are partial. They do not resolve all the issues (except, perhaps, for the total situation called for in the fourth thesis). Because of the partiality of situations, we must distinguish betwen two ways a situation s can fail to make a given state of affairs σ factual:

- s may make the dual of σ factual.
- s may fail to resolve the σ issue one way or the other.

2.3.3 Parameters, anchors and compound infons

For the purposes of a theory of information, one needs a wider class of entities, that includes states of affairs and other entities built from them by abstraction and combination. We call this wider class *infons.* All states of affairs are infons, but not vice versa. [11]

We assume a set of *parameters* corresponding to individuals and locations.[12] Where $\ll \ldots, a, \ldots \gg$ is a state of affairs with a as a minor constituent, and

[11] The term "infon" is dne to Keith Devlin.

[12] In "What Are Parameters?" (in preparation), we address the issue of what parameters are, and whether they are a necessary part of situation theory.

x is a parameter, $\ll \ldots, \mathrm{x}, \ldots \gg$ is an *infon*. The step from states of affairs (or nonparametric infons) to parametric infons is a sort of abstraction. To get from the infons back to states of affairs, we need *anchors*. An anchor is a partial function from the domain of parameters to appropriate objects. Where f is an anchor, $\ll \ldots, \mathrm{x}, \ldots \gg [f] = \ll \ldots, f[\mathrm{x}], \ldots \gg$.

An anchor f *satisfies* an infon i relative to a situation s iff $s \models i[f]$. An anchor f satisfies an infon i *simpliciter* iff $\models i[f]$, that is, if there is a situation s such that $s \models i[f]$.

In considering which ways of combining infons are allowed, we are guided by the *principle of persistence*. Given a partial order $\trianglelefteq$ (intuitively, "part of") on situations, two situations, and an infon σ:

$s_1 \trianglelefteq s_2 \,\&\, s_1 \models \sigma \Rightarrow s_2 \models \sigma$.

Intuitively, if a situation supports two infons, any larger situation of which it is a part will also, so the meet of two infons is permitted. If a situation supports one of two infons, any larger situation of which it is a part should also, so the join of two infons is permitted. Indeed, persistence holds for meets and joins of arbitrary sets of infons. We assume a complete lattice of infons. In what follows, we only need the meet of two infons, however. For the meet of two infons σ, $\sigma\prime$, we use the notation $(\sigma \wedge \sigma\prime)$; for the general notion of the *meet* of a set of infons we use $\bigwedge I$.

- $s \models \sigma_1 \wedge \sigma_2$ iff $s \models \sigma_1$ *and* $s \models \sigma_2$.
- f satisfies $\bigwedge I$ iff $i[f]$ is factual for each $i \in I$.

If there is an individual, such that a situation supports an infon with a parameter anchored to that individual, any larger situation of which it is a part will also, so the existentialization of an infon is permitted. For the *existentialization* of an infon with respect to parameter x, $\exists \mathrm{x}(i)$, where $i = \sigma(x)$:

- f satisfies $\exists \mathrm{x}(i)$ iff for some object a, $i[f_{\mathrm{x}/a}]$ is factual.

Suppose, however, that we tried to introduce an operation $\forall$, taking parameters and infons into infons, and such that, for example,

$s \models \forall \mathrm{x} \ll \sigma(\mathrm{x}); 1 \gg$

if for every object a in s, s supports the fact $\sigma(a)$. By an object being in a situation, we simply mean that it is a minor constituent of some state of affairs made factual by s. Still, s will no doubt be a part of many larger situations s', in which there are objects such that $s' \not\models \sigma(a)$. By persistence, there can be no such infon-forming operator.[13]

2.3.4 Types, constraints and involvement

We now define the notions that are at the heart of our account of information.

Where σ is a state of affairs, $[s|\ s \models \sigma]$ is the *type of situation* that supports σ. Where i is a parametric infon, $[s|\ s \models i]$ is a *parametric type*, and i is the *condition* of T (*cond(T)*). A situation s is *of* parametric type T relative to f if $s \models i[f]$, where i is the condition of T and f is defined on all of the parameters of i.

Since parametric infons and parametric types are the entities most used from now on, we shall mean parametric types when we say 'types'; nonparametric types may be thought of as the special case.

We take constraints to be states of affairs with types of situations as constituents. *Simple involvement* is a binary relation. If T involves T', then for every situation of type T, there is a situation of type T'.[14] We write:

$$\ll \mathit{Involves},\ T,\ T',\ 1 \gg$$

Relative involvement is a ternary relation. If T involves T' relative to T'', then, for any pair of situations of the first and third types, there is a situation of the second type. We write:

$$\ll \mathit{Involves}_R,\ T,\ T',\ T'';\ 1 \gg$$

For simplicity's sake, we assume that constraints are unlocated. In particular the notion of relative involvement (relative constraints) should not be confused with that of *conditioned constraints*, introduced in *Situations and Attitudes*.[15]

[13] One could imagine allowing some form of bounded universal quantification, but what should act as the bound? It could, of course, be an arbitrary set, quite unrelated to the situation s. Or it could be a subset of the collection of all objects that are in s. In any event, for present purposes we simply ignore such possibilities.

[14] Note that the definition does not require that when s is of type T that *it* also be of type T'. There are also *negative constraints*. We posit a relation of *preclusion* between types of situation. One type of situation precludes another just in case if there is a situation of the first type, then there is no situation of the second.

[15] And discussed further in Jon Barwise's "The Situation in Logic-II: Conditionals and Conditional Information," CSLI Report No. 21, 1985; reprinted in *The Situation in Logic*, CSLI Lecture Notes, No. 17, 1989.

