

PDF version of the entry  
Logic and Information  
<https://plato.stanford.edu/archives/win2016/entries/logic-information/>  
from the WINTER 2016 EDITION of the

## STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta	Uri Nodelman	Colin Allen	R. Lanier Anderson
Principal Editor	Senior Editor	Associate Editor	Faculty Sponsor

Editorial Board  
<https://plato.stanford.edu/board.html>

Library of Congress Catalog Data  
ISSN: 1095-5054

**Notice:** This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

*Stanford Encyclopedia of Philosophy*  
Copyright © 2016 by the publisher  
The Metaphysics Research Lab  
Center for the Study of Language and Information  
Stanford University, Stanford, CA 94305

Logic and Information  
Copyright © 2016 by the authors  
Maricarmen Martinez and Sebastian Sequoiah-Grayson

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

## Logic and Information

*First published Mon Feb 3, 2014*

The explicit inclusion of the notion of *information* as object of logical study is a rather recent development, when one compares with notions such as proof, truth, consequence, or algorithm, which are also central to logical practice. It was by the beginning of the present century that a sizable body of existing technical and philosophical work (with precursors that can be traced back to the 1930s) coalesced into the new emerging fields of logic and philosophy of information. This entry is mainly about the logic. It surveys major logical approaches to the notion of information, according to three complementary stances: information-as-range, information-as-correlation, and information-as-code.

The core intuition leading the INFORMATION-AS-RANGE stance is that an informational state may be characterised by the range of possibilities or configurations that are compatible with the information available at the moment. Acquiring new information corresponds to shrinking that range, thus reducing uncertainty about the actual configuration of affairs. With this understanding, the setting of possible-world semantics for epistemic modal logics proves adequate for the study of various semantic aspects of information. A prominent theme here is *information update*, which many times occurs in *social settings* due to *interaction between agents* according to a variety of *epistemic actions*.

The INFORMATION-AS-CORRELATION stance focuses on information flow as it is licensed within structured systems formed by parts that are systematically correlated. For example: the number of rings of a tree trunk can give you information about the time when the tree was born in virtue of certain regularities of nature that ‘connect’ the past and present of trees.

The central themes of this stance include the *aboutness*, *situatedness*, and *accessibility of information in structured information environments*.

The key concern of the third stance, INFORMATION-AS-CODE, is the *syntax-like structure* of information pieces (their *encoding*) and the *inference and computation processes* that are licensed by virtue (among other things) of that structure. A most natural logical setting to study these informational aspects is proof theory. Some *substructural logics*, in particular relevant logic and some linear logics, have been given in recent times interpretations as logics that capture relevant aspects of the information-as-code stance.

The three stances are by no means incompatible, but they are not necessarily reducible to each other either. This will be expanded on later in the entry, and some further topics of research will be illustrated, but for a preview of how the three stances can live together, take the case of a structured information system composed of several parts. First, the correlations between the parts naturally allow for ‘information flow’ in the sense of the information-as-correlation stance. Second, they also give raise to local ranges of possibilities, since the local information available at one part will be compatible with just a certain range of global states of the system. Third, the combinatorial, syntax-like, proof-theoretical aspects of information can be brought to this setting in various ways. One of them is treating the correlational flow of information as a sort of combinatorial system by which local information states are combined in syntactic-like ways fitting a particular interpretation of relevance logic. Or one could explicitly add code-structure to the modelling, for example by assigning local deductive calculi to either the components or the local states of the system.

- 1. Semantic Information as Range
  - 1.1 Epistemic logic

- 1.2 Dynamic epistemic logic, information change
- 1.3 Quantitative Approaches
- 2. Information as Correlation: Situation Theory
  - 2.1 Situations and Supporting Information
  - 2.2 Information flow and constraints
  - 2.3 Distributed information systems and channel theory
- 3. Information as Code
  - 3.1 Categorical Information Theory
  - 3.2 Substructural logics and information flow
  - 3.3 Related Approaches
- 4. Connections Between the Approaches
  - 4.1 Ranges and correlations
  - 4.2 Code and correlations
  - 4.3 Code and ranges
- 5. Special topics
  - 5.1 Information Structures and Equivalence
  - 5.2 Negative information
- 6. Conclusion
- Bibliography
- Academic Tools
- Other Internet Resources
- Related Entries