2.3.5 Propositions, and more types

We take propositions to be nonlinguistic abstract objects that have absolute truth values. From the perspective of situation theory, this means that a proposition requires not only a type—that which corresponds or doesn't correspond to the way things are—but also a situation for the type to correspond to. Two basic kinds of propositions are recognized in situation theory. An *Austinian proposition* is determined by a situation and a type, and is true if the situation is of the type. A *Russellian proposition* is determined by a type alone, and is true if some situation or other is of that type. If we adopted the fourth thesis, that there is a (unique) total situation, Russellian propositions could be taken to be Austinian propositions determined by this total situation.[16]

Propositions are not infons. Infons characterize situations; propositions are truth bearers. We shall assume that for each type of situation and each situation there is an Austinian proposition that is true just in case that situation is of that type. In what follows, we shall limit ourselves to Russellian propositions; with respect to them, we shall assume that for each type, there is a proposition that is true just in case some situation is of that type.[17]

Given our commitment to the persistence of infons, infons are closed under meet, join, and existentialization. Thus, if we want there to be a domain of objects closed under other logical operators, we must posit at least one algebra of objects more complex—closed under more operations—than that of infons. We posit two such: an algebra of types and an algebra of propositions. For the purpose of presenting a sketch of our account of information, we need no more complex types than those already furnished by abstraction over infons. For the sake of the comparison with other approaches, however, we do need one thing more—negative types and negative propositions.

Not all types are based on infons. Where σ is an infon, there is a type $\neg T_\sigma = [s | s \not\models \sigma]$. We can extend the notion of persistence to types. Infon-based types are persistent; those like $\neg T_\sigma$, not so based, need not be. In fact, we assume that compound types are formed by abstracting over propo-

[16]See Jon Barwise and John Perry, *Situations and Attitudes*, pp. 139-40. The distinction is further clarified and the present terminology introduced in Jon Barwise and John Etchemendy, *The Liar: An Essay in Truth and Circularity*, Oxford University Press, 1987, where it plays a key role in the treatment of semantic paradox. See also the postscript to the 2nd edition, forthcoming.

[17]This last is a strong assumption, that can lead to paradox. We shall not concern ourselves with such issues in the present essay; see Barwise and Etchemendy's *The Liar*.

sitions. Thus, in the case at hand we posit, for every Russellian proposition $(\models T)$, a negation $(\neg \models T)$.[18] We will assume that the domain of situations with which we work is large enough to mimic the world, relative to the types we allow. In particular, the domain of situations is large enough so that for each type T, the Russellian proposition determined by T is either true or false and the denial of a proposition is true iff the proposition is false. These assumptions, together with the foregoing, allow us to derive a Boolean algebra of propositions.

Infons may have individuals and locations as constituents. When an infon with an individual or location as a constituent is the condition of a type, then we also say that the type has that individual or location as a constituent, as does the proposition determined by that type. A proposition whose type contains no such constituents, because each argument role has been quantified over, is *general*, in David Kaplan's terminology; others are *singular*.[19] We shall say that a singular proposition is *about* its constituents.

2.4 Information

We now turn to constructing our theory of information within the version of situation theory just sketched.

Let C be some constraint. The fact σ *carries the pure information that* P *relative to* C iff

1. $C = \ll Involves, T, T'; 1 \gg$.
2. For any anchor f such that $\sigma = cond(T)[f]$, $P =$ the proposition that $\exists s'(s' \models \exists a_1, ...a_n(cond(T')[f]))$.

Informally, we would have the following in the case of the x-ray:

> The x-ray's being Φ-ish indicates that there is a dog, of whom this is an x-ray, and that dog has a broken leg.[20]

[18] This way of proceeding leads to great complexities in the case of compound Austinian propositions, in that the component propositions can concern distinct situations. This difficulty, at least, is avoided by the limitation to Russellian propositions.

[19] See Kaplan's "Demonstratives." This widely circulated and influential manuscript is to be published in *Themes from Kaplan*, edited by J. Almog, J.R. Perry, and H. Wettstein, Oxford University Press, 1989.

[20] In what follows, we shall simply assume that the x-ray's being Φ-ish indicates that it is of a dog's leg.

We are often interested in more specific information. For instance, to guide her action appropriately, the vet has to know which dog. It is not enough for her to be acquainted with the indicating fact and aware of the constraint. She must know that the x-ray was of Patches's leg; she must know that the information it carries is *about* Patches—what we have called *incremental* information. To capture the notion of incremental information, we need a more complex constraint, one of relative involvement. The additional type in such constraints is the *connecting type*, the type of the *connecting situation.* We also call such constraints *relative constraints.*

Let C be some relative constraint. Then the fact σ *carries the incremental information that P relative to C and the fact σ'* iff

1. $C = \ll Involves_R,\ T,\ T',\ T'';\ 1 \gg$.

2. For any anchor f such that $\sigma = cond(T)[f] \wedge \sigma\prime = cond(T'')[f]$, $P =$ the proposition that $\exists s'(s' \models \exists a_1, \ldots, a_n(cond(T')[f]))$.

Again, informally, in our case, the connecting fact is that the x-ray in question is of Patches's leg, and it is in virtue of Patches's being a constituent of this fact that he is a constituent of the indicated proposition, the proposition that Patches's leg is broken.[21]

2.4.1 An application of the theory

Let's now apply the theory more formally and fully to our example involving Patches's leg and the x-ray. We can consider this as a case of *pure information* or of *incremental information.*

In both cases, the indicating fact σ is the x-ray's being of a certain type at t. When we consider the pure information, we have in mind the following simple constraint: if there is a state of affairs consisting of some x-ray's having the pattern Φ at some time t, then there is a state of affairs involving a dog's leg having been the object of that x-ray and that leg's being broken at t. So the indicated proposition is that there is a dog of which this is the x-ray, and it has a broken leg. The pure information is about the x-ray, but not about Patches, or his leg.

For the sake of simplicity, we have assumed that the x-ray is developed essentially instantaneously. We are also making a much more important

[21] As we shall see, this reflects the fact that the anchor for the connecting type, that is for T'', must assign Patches to the role of being the object whose leg is x-rayed and thus the indicated type, T'—and the indicated proposition—will be about him.

simplifying assumption. Very few useful regularities meet the Humean condition of *constant* conjunction. Most regularities are *local*; they hold or are valid only within a limited range of environments. Another way to put this is that they are valid only conditionally.[22] Moreover, in many cases it is (practically) impossible to fold the relevant conditions into the antecedent type. Even where it is possible, it is often misleading, leading to a mischaracterization of the capacities or functionality of the information-sensitive device or agent. In what follows, however, we idealize and treat all the relevant constraints as though they were unconditioned.[23]

We now return to the main theme. Using the resources of situation theory, we represent the simple constraint as follows:

$$T = [s|\, s \models \ll \text{X-ray, x, t; } 1 \gg \;\wedge \ll \text{Has-pattern-}\Phi\text{, x, t; } 1 \gg]$$

$$T' = [s|s \models \ll \text{Is-xray-of, x,y,t;} 1 \gg \wedge \ll \text{Has-Broken-Leg,y,t;} 1 \gg]$$

$$C = \ll \text{Involves, } T,\ T';\ 1 \gg$$

The indicating situation, σ, is

$$\ll \text{X-ray, } a,\ t';\ 1 \gg \;\wedge\; \ll \text{Has-pattern-}\Phi,\ a,\ t';\ 1 \gg$$

where a is the x-ray and t' the time. We assume that σ is factual, that is that $\exists s(s \models \sigma)$. Now let f be any anchor defined on x and t (at least) such that

$$\sigma = cond(T)[f] = \ll \text{X-ray, x, t; } 1 \gg \;\wedge \ll \text{Has-pattern-}\Phi\text{, x, t; } 1 \gg [f]$$

Thus, $f(\mathrm{x}) = a$ and $f(\mathrm{t}) = t'$. Then $P =$ the proposition that

$$\exists s'(s' \models \exists y(\ll \text{Is-xray-of, x, } y\text{, t; } 1 \gg \wedge \ll \text{Has-Broken-Leg, } y\text{, t; } 1 \gg)[f])$$

Thus P is the proposition that the state of affairs that consists of some dog being the object of a, the x-ray in question (at t', the time in question) and that dog's having a broken leg (at the time in question) is factual. Or, more simply, it is the proposition that there is some dog whose leg is depicted by a at t' and whose leg is broken at t'.

[22] Compare the discussion of *conditioned constraints* in *Situations and Attitudes*.

[23] A related point can be made about the work of Halpern-Moses and of Rosenschein, except that the former are dealing with a design or specification problem. They can limit consideration to those runs of a distributed system in which everything goes according to the specifications. See below.

When we consider this as a case of incremental information, we have in mind the relative constraint that if an x-ray is of this type, and it is the x-ray of a dog, then that dog had a broken leg at the time the x-ray was taken. The fact that the x-ray was of Patches is the connecting fact, and the incremental informational content is the proposition that Patches has a broken leg. This proposition is about Patches, but not about the x-ray.

The relevant relative constraint is:

$$C' = \ll Involves_R,\ T,\ T',\ T'';\ 1\gg$$

where T, the indicating type is as before. T', the indicated type is:

$$[s|\ s \models \ll \text{Has-Broken-Leg},\ \mathbf{y},\ \mathbf{t};\ 1\gg]$$

and T'', the connecting type is:

$$[s|\ s \models \ll \text{Is-xray-of},\ \mathbf{x},\ \mathbf{y},\ \mathbf{t};\ 1\gg]$$

As before, σ is:

$$\ll \text{X-ray},\ a,\ t';\ 1\gg\ \wedge\ \ll \text{Has-pattern-}\Phi,\ a,\ t';\ 1\gg$$

Again, we assume that σ is factual. Further, we assume that the connecting state of affairs, σ' is factual. Where b is Patches, σ' is:

$$\ll \text{Is-xray-of},\ a,\ b,\ t';\ 1\gg$$

Any anchor f, such that $\sigma = cond(T)[f]$ *and* $\sigma\prime = cond(T')$, must be defined on the parameter $\mathbf{y}$ of the connecting type, in particular, it must anchor $\mathbf{y}$ to Patches. Thus, for any such anchor f, the proposition carried incrementally by σ relative to C and $\sigma\prime$ is the proposition that

$$\exists s''(s'' \models \ll \text{Has-Broken-Leg},\ b,\ t';\ 1\gg)$$

This is a singular proposition about Patches, and not at all about the x-ray. And it is, after all, Patches that we're concerned about.

3 The Halpern-Moses, Rosenschein Conceptions

We turn now to brief descriptions of the framework developed by Halpern and Moses and to an even briefer summary of Rosenschein's theory of situated automata.

This research is of interest for a number of reasons, but for present purposes what is key is that the operator(s) introduced and studied are really best thought of in terms of one thing's carrying information that P, and not in terms of a person's knowing that P. There are (at least) three reasons for saying this.

The phenomena being modeled can be characterized as follows: there is a computational system with a number of distinct components or subsystems. Such a component might be a register within a machine, or a processor in a distributed network. These kinds of application are fairly far removed from the phenomena that Hintikka and other philosophers have had in mind. They were interested in the logic of attributions of knowledge to people and/or to idealized versions of people. However anthropocentric or retrograde our reluctance to attribute beliefs or knowledge to imaginable robots, it is perhaps not overfastidious to resist such attributions to their component registers. The same goes (even) for VAXen hooked together in distributed systems.

The second reason has already been touched on. The characteristic $S5$ axiom and its weaker cousin the symmetry axiom seem completely implausible for knowledge. But we shall see that they are not so implausible for at least one notion if carrying information. This then is our third reason for looking at these proposed formalizations of knowledge as formalizations instead of some notion of carrying information.

3.1 Distributed systems

We shall draw on the exposition in the paper by Halpern and Moses, *Knowledge and Common Knowledge in a Distributed Environment* (the revised version).[24]

A distributed system is modeled as a finite collection $\{p_1, p_2, \ldots, p_n\}$ of processors that communicate with each other by sending messages along a communication network connecting them. A *run* r of a distributed system

[24] IBM Research Note, IBM RJ 4421, 1984. The reader might also want to consult Halpern and Fagin, "I'm OK if You're OK," *Journal of Philosophical Logic* 17, No. 4 (November 1988), pp. 329-354.

is an execution of the system, starting from time 0, when the first p_i in the system wakes up, and going on forever. A *point* is a pair (r, t), where r is a run and t a time ≥ 0. A run is characterized by associating for each processor in the system a *local history* $h(p_i, r, t)$ at each point of r. The local history of a processor p_i is empty until the time $t_0(i, r)$ that the processor first either sends or receives a message; the processor's local state at $t_0(i, r)$ is its initial state. For $t \geq t_0(i, r)$, the history $h(p_i, r, t)$ consists of p_i's initial state and the sequence of messages p_i has sent or received, in the order they were sent/received up to, but not including, point (r, t). (It is assumed that these messages are finite in number.)