---

## 1. Semantic Information as Range

The understanding of semantic information as range has its origins in Bar-Hillel and Carnap's theory of semantic information, Bar-Hillel and Carnap (1952). It is here that the *inverse range principle* is given its first articulation with regard to the informational content of a proposition. The inverse range principle states that the more information carried by a

proposition, the less likely it is that the proposition is true. Similarly, the more likely the truth of a proposition, the less information it carries.

The likelihood of the truth of a proposition connects with semantic information as range via a possible worlds semantics. For any contingent proposition, it will be supported by some possibilities (those where it is true) and not supported by others (those where it is false). Hence a proposition will be supported by a range of possibilities, an “information range”. Now suppose that there is a probability distribution across the space of possibilities, and for the sake of simplicity suppose that the distribution is uniform. In this case, the more worlds that support a proposition, the likelier the propositions truth, and, via the inverse relationship principle, the less information it carries. Although information as range has its origins in quantitative information theory, its role in contemporary qualitative logics of information cannot be overstated.

Consider the following example due to Johan van Benthem (2011). A waiter in a cafe receives the order for your table—an espresso and a soda. When the waiter arrives at your table, he asks “For whom is the soda?”. After your telling him that the soda is for you and his giving you your soda, the waiter does not need to ask about the espresso, he can just give it to your cafe-partner. This is because the information gained by the waiter from your telling him that you ordered the soda allows him to eliminate certain open possibilities from the total range of possibilities such that only one is left—your friend ordered the espresso.

Logics of information distinguish regularly between *hard information* and *soft information*. Hard information is *factive*, and unrevisable. Hard information is often taken to be the object of *knowledge*. In contrast to hard information, soft information is *non-necessarily-factive*, hence revisable in the presence of new information. Soft information, in virtue of its revisability, corresponds very closely to the information that falls inside

the scope of belief as opposed to knowledge. The terms knowledge and belief are conventional, but on the context of information flow, the hard/soft information reading is convenient on account of it bringing the informational phenomena to the foreground. Although both hard and soft information are important for our epistemic and doxastic success, in this section we will concentrate mainly on logics of hard information flow.

In section 1.1 we will see both how it is that classic epistemic logics exemplify the flow of hard information within the information as range framework. In section 1.2 we will extend our exposition from logics of hard information-gain to logics of the actions that facilitate the gain of such hard information, dynamic epistemic logics. At the end of Section 1.2, we will expound the important phenomenon of *private information*, before examining how it is that information as range is captured in various *quantitative frameworks*.

## 1.1 Epistemic logic

In this section we will explore how it is that the elimination of possibilities corresponding to information-gain is the starting point for research on logics of knowledge and belief that fall under the heading of *epistemic logics*. We will begin with classic single-agent epistemic logic, before exploring multi-agent epistemic logics. In both cases, since we will be concentrating on logics of knowledge as opposed to logics of belief, the information-gained will be hard information.

Consider the waiter example in more detail. Before receiving the hard information that the soda is for you, the waiter's knowledge-base is modelled by a pair of worlds (hereafter *information states*)  $x$  and  $y$  such that in  $x$  you ordered the soda and your friend the espresso, and in  $y$  you ordered the espresso and your friend the soda. After receiving the hard information that the soda is for you,  $y$  is eliminated from the waiter's

knowledge-base, leaving only  $x$ . As such, the reduction of the range of possibilities corresponds to an information-gain for the waiter.

Although epistemic logic was conceived traditionally as a logic of *knowledge*, one natural interpretation of the formalism is in terms of a logic modelling the handling of hard semantic information, an approach given contemporary expression in Hintikka (2007). That this corresponds to an understanding of semantic information as range is most easily appreciated by noting that an agent's knowledge-base is represented by an accessibility relation ranging across information-states (possible worlds). Consider the truth condition for *agent  $\alpha$  knows that  $\phi$* , written  $K_\alpha\phi$ :

$$(1) \ x \models K_\alpha\phi \text{ iff for all } y \text{ s.t. (such that) } R_\alpha xy, y \models \phi$$

The accessibility relation  $R$  is an equivalence relation connecting  $x$  to all information states  $y$  such that  $y$  is indistinguishable from  $x$ , *given  $\alpha$ 's hard information at that state  $x$* . So, if  $x$  was the waiter's information state before being informed that you ordered the soda,  $y$  would have included the information that you ordered the espresso, as each option was as good as the other until the waiter was informed otherwise. There is an implicit assumption at work here—that some state  $z$  say, where you ordered both the soda and the espresso, is not in the waiter's information-range. That is, the waiter knows that  $z$  is not a possibility. Once informed however, the information states supporting your ordering the espresso are eliminated from the range of information corresponding to the waiter's knowledge.

Basic modal logic extends propositional formulas with modal operators such as  $K$ , where  $\mathbf{K}$  is the set of all Kripke Models then we have the following axioms:

$$(A1) \ \mathbf{K} \models K_\alpha\phi \wedge K_\alpha(\phi \rightarrow \psi) \rightarrow K_\alpha\psi$$

$$(A2) \ \mathbf{K} \models \phi \Rightarrow \models K_\alpha\phi$$

(A1) states that hard information is closed under (known) implications. Since the first conjunct states that all states accessible by  $\alpha$  are  $\phi$  states,  $\alpha$  possesses the hard information that  $\phi$ , hence  $\alpha$  also possesses the hard information that  $\psi$ . (A2) states that if  $\phi$  holds in the set of all models, then  $\alpha$  possesses the hard information that  $\phi$ . In other words, (A2) states that all tautologies are true and unrevisable, and (A1) states that  $\alpha$  knows the logical consequences of all propositions that  $\alpha$  knows (be they tautologies or otherwise). That is, the axioms state that the agent is logical omniscient, or an *ideal reasoner*, a property of agents that we will return to in detail in the sections below.<sup>[1]</sup>

The formalism above is basic single-agent epistemic logic. But reasoning and information flow are very often *multi-agent affairs*. Consider again the waiter example above. Importantly, the waiter is only able to execute the relevant reasoning procedure corresponding to a restriction of the range of information states *on account of* your announcement to him with regard to the espresso. That is, it was the verbal interaction between several agents that facilitated the information flow that enabled the logical reasoning to be undertaken.

It is at this point that multi-agent epistemic logic raises new questions regarding the information in a group. “Everybody in  $G$  possesses the hard information that  $\phi$ ” (where  $G$  is any group of agents from a finite set of agents  $G^*$ ) written as  $E_G\phi$ .  $E_G$  is defined for each  $G \subseteq G^*$  in the following manner:

$$(2) \ E_G\phi = \bigwedge_{\alpha \in G} K_\alpha\phi$$

This form of group knowledge is importantly different from *common knowledge* (Lewis 1969; Fagin et al. 1995). Common knowledge is the condition of the group where *everybody knows that everybody knows that*

*everybody knows ... that*  $\phi$ . In other words, common knowledge concerns the hard information that each agent in the group possesses about the hard information possessed by the other members of the group. That everybody in  $G$  possesses the hard information that  $\phi$  does not imply that  $\phi$  is common knowledge. With group knowledge each agent in the group may possess the same hard information (hence achieving group knowledge) without necessarily possessing hard information about the hard information possessed by the other agents in the group. As noted by van Ditmarsh, van der Hoek, and Kooi (2008: 30), “the number of iterations of the  $E$ -operator makes a real difference in practice”.  $C_G\phi$ —the common knowledge *that*  $\phi$  for members of  $G$ , is defined as follows:

$$(3) \quad C_G\phi = \bigwedge_{n=0}^{\infty} E_G^n\phi$$

To appreciate the difference between  $E$  and  $C$ , consider the following “spy example” (originally Barwise (1988) with the envelope details due to Johan van Benthem.