> Corresponding to every distributed system, given an appropriate set of assumptions about the properties of the system and its possible interaction with its environment, there is a natural set R of all possible runs of the system. We identify a distributed system with such a set R of its possible runs. (p. 11)

One especially interesting and important family of constraints on runs is represented by deterministic *protocols.*

> A protocol is a function specifying what actions a processor takes at any given point as a function of the processor's internal state. Since a processor's internal state is determined by its history, we simply define a protocol to be a deterministic function specifying what messages the processor should send at any given instant, as a function of the processor's history. (p. 11)

In the applications we have in mind, the only actions the processors can perform are transmissions of messages to other processors. These messages might be about the environment in which the entire system is embedded, but there is no (other) interaction with the environment. The environment remains unchanged throughout a run. Indeed, the content of the messages plays no role in the specification or analysis of the distributed system. This is perfectly natural given that the purpose of the analysis is to prove properties of any of a class of distributed systems, running any of a wide variety of protocols. In any case, either the content of the messages is not intended to have any effect on the behavior of any of the processors, or what effects messages have in virtue of their contents is ignored.

Given all this, it is clear that in so far as we talk of a processor's carrying information at all, it will not be information about the wider environment that we have in mind, but information about the system and its components. Of course, Halpern and Moses talk not of carrying information, but of knowledge. For the moment, we shall follow them in this.

3.1.1 Ascribing knowledge to distributed systems

Halpern and Moses here and elsewhere stress that there is no unique "correct" answer to the question *What does it mean to say that a processor knows a fact* ϕ? Correctness is first a matter of appropriateness to the given application; beyond that, of course, relative correctness depends on the fruitfulness of the proposed analysis. They do, however, propose a general framework, *view-based* knowledge intepretations, within which to define various notions of knowledge appropriate to various concerns within the study of distributed systems.

They assume an underlying language $\mathcal{L}_0$ for expressing the *ground* facts; these are facts about the state of the system that do not "explicitly involve processors' knowledge." Examples of such formulas are "The value of register x is 0" or "Processor p_i sent the message m to processor p_j." Note that these are all facts about the distributed system. Formally, a ground fact expressed by a formula ϕ of $\mathcal{L}_0$ is modeled by a set of points $\pi(\phi) \subseteq R\,X\,[0, \infty)$. ϕ *holds* at a point (r, t) just in case $(r, t) \in \pi(\phi)$. They then extend $\mathcal{L}_0$ to a language $\mathcal{L}$ closed under Boolean connectives and under the family of operators K_i, for each p_i.[25]

> At every point each processor is assigned a *view*, and [we] say that two points are *indistinguishable* to the processor if it has the same view in both. A processor is then said to *know* a fact at a given point exactly if the fact holds at all of the points that the processor cannot distinguish from the given one. Roughly speaking, a processor knows all of the facts that (*information theoretically*) follow from its view at the current point....The approach taken to defining knowledge in view-based systems is closely related to the possible-worlds approach taken by Hintikka. For us, the "possible worlds" are the points in the system; the "agents" are the processors. A processor in one world (i.e., point) considers another world possible if it has the same view in both. (p. 13)

[25] Halpern and Moses's language also includes operators for *common knowledge* and *implicit knowledge*, where the latter is understood as the result of combining the knowledge of all the processors.

That was "roughly speaking". Formally, they posit: a *view function* v for system R that assigns a view (from some set Σ of views) to every p_i and point. "The structure of Σ is not relevant at this point." It is enough to impose the following requirement:

> whenever $h(p_i, r, t) = h(p_i, r', t')$, it must also be the case that $v(p_i, r, t) = v(p_i, r', t')$.

What justifies this requirement is that the local history of a processor encompasses all of the events, the sending and receiving of messages, that can affect the processor. Indeed, researchers often talk of the local history as capturing all the events that a processor can *observe*. Thus, it makes sense to have a processor's view at a point be a function of its local history at that point. A natural enough structure for Σ is that it is a set of *global states* of the system at a point; a global state being represented as a tuple $\langle l_1, \ldots l_i \ldots \rangle$ of the local states of the processors at that point. So a processor's view at a point in a run is, informally speaking, what of the global state of the system the processor can see (from its *point of view*, of course) at that point. What a processor knows at a point is what must be true about the system, given what it sees of the system at that point.

A *view-based knowledge interpretation* I is a triple (R, π, v). If I is a (view-based) knowledge interpretation and (r, t) a point of R (that is, $r \in R$), then (I, r, t) is a *knowledge point*. The crucial clause of the inductive definition of truth of a formula $\phi = K_i\psi$ of $\mathcal{L}$ at a knowledge point is as follows:

> $(I, r, t) \models K_i\psi$ *iff* $(I, r', t') \models \psi$ for all points satisfying $v(p_i, r, t) = v(p_i, r', t')$.

This clause "captures the fact that a processor's p_i's knowledge at a point (r, t) is completely determined by its view $v(p_i, r, t)$. The processor does not know ϕ in a given view exactly if there is [even] one point (in R) at which the processor has that same view, and ϕ does not hold."

Given that the view of a processor at a point is determined by its local history to that point and that its knowledge at that point is completely determined by its view, a processor cannot distinguish among the range of ways, the possible runs of the system, that eventuate in its local history, hence its view, at that point. That the K_i's are $S5$ operators follows immediately from the fact that the relation of a processor p_i having the same view at two points—of two points being indistinguishable to p_i—is an equivalence relation. So what a processor knows at a point is characterized in terms of what global states of the system to which it belongs are indistinguishable to the processor at that point—what states of the system are compatible with the processor's *evidence* at that point. Notice, too, that what actions a processor performs according to a protocol for the system (what messages it sends) are a function of its local history. Thus, just as for knowledge, if two points are indistinguishable to p_i, then the actions performed by p_i at these two points must be the same.