There are a group of competing spies at a formal dinner. All of them are tasked with the mission of acquiring some secret information from inside the restaurant. Furthermore, it is common knowledge amongst them that they want the information. Given this much, compare the following:

- Each spy knows that the information is in an envelope on one of the other tables, but they don't know that the other spies know this (i.e., it is not common knowledge).
- It is common knowledge amongst the spies that the information is in the envelope.

Very obviously, the two scenarios will elicit very different types of behaviour from the spies. The first would be relatively subtle, the latter

dramatically less so. See Vanderschraaf and Sillari (2009) for further details.

## 1.2 Dynamic epistemic logic, information change

As noted above, the waiter example from the beginning of this section is as much about information-gain via announcements as it is about information structures. In this section, we will outline logics of actions that result in information increase.

Hard information flow can be facilitated by more than one epistemic action. Two canonical examples are *announcements* and *observations*. When “announcement” is restricted to *true and public announcement*, its result on the receiving agent's knowledge-base is similar to that of an observation (on the assumption that the agent believes the content of the announcement). The public announcement that  $\phi$  will restrict the model of the agent's knowledge-base to the information states where  $\phi$  is true, hence “announce  $\phi$ ” is an *epistemic state transformer* in the sense that it transforms the epistemic states of the agents in the group, (van Ditmarsh, van der Hoek, and Kooi 2008: 74).<sup>[2]</sup>

Dynamic epistemic logics extend the language of epistemic logics with dynamic operators. In particular, *public announcement logic* (PAL) extends the language of epistemic logics with the dynamic announcement operator  $[\phi]$ , where  $[\phi]\psi$  is read “after announcement  $\phi$ , it is the case that  $\psi$ ”. The key *reduction axioms* of PAL are as follows:

- (RA1)  $[\phi]p$  iff  $\phi \rightarrow p$  (where  $p$  is atomic)
- (RA2)  $[\phi]\neg\phi$  iff  $\phi \rightarrow \neg[\phi]\psi$
- (RA3)  $[\phi](\psi \wedge \chi)$  iff  $[\phi]\psi \wedge [\phi]\chi$
- (RA4)  $[\phi][\psi]\chi$  iff  $[\phi \wedge [\phi]\psi]\chi$
- (RA5)  $[\phi]K_\alpha\psi$  iff  $\phi \rightarrow K_\alpha(\phi \rightarrow [\phi]\psi)$

RA1–RA5 capture the properties of the announcement operator by connecting what is true before the announcement with what is true after the announcement. The axioms are named ‘reduction’ axioms because the left-to-right hand direction reduces either the number of announcement operators or the complexity of the formulas within their scope. For an in depth discussion see Pacuit (2011). RA1 states that announcements are truthful. RA5 specifies the epistemic-state-transforming properties of the announcement operator. It states that  $\alpha$  knows that  $\psi$  after the announcement that  $\phi$  iff  $\phi$  implies that  $\alpha$  knows that  $\psi$  will be true after  $\phi$  is announced in all  $\phi$ -states. The “after  $\phi$  is announced” condition is there to account for the fact that  $\psi$  might change its truth-value after the announcement. The interaction between the dynamic announcement operator and the knowledge operator is described completely by RA5 (see van Benthem, van Eijck, and Kooi 2006).

Just as adding the *common knowledge* operator  $C$  to multi-agent epistemic logic extends the expressive capabilities of multi-agent epistemic logic, adding  $C$  to PAL results in the more expressive *public announcement logic with common knowledge*, (PAC). The exact relationship between public announcements and common knowledge is captured by the *announcement and common knowledge rule* of the logic PAC as the following:

(4) From  $\chi \rightarrow [\phi]\psi$  and  $(\chi \wedge \phi) \rightarrow E_G\chi$ , infer  $\chi \rightarrow [\phi]C_G\psi$ .

Again, PAC is the dynamic logic of hard information. The epistemic logics dealing with *soft information* fall within the scope of *belief revision theory* (van Benthem 2004; Segerberg 1998). Variants of PAL that model soft information augment their models with plausibility-orderings on information-states (Baltag and Smets 2008). These orderings are known as *preferential models* in non-monotonic logic and belief-revision theory. The logics can be made dynamic in virtue of the orderings changing in the face of new information (which is the mark of soft information as opposed to

hard information). Such plausibility-orderings may be modelled qualitatively via partial orders etc., or modelled quantitatively via probability-measures. Such quantitative measures provide a connection to a broader family of quantitative approaches to semantic information that we will examine.

**Private information.** Private information is an equally important aspect of our social interaction. Consider scenarios where the announcing agent is aware of the private communication whilst other members of the group are not, such as emails in Bcc. Consider also scenarios where the sending agent is *not* aware of the private communication, such as a surveillance operation. The system of *dynamic epistemic logic* (DEL) models events that turn on private (and public) information by modelling the agents' information concerning the events *taking place* in a given communicative scenario (see Baltag et al. 2008; van Ditmarsh et al. 2008; and Pacuit 2011).

Importantly, modal information theory approach to multi-agent information flow is the subject of a great amount of research. The semantics is not always carried out in relational terms (i.e., with Kripke Frames) but is done often algebraically (see Blackburn et al. 2001 for details of the algebraic approach to modal logic). For more details on algebraic as well as type-theoretic approaches, see the subsection on algebraic and other approaches to modal information theory in the supplementary document Abstract Approaches to Information Structure.

### 1.3 Quantitative Approaches

Quantitative approaches to *information as range* also have their origins in the inverse relationship principle. To restate—the motivation being that the less likely the truth of a proposition as expressed in a logical language with respect to a particular domain, the greater the amount of information

encoded by the relevant formula. This is in contrast to the information measures in the *mathematical theory of communication* (Shannon 1950) where such measures are gotten via an inverse relationship on the expectation of the receiver  $R$  of the receipt of a signal from some source  $S$ .

Another important aspect of the theory of classical semantic information, is that it is an entirely *static* theory—it is concerned with the semantic informational content and measure of particular formulas, and not with information *flow* in any way at all.

The formal details of semantic information theory turn on the probability calculus. These details may be left aside here, as the obvious conceptual point is that logical truths have a truth-likelihood of 1, and therefore an information measure of 0. Bar-Hillel and Carnap did not take this to mean that logical truths, or deductions, were without information yield, only that their theory of semantic information was not designed to capture such a property. They coined the term *psychological information* for the property involved. See Floridi (2013) for further details.

A quantitative attempt at specifying the information yield of deductions was undertaken by Jaakko Hintikka with his theory of *surface information* and *depth information* (Hintikka 1970, 1973). The theory of surface and depth information extends Bar-Hillel and Carnap's theory of semantic information from the monadic predicate calculus all the way up to the full polyadic predicate calculus. This itself is a considerable achievement, but although technically astounding, a serious restriction of this approach is that it is only a fragment of the deductions carried out within full first-order logic that yield a non-zero information measure. The rest of the deductions in the full polyadic predicate calculus, as well as all of those in the monadic predicate calculus and propositional calculus, measure 0, (see Sequoiah-Grayson 2008).

The obvious inverse situation with the theory of classical semantic information is that logical contradictions, having a truth-likelihood of 0, will deliver a maximal information measure of 1. Referred to in the literature as the *Bar-Hillel-Carnap Semantic Paradox*, the most developed quantitative approach to addressing it is the theory of *strongly semantic information* (Floridi 2004). The conceptual motivation behind strongly semantic information is that for a statement to yield information, it must help us to narrow down the set of possible worlds. That is, it must assist us in the search for the actual world, so to speak (Sequoiah-Grayson 2007). Such a *contingency requirement on informativeness* is violated by both logical truths and logical contradictions, both of which measure 0 on the theory of strongly semantic information. See Floridi (2013) for further details.