Halpern and Moses note that there is a family of view-based knowldege interpretations. Two of special interest are: the *complete history* interpretation, and the *bounded memory* or *local state* interpretation. On the complete history interpretation, $v(p_i, r, t) =_{\mathrm{df}} h(p_i, r, t)$. "That is, the processor's complete history is taken to be the view on which the processor's knowledge is based ... In a precise sense, it provides the processors with more knowledge about the ground formulas than any other interpretation could possibly do. This is one of the reasons it is particularly well suited for proving possibility and impossibility of achieving certain goals in distributed systems, and for the design and analysis of distributed protocols."

The bounded-memory interpretation allows for the fact that processors are finite state machines that might "forget" some of the facts they have previously observed. The view or *local state* of such a processor is thus some proper subfunction of its local history. "In particular, if a processor can arrive at a given state by two different message histories, then, once in that state, the processor's knowledge cannot distinguish between these two 'possible pasts'."

In all such interpretations, the equivalence relation among views is "internalist" in a certain sense. What counts for both knowledge (information) and control (action) are what possibilities for the global state are left open (dually: ruled out) by a processor's local history or local state. Again, very different global histories might all eventuate in the same local history for some processor up to some point. What that processor knows at that point is characterized by what is true of the system in all such global histories.

3.2 Situated automata

Rosenschein is interested in characterizing the internal states of computational processes or agents embedded in an environment in terms of what information those states carry about the environment.[26] In what follows, we shall assume that the agent is a robot. The environment can be in any of range R of possible states; the local state of a component of the robot is characterized by the range of such possibilities compatible with the component's being in that state. Here is a natural case that is useful to fix ideas: a component, say a particular register p_i, is connected to a sensor designed to respond differentially to one feature (or cluster of features) of its surroundings. At any time during the device's operation, p_i takes on (contains) one of a range of values—to simplify, let these be 0 or 1. The designer of the robot is to bring it about, somehow or other, that the value or *contents* of the location is 1 when and only when some environmental condition of interest obtains. Thus, only certain pairings of environment state and component state are possible; in particular, let us assume that the value of p_i is 1 iff the robot is facing a wall two feet ahead. This is handy in case one can also bring it about that p_i can be connected with control components of the robot, say those that control the movements of its wheels, in such a way that p_i's being in that state, having value 1, plays a role in causing the wheels to stop. Again, this is handy if you (the designer or the user) want the robot to stop whenever it finds a wall two feet ahead. In this way the local state of the register becomes a useful *information and control state.*

In the above example, we imagined that it was possible for the robot designer to bring it about that the local state of a particular component could be made to covary reliably with a certain fact: namely that the robot is facing a wall two feet ahead. This possiblity is conditional on various factors. Abstracting from the details, we can say that with respect to this robot and this characterization of the local state of one of its components (or some group thereof), we are interested only in that class of environments in which those conditioning factors obtain. This delimitation is captured, informally, by confining the set of possible worlds for a given application to those in which the appropriate laws hold. This is the analogue of the determination of an appropriate set R of possible runs.

Rosenschein sums this up as follows:

> Because of constraints between a process and its environment, not every state of the process-environment pair is possible in

[26]The account that follows is based on Stan Rosenschein, "Formal Theories of Knowledge in AI and Robotics", CSLI Report No. 84, 1987, and Leslie Kaelbling and Stan Rosenschein, "The Synthesis of Digital Machines with Provable Epistemic Properties," CSLI Report No. 83, 1987.

> general. A process [component] x is said to know a proposition ϕ (written $K_x\phi$) in a situation where its internal state is v if ϕ holds in all possible situations in which x is in v.[27]

Rosenschein's formal setup is an extension of that of Halpern and Moses. His language includes temporal operators as well; in Halpern and Moses, time is absorbed into the model structure. Moreover, Rosenschein separates out a domain of *locations* and a domain of values, in the style of the semantics of programming languages. We ignore the complications.

Rosenschein's approach, then, is analogous to the case in which memory is bounded. One defines an equivalence relation (of indistinguishability) between local states of a processor and then defines what the processor knows in a given state as what is true whenever the processor is in a state indistinguishable from (of the same kind as) the given state.

4 Comparative Axiomatics

4.1 The Halpern-Moses, Rosenschein approach and the axioms

It is quite clear which of the axioms (1) through (3) and which rules of inference are validated by the Halpern-Moses and Rosenschein conceptions. All of them; indeed, all of them, and then some. In particular, axiom scheme (4) is also validated.[28]

The validation is quite direct. The crucial elements are that the approach is carried out within standard modal model theory and that the accessibility relation is taken to be that of indistinguishability between possible runs. Given that the view of a processor at a point is determined by its local history to that point and that its knowledge at that point is completely determined by its view, a processor cannot distinguish among the possible runs of the system, that eventuate in its local history, hence its view, at that point. That the K_i's are $S5$ operators follows immediately.

This is one of the few cases in intensional logic in which the natural order of analysis leads from constraints on a binary relation (indistinguishability,

[27] "The Synthesis of Digital Machines with Provable Epistemic Properties", p.84

[28] As noted, in Halpern-Moses, the axiomatization covers other notions as well—those of *implicit* and *common knowledge*—and they show that $S5$ is not a complete axiomatization of those notions.

given the state of a component device) defined on an independently well-motivated and easily grasped set of items (possible runs of a system subject to certain constraints or specifications) to the axiomatization, rather than vice-versa.[29] In this respect, the earlier work of Hintikka and others on epistemic logics is more typical. Such work starts with the attempt to figure out what principles govern a certain notion and then moves to models.

4.2 The Israel-Perry approach and the axioms

What of the Israel-Perry approach? We shall be dealing at first only with pure information. We want to check which of the principles (1) through (4) hold of this notion. To simplify the comparison, we shall assume a language adequate for classical sentential logic. We shall ignore the fact that the simplest form of pure information content is:

$(\models \exists \mathrm{x}(\ll R, a, \mathrm{x}; 1 \gg \wedge \ll F, \mathrm{x}; 1 \gg))$

Instead we shall associate with the sentence letters of the language, propositions of the form $(\models T_\sigma)$, where σ is a basic state of affairs with positive polarity. Every basic state of affairs has a dual. As noted earlier, we are foregoing aspects of partiality crucial to situation theory. We assume that no issue is unresolved; that is, for every basic state of affairs σ, either it or its dual is factual. A sentence letter gets assigned false just in case the dual of the state of affairs of the associated proposition is factual and this is the condition, $(\neg \models T_\sigma)$, for the truth of the negation of the sentence.

This fixes only one valuation, the one that corresponds to the way the world actually is. This is an account of an interpreted language, not of the possible valuations of an uninterpreted one. The end result is to allow us to validate all the classical logical truths of our language as true (simpliciter). Our account makes no room for alternative possible worlds; the locus of possibility, to repeat, is in the relation between the one actual world and various ways (parts of) it might be. We take intensional (modal) facts to be facts about the actual world and the objects in it, not nonintensional (nonmodal) facts about other worlds. We are not trying to show any equivalence between the situation-theoretic structures and Kripke-style structures for sentential modal logic. We are describing an interpreted intensional (modal) language with an eye to seeing which principles, as expressed in that language, we accept.[30]

[29] Temporal logic also provides natural model structures for modal languages.

[30] On the other hand, we can indeed model appropriate situation theoretic structures in set theory and obtain analogues of modal models for our language. For more on this, see

Still pursuing the groundwork necessary for the comparison, the compound infons (states of affairs) that enter into our analysis don't enter into our language as propositions. Rather, they are implicit in our intensional operators. In the Halpern-Moses setup, there is a family of indexed operators, the indices being associated with processors—the components of the distributed system. One way of presenting Rosenschein's account would have a doubly indexed set of operators—one component for components (devices, processors) as in Halpern-Moses, the other for times. Modal operators can be introduced in a presentation of the Israel-Perry approach, by a family of multiply indexed operators. One would index by indicating fact (which includes a temporal constituent) and by constraint. So one might see operators that look like this: $\Box_{f,C}$. We assume that the fact is always *meaningful* for the given constraint; that is, the fact is of the first type of the associated pure constraint.[31] As usual, if we keep that to which we are relativizing fixed, we would drop subscripts. Again, we do not favor such a language; we would expect rather to express the relevant principles in a (many-sorted first-order) language for situation theory.

Now on to the axioms and rules. Skipping ahead to the easy case, it is clear both for pure and incremental information that if some fact carries the information that P, relative to some constraint, and perhaps some connecting facts, then it is the case that P. So axiom scheme (2) is validated.

What of scheme (1), distributivity over the conditional? If we were interested in the nonrelative notions of the information a fact carries relative to some constraint or other or to the notion of the total information carried by a fact relative to all actual constraints, perhaps there would be no problem with (1). But neither of those is the relation we have been studying. If an indicating fact carries the information, relative to some constraint C, that $(P \rightarrow Q)[\equiv (\neg P \vee Q)]$, then if that fact carries the information that P, relative to C, does it carry the information that Q, relative to that same constraint?

This question suggests another. Can a fact carry the information that $(P \vee Q)$ without carrying the information that P or the information that Q? Implicitly, we have been supposing that it can, in so far as we have been supposing that the operator associated with *carries the information that P* is

---

Barwise and Etchemendy, *The Liar*.

[31] For incremental information, we would index by indicating fact, constraint and connecting fact. So the operators might look like this: $\Box_{f,C,c}$. Given the unhappiness of this notation, the reader will be pleased to know that we will not be exploring the logic of incremental information in this essay.

a *necessity-style operator*. It can be necessary that $(P \vee Q)$ without it being necessary either that P or that Q. But this does not hold for possibility. How is it with respect to *σ carries the information that*? We have just noted that this acts like $\Box$, and unlike $\Diamond$, in being truth-involving. We shall return to this question in §4.2.1; but, for now, we shall continue to assume that our operator is akin to the $\Box$ with respect to its interaction with $\vee$. To return to the main theme, it does seem reasonable to assume that schema (1) is validated.

How about axiom scheme (3)? It is hard to get one's head clear about this. Here's what the principle says, in a given instance. If the x-ray's being such and such carries the information that there is a dog with a broken leg, relative to C, then, given C, the x-ray's being such and such carries the information that the x-ray's being such and such carries the information that there is a dog with a broken leg, given C. Does any iteration of this operator make sense?

Even if one has in mind the relative notion of the information a fact carries, given a certain constraint, one might very well reason as follows. We are given that a certain device, within a larger system of devices (or a world) subject to certain constraints, is in a certain state at a certain time. We can now say that there are three kinds of truths:

(i) Those that are true, quite independent of how the system works or what state the device is in.

(ii) Those whose truth depends in some way on how the system works, but are independent of what state the particular device is in.

(iii) Those whose truth depends both on how the system works and on what state the device is in.

Of the first, some are merely contingent truths about objects quite remote from the system or about features of the system or its components that don't enter into the relevant constraints (as e.g., the color of the x-ray technician's uniform). But some of these truths are necessary truths. About these, one might accept the principle that every fact carries the information that p, where p is a necessary truth. This would involve accepting the following. For any appropriate fact f and constraint C:

- $\Box p \rightarrow \Box_{f,C}\, p$

where the unsubscripted $\Box$ expresses plain old (metaphysical) necessity.[32]

The proposition that if the system or the world is governed by a certain set of constraints, then the fact that the x-ray has the pattern Φ carries the information that there is a dog with a broken leg *is* a necessary, nonlogical truth. The fact itself is just a contingency; the proposition that the x-ray has the Φ pattern is a contingent one. That the constraint in question is actual is itself a contingency; the proposition that the one type of situation involves the other is a contingent truth, if true. But, the proposition that, if the relevant constraint is actual, then the x-ray's being Φ-ish carries the information that there is a dog with a broken leg is, if true, a metaphysically necessary truth. So, where f is the fact about the x-ray, C the relevant constraint(s), and p is the sentence to the effect that there is a dog with a broken leg, we would have:

- $\Box(\Box_{f,C}\, p)$
- $\Box_{f,C}(\Box_{f,C}\, p)$

Of course, it is also true that one would have this for any appropriate fact-constraint pair. That is, let $f' \neq f$, $C' \neq C$, and let it be the case that f carries the information that q. Then

- $\Box_{f',C'}\, q$
- $\Box_{f,C}(\Box_{f',C'}\, q)$
- $\Box_{f',C'}(\Box_{f,C}\, p)$

It is reasonable to accept a similar thesis with respect to truths of type (ii), limited this time to those fact-constraint pairs whose first members are facts involving possible states of the system in question and whose second members are the system constraints.[33] Let p be a proposition such that, given the way the system works, that is given constraint(s) C, then no matter what state the device is in, it is the case that p. Thus $\Box_{f,C}\, p$ and $\Box_{f,C}\Box_{f,C}\, p$.