## 2. Information as Correlation: Situation Theory

In general, the correlational take on information looks at *structured information environments* and how information flow is enabled within them by virtue of the systematic ways in which their parts are connected. An equally important theme in this view is the *aboutness of information*: the pattern of pixels that appear on the screen of a computer gives information (not necessarily complete) *about* the sequence of keys that were pressed by the person who is typing a document, and even a partial snapshot of the clear starry sky your friend is looking at now will give you information *about* his possible locations on Earth at this moment. These are typical examples involving structured environments in which a part can carry information *about* another one. The focus on structured environments and the aboutness of information goes also hand in hand with a third main topic of this approach, namely the *situatedness of information*, that is, its dependency on the particular setting on which an informational signal occurs. Take the starry sky as an example again: the

same pattern of stars, at different moments in time and locations in space will in general convey different information about the location of your friend.

Historically, the first paradigmatic setting of correlated information was studied by Shannon (1948). He considered a communication system formed by two information sites, a source and a receiver, which are connected via a noisy channel. Shannon's concern was purely quantitative. He gave conclusive and extremely useful answers to questions having to do with how to construct communication codes to be used in the source and receiver sites so that the effectiveness of communication can be maximised (in terms of bits of information that can be transmitted) while minimizing the possibility of errors caused by channel noise.

The logical approach to information as correlation builds on Shannon's ideas but is naturally concerned with the qualitative aspects of information flow that were highlighted above and that Shannon's theory does not address, such as: *what* information *about* a 'remote' site (remote in terms of space, time, perspective, etc.) can be drawn out of information that is directly available at a 'proximal' site? This leads to non-standard forms of inference that involve several situations at the same time and stress the *aboutness* of information.

*Situation theory* (Barwise and Perry 1983; Devlin 1991) is the major logical framework so far that has made these ideas its starting point for an analysis of information. Its origin and some of its central insights can be found in the project of naturalization of mind and the possibility of knowledge initiated by Fred Dretske (1981), which soon influenced the inception of situation semantics in the context of natural language (see Kratzer 2011).

Technically, there are two kinds of developments in situation theory:

- a. Set-theoretic and model-theoretic frameworks based on detailed ontologies suitable for modelling informational phenomena in concrete applications.
- b. A mathematical theory of information flow as enabled by lawful *channels* that connect parts of a whole. This theory takes a more abstract view on information as correlation applicable (in principle) to all sorts of systems that can be decomposed in interrelated parts.

The next three subsections survey some of the basic notions about the sites of information in situation theory (situations), the basic notion of information flow based on correlations between situations, and finally the mathematical theory of classifications and channels mentioned in (b).

## 2.1 Situations and Supporting Information

The ontologies in (a) span a wide spectrum of entities. They are meant to reflect the way in which an agent carves up a system. Here "a system" can be the world, or a part or aspect of it, and the agent (or kind of agent) can be an animal species, a device, a theorist, etc. The list of basic entities includes individuals, relations (which come with roles attached to them), temporal and spacial locations, and various other things. Distinctive among them are the *situations* and *infons*.

Roughly speaking, situations are highly structured parts of a system, such as a class session, a scene as seen from a certain perspective, a war, etc. Situations are the basic supporters of information. Infons, on the other hand, are the informational issues that situations may or may not support. The simplest kind of informational issue is whether some entities  $a_1, \dots, a_n$  stand (do not stand) in a relation  $R$  when playing the roles  $r_1, \dots, r_n$ , respectively. Such basic infon is usually denoted as

$$\langle\langle R, r_1 : a_1, \dots, r_n : a_n, i \rangle\rangle.$$



where  $i$  is 1 or 0, according to whether the issue is positive or negative.

Infons are not intrinsic bearers of truth, they are not claims either. They are simply informational issues that may or may not be supported by particular situations. We'll write  $s \models \sigma$  to mean that the situation  $s$  supports the infon  $\sigma$ . As an example, a successful transaction whereby Mary bought a piece of cheese in the local market is a situation that supports the infon

$$\sigma = \langle\langle \text{bought}, \text{what} : \text{cheese}, \text{who} : \text{Mary}, 1 \rangle\rangle.$$

This situation does *not* support the infon

$$\langle\langle \text{bought}, \text{what} : \text{cheese}, \text{who} : \text{Mary}, 0 \rangle\rangle$$

because Mary did buy cheese. Nor does the situation support the infon

$$\langle\langle \text{landed}, \text{who} : \text{Armstrong}, \text{where} : \text{Moon}, 1 \rangle\rangle,$$

because Armstrong is not part of the situation in question at all.