[32] If it is a necessary truth that p, then it is a necessary truth that it is. So we also have $\Box_{f,C}\Box\, p$. In any event notice that neither $(\Box_{f,C}\, p \rightarrow \Box(\Box_{f,C}\, p))$ nor $(\Box_{f,C}\, p \rightarrow \Box_{f,C}(\Box\, p))$ is the principle we are after.

[33] On the Halpern-Moses approach, this limitation is vacuous. The system is the world: more properly, each possible world is a possible run of the system, subject to (all) the constraints expressed in the specification of the system. The same is true for Rosenschein. For him, the world is the system; the constraints in question, all the constraints. We return to this theme in the conclusion.

It is hard to know what to make of this line of reasoning. One thing is clear, though. It is arguably more plausible if one has in mind the notion of the total (pure) information carried by a fact. Further, it is not at all clear there is any other reason for accepting schema (3) as valid for the relational notion we have in mind. We reject (3), though tentatively.

At best the same fate awaits schema (4). Indeed, a worse fate awaits it. We reject (4). We shall not argue directly for this. Rather we shall simply ask the reader to ponder the following candidate principles:

- 4′. $(\neg\Box_{\mathfrak{f},\mathcal{C}}\neg\Box_{\mathfrak{f},\mathcal{C}}\, p \rightarrow \Box_{\mathfrak{f},\mathcal{C}}\, p)$
- B. $(\neg\Box_{\mathfrak{f},\mathcal{C}}\neg\Box_{\mathfrak{f},\mathcal{C}}\, p \rightarrow p)$

(4′) is merely the (slightly more illuminating) contrapositive of (4). (B) is the so-called symmetry schema; given the principles we have accepted so far (and including necessitation), rejecting (B) means rejecting (4). There is adequate reason for rejecting (B)—unless, perhaps, one has in mind the notion of the total information carried by a state.

What of necessitation? Is it a sound rule of inference—sound, that is, with respect to truth?[34] Very much the same line of reasoning sketched above with respect to (3) can be brought to bear on this question. Again, there is a plausible story according to which, where p is a necessary truth—logical or nonlogical—then, for any fact and constraint, $\Box_{\mathfrak{f},\mathcal{C}}\, p$ is also a necessary truth. With the same limitations in force as in the case of (3), where p is any necessary truth of the second sort, so too is $\Box_{\mathfrak{f},\mathcal{C}}\, p$.

Unrestricted necessitation is certainly not sound. Indeed a similar restriction on necessitation is often in force in applied epistemic logics. Let a be some agent of interest. Let p be some truth known to the theorist, but not known by the theorist to be known to a. The theorist may even know that p is unknown to a. The theorist may want to include p as an axiom in his theory of a and a's knowledge. Then he won't accept the rule: if $\vdash p$, then $\vdash K_a\, p$. Rather, it is sound if restricted to logical truths—including truths of the relevant (pure) logic of knowledge.

If one is thinking in terms of standard modal systems, what we are left with is a choice. On the one hand, if we accept the line of reasoning sketched above, we arrive at $S4$. On the other, if we reject it totally, we arrive at the modal system M (or T), but without necessitation. If we reject it less totally, we do arrive at T (or M). But there is a further consideration.

[34] Of course, this is not the usual question, but we are analyzing an interpreted language, and are interested in truth *simpliciter*, not truth-in-a-model.

4.2.1 A can of worms

M, $S4$, and $S5$ are *classical* modal systems; they yield as a derived rule, *RE*: if $\vdash (P \leftrightarrow Q)$, then $\vdash (\Box P \leftrightarrow \Box Q)$. This rule validates replacement of provable logical equivalents, and modal contexts do support substitution of logical and or necessary equivalents. This does not seem to be true for reports of cognitive attitudes, such as *x knows that.* Indeed this is widely perceived as a major problem with using standard (normal or classical) modal logics for axiomatizing or characterizing cognitive notions. How about information contexts? On none of the three analyses presented here, are these cognitive contexts; they involve no essential reference to the mental states of agents.

Important evidence in regard to the transparency to substitution by logically equivalent sentences would be judgments as to the validity of arguments such as the following:

> (a) The tree's having 100 rings carries the information that the tree is 100 years old.
>
> (b) So, the tree's having 100 rings carries the information that the tree is 100 years old and that either Palo Alto is the capital of California or it is not the case that Palo Alto is the capital of California.[35]

Though we believe that such arguments are invalid, we do not base any claim on our intuitions as to how such judgments would actually go. Indeed, we have made no direct claims about the semantics of such natural language contexts. We have sketched an account that allows us both to have propositions as the range of the *carries information* relation and to distinguish propositions that must be true or false together. This account allows us to hold that the relation distinguishes between logically or necessarily equivalent propositions just as it does between materially equivalent propositions. We haven chosen to exploit that opportunity. Again, one can imagine a line of reasoning that supports substitutivity of logical equivalents; and again, that line will be the more plausible, the more one really has in mind a different notion: that of the total information carried by a fact.

If we were to mount an argument that the context does not support substitution of logical equivalents, it might rest on the following principles: If σ carries the information (relative to some fixed constraint) that P and

[35]The reader may replace "carries the information" with "really indicates".

Q, it carries the information that P. Moreover, and more controversially, if σ carries the information that P or Q, then it must either carry the information that P or carry the information that Q. So in this respect, the logic of the notion is akin to that of possibility, rather than necessity. But earlier, in §4.2, we decided to assume just the opposite.

Before addressing this conflict, let us apply these principles to the above argument. If we accepted (b) above, we would have it that the tree's having 100 rings carries the information that either Palo is the capital of California or it is not. By the second and more controversial principle, it follows that either the tree's having 100 rings carries the information that Palo Alto is the capital of the state or it carries the information that it is not the capital of the state. Neither is true. Relative to the implied constraint, the number of rings carries no information about the government of California.

The problem is that the logic of σ *carries the information that* seems to be mixed. We have already seen that it is akin to a necessity operator in being truth-involving. Moreover, if σ both carries the information that P and that Q, relative to a single fixed constraint, then relative to that same constraint, it carries the information that P and Q. In this respect also, the operator acts like a necessity, not a possibility, operator. On the other hand, there is some intuitive backing for the principle that σ carries the information that $PorQ$ *iff* it carries the information that P or it carries the information that Q.

In this regard, it might be useful to compare the following operator (scheme): σ *brings it about that*—where σ is particular act, consisting of something's doing something at some time and place. This also involves a notion of causal or lawlike connection. Surely if σ brings it about that P, then P. It seems also that, if σ brings it about that P and brings it about that Q, then it brings it about that P and Q. Now, what of the following: it cannot be the case that σ brings it about that P or Q unless it either brings it about that P or brings it about that Q? Intuitions waver on this point.

One major reason for our not opting to argue for the nonclassical nature of the operator associated with "σ carries the information that..." is that we wanted to avoid opening this can of worms. In what follows, we shall continue to assume that the associated operator is a necessity-style operator, at least with respect to its interaction with "and" ($\wedge$) and "or" ($\vee$). We hope to address these issues in a later paper.

4.3 The status of constraints

Though carrying information is, as we have noted, truth-involving, it cannot be completely characterized in terms of truth. One needs as well some notion of a lawlike connection, a constraint, between information-carrying states or signals and information content. And in all three accounts, some such notion is introduced, but in quite different ways.

There is no representation within the modal frameworks developed and deployed by Halpern-Moses and Rosenschein of the constraints obeyed by the systems they analyze. The notion of a component's carrying information (*knowing that*), like that of necessity (in a model), is defined in terms of truth (in a model). The logical constraints are built into the logic, that is, into the inductive definition of truth in a model. The relativization to the relevant nonlogical, or *design*, constraints is captured only informally, by an understanding of the appropriate domain of the model structure (the set of possible worlds) as characterized by the specification of a given system—that is, by identifying the set of possible worlds with the set of possible runs (points) of the system so specified.[36] Such a specification might be given in some formalism, but it would play no direct role in the logic of information flow in distributed systems.

The situation is more complicated in the case of the work on situated automata. There, too, the relativization to particular constraints operative in the target environments is realized by way of a delimitation of the set of possible worlds. Of course, in the extreme case, this might be thought of, for example, as the set of all physically possible worlds. But this characterization is unlikely to be of much use. In realistic applications, the designer will neither know nor need to know all physical laws. Rather, he will have in mind a range of environments and tasks that, together with some specification of the general capacities of the device components, will suggest a set of relevant constraints. The designer is to express these constraints in a formal theory of the environment and of the desired interactions between the environment and the device. This theory then is a crucial input to a semiautomated design and compilation process.[37]

In the situation theoretic framework, on the other hand, constraints

[36] Halpern-Moses aren't interested in the set of all possible runs (or points) of all possible distributed systems—assuming such a notion made sense.

[37] For more on this, see "The Synthesis of Digital Machines with Provable Epistemic Properties" and Rosenschein's "Synthesizing information-tracking automata from environment descriptions," to appear in the *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, 1989.

(relations between types of situation) are "first-class citizens" within the theory—any appropriate situation theoretic structure for a theory of information will include a subdomain of types of situations and constraints between them, the latter simply being a special kind of state of affairs. This allows us to treat the relativity to constraints that we find so important explicitly.

The comparison we have been running through should not lead one to believe that Kripke-style model structures for modal languages are appropriate structures for the theory of information presented in section 2. The present comparison has only to do with axiomatics.

5 Conclusion

In all the cases where the logic of the relational notion carrying information differs from $S5$, the key point is precisely the relativity to constraints.

Indeed, we suggest that Halpern-Moses and Rosenschein should be thought of as axiomatizing the following notion: *the total pure information carried by a fact or by the state of an information-bearing object.* Fix all the constraints and some state of some entity: how might a world (system) be that honors those constraints? What possibilities are ruled out by the state of the entity; how does the state of the entity discriminate amongst all the possibilities? Posing the question this way leads naturally to postulating an equivalence relation among states.[38]

Let us return to the point, made earlier, about the difference between $S5$ and the normal modal logics weaker than $S5$. In $S5$, necessity is a property of propositions; analogously, in an $S5$ logic of information, the total information a state carries is a property of that state.[39] It should come as no surprise, if the notion one is analyzing is an ineliminably relational one, that the appropriate modal structures are going to be ineliminably relational. That is, they will not be $S5$ structures.

So far, nothing has been made of the distinction between pure and incremental information, or more properly between the relations of a fact's

[38] If no further, connecting facts are fixed, then the entity's being in one or another state cannot impose any conditions on any specific, remote objects. Of course, the constraints can impose general conditions on the world (environment).

[39] It seems to be an intrinsic property of the state; of course, it really isn't. In terms of model structures, what is required is that the possible worlds (possible runs) of a structure are worlds in which all the actual constraints (resp., all the design constraints or specifications that determine the system) are operative.

carrying the pure information that P relative to a constraint and the relation of a fact's carrying the information that P, relative to a constraint and a set of connecting facts.[40]

The logic of incremental information has yet to be explored.

6 Bibliographical Note

In section 2, we presented a brief sketch of parts of situation theory. For more, and also differing, detail the reader should consult Jon Barwise, "Notes on a model for a theory of situations" (1987, unpublished ms.), "A type-free theory of types and propositions" (1988, unpublished ms.), "Situations, Facts, and True Propositions," in *The Situation in Logic*, CSLI, Lecture Notes, No. 17, 1989, pp. 221-254, "Branch points in situation theory", in *ibid*, pp. 255-276; Keith Devlin, *Logic and Information* (1989, forthcoming); Tim Fernando, "A logical space of abstract situations" (1989, to appear in *Proceedings of the First STASS Conference*); and Dag Westerstähl, "Parametric Types and Propositions in Situation Theory" (1989), Report No. 89-7, of the Department of Philosophy, Göteborg University.

[40]It is not at all clear that there is a useful notion of the total incremental information carried by a fact.