The discrimination or individuation of a situation by an agent does not entail that the agent has full information about it: when we wonder whether the local market is open, we have individuated a situation about which we actually *lack* some information. See Textor (2012) for a detailed discussion on the nature of situation-like entities and their relation with other ontological categories such as the possible worlds used in modal logic.

Besides individuals, relations, locations, situations and basic infons, there are usually various kinds of parametric and abstract entities. For example, there is a mechanism of *type abstraction*. According to it, if  $y$  is a parameter for situations, then

$$T_y = [y \mid y \models \langle\langle \text{bought}, \text{what} : \text{cheese}, \text{who} : x, 1 \rangle\rangle]$$

is the type of situations where somebody buys cheese. There will be some basic types in an ontology, and many other types obtained via abstraction, as just described.

The collection of ontology entities also include propositions and constraints. They are key in the formulation of the basic principles of information content in situation theory, to be introduced next.

## 2.2 Information flow and constraints

The following are typical statements about “information flow” as studied in situation theory:

[E1] The fact that the dot in the radar screen is moving upward indicates that flight A123 is moving northward.

[E2] The presence of footprints of pattern  $P$  in Zhucheng indicates that a dinosaur lived in the region millions of years ago.

The general scheme has the form

[IC] *That  $s : T$  indicates that  $p$ .*

where  $s : T$  is notation for “ $s$  is of type  $T$ ”. The idea is that it is concrete parts of the world that act as carriers of information (the concrete dot in the radar or the footprints in Zhucheng), and that they do so by virtue of being of a certain type (the dot moving upward or the footprints showing a certain pattern). What each of these concrete instances indicates is a fact about another correlated part of the world. For the issues to be discussed below it will suffice to consider cases where the indicated fact—that is  $p$  in the formulation of [IC]—is of the form  $s' : T'$ , as in the radar example.

The conditions needed to verify informational signalling in the sense of [IC] rely on the existence of law-like *constraints* such as natural laws, necessary laws such as those of math, or conventions thanks to which (in part) one situation may serve as carrier of information about another one. Constraints specify the *correlations* that exist between situations of various types, in the following sense: if two types  $T$  and  $T'$  are subject to the constraint  $T \Rightarrow T'$ , then for every situation  $s$  of type  $T$  there is a relevantly connected situation  $s'$  of type  $T'$ . In the radar example, the relevant correlation would be captured by the constraint  $\text{GOINGUPWARD} \Rightarrow \text{GOINGNORTH}$ , which says that each situation where a radar point moves upward is connected with another situation where a plane is moving to the north. It is the existence of this constraint that allows a particular situation where the dot moves to indicate something about the connected plane situation.

With this background, the verification principle for information signalling in situation theory can be formulated as follows

**[IS Verification]**  $s : T$  indicates that  $s' : T'$  if  $T \Rightarrow T'$  and  $s$  is *relevantly connected* to  $s'$ .

The relation  $\Rightarrow$  is transitive. This ensures that Dretske's Xerox principle holds in this account of information transfer, that is, there can be no loss of semantic information through information transfer chains.

**[Xerox Principle]:** If  $s_1 : T_1$  indicates that  $s_2 : T_2$  and  $s_2 : T_2$  indicates that  $s_3 : T_3$ , then  $s_1 : T_1$  indicates that  $s_3 : T_3$ .

The [IS Verification] principle deals with information that *in principle* could be acquired by an agent. The access to some of this information will be blocked, for example, if the agent is oblivious to the correlation that exists between two kinds of situations. In addition, most correlations are not absolute, they admit exceptions. Thus, for the signalling described in

[E1] to be really informational, the extra *condition* that the radar system is working properly must be met. Conditional versions of the [IS Verification] principle may be used to insist that the carrier situation must meet certain background conditions. The inability of an agent to keep track of changes on these background conditions may lead to errors. So, if the radar is broken, the dot on the screen may end up moving upward while the plane is moving south. Unless the air controller is able to recognise the problem, that is, unless she realises that the background conditions have changed, she may end up giving absurd instructions to the pilot. Now, instructions are tied to actions. For a treatment of actions from the situation-theoretical view, we refer the reader to Israel and Perry (1991).

## 2.3 Distributed information systems and channel theory

The basic notion of information flow sketched in the previous section can be lifted to a more abstract setting in which the supporters of information are not necessarily situations as concrete parts of the world, but rather any entity which, as in the case of situations, can be classified as being of or not of certain types. The mathematical theory of distributed systems (Barwise and Seligman 1997) to be described next takes this abstract approach by studying information transfer within distributed systems in general.

A model of a distributed system in this framework will actually be a model of a *kind* of distributed system, hence the model of the radar-airplane system that we will use as a running example here will actually be a model of radar-airplane *systems* (in plural). Setting such a model requires describing the architecture of the system in terms of its parts and the way they are put together into a whole. Once that is done, one can proceed to see how that architecture enables the flow of information among its parts.

A part of a system (again, really its kind) is modelled by saying how particular instances of it are classified according to a given set of types. In other words, for each part of a system one has a *classification*

$$\mathbf{A} = \langle \text{Instances}, \text{Types}, \models \rangle,$$

where  $\models$  is a binary relation such that  $a \models T$  if the instance  $a$  is of type  $T$ . In a simplistic analysis of the radar example, one could posit at least three classifications, one for the monitor screen, one for the flying plane, and one for the whole monitoring system:

$$\mathbf{Screens} = \langle \text{Monitor-Screens}, \text{Types of Screen Configurations}, \models_M \rangle$$

$$\mathbf{Planes} = \langle \text{Flying Planes}, \text{Types of Flying Planes}, \models_P \rangle$$

$$\mathbf{MonitSit} = \langle \text{Monitoring Situations}, \text{Types of Monitoring Situations}, \models_M \rangle$$

A general version of a ‘part-of’ relation between classifications is needed in order to model the way parts of a system are assembled together. Consider the case of the monitoring systems. That each one of them has a screen as one of its parts means that there is a function that assigns to each instance of the classification **MonitSit** an instance of **Screens**. On the other hand, all the ways in which a screen can be classified (the types of **Screens**) intuitively correspond to ways in which the whole screening system could be classified: if a screen is part of a monitoring system and the screen is blinking, say, then the whole monitoring situation is intuitively one of the type ‘its screen is blinking’. Accordingly, a generalised ‘part-of’ relation between any two arbitrary classifications **A**, **C** can be modelled via two functions

$$f^\wedge : \text{Types}_A \rightarrow \text{Types}_C \text{ and } f^\vee : \text{Instances}_C \rightarrow \text{Instances}_A,$$

the first of which takes every type in **A** to its counterpart in **C**, and the second of which takes every instance  $c$  of **C** to its **A**-component.<sup>[3]</sup>

If  $f : \mathbf{A} \rightarrow \mathbf{C}$  is shortcut notation for the existence of the two functions above (the pair  $f$  of functions is called an *infomorphism*), then an arbitrary distributed system will consist of various classifications related by infomorphisms. For our purpose it will suffice here to consider three classifications **A**, **B**, **C** together with two infomorphisms

$$f : \mathbf{A} \rightarrow \mathbf{C} \text{ and } g : \mathbf{B} \rightarrow \mathbf{C}.$$

So a simple way to model the radar monitoring system would consist of the pair

$$f : \mathbf{Screens} \rightarrow \mathbf{MonitSit} \text{ and } g : \mathbf{Planes} \rightarrow \mathbf{MonitSit}.$$

The common codomain in these cases (**C** in the general case and **MonitSit** in the example) works as the core of a *channel* that connects two parts of the system. The core determines the correlations that obtain between the two parts, thus enabling information flow of the kind discussed in section 2.2. This is achieved via two kinds of links. On the one hand, two instances  $a$  from **A** and  $b$  from **B** can be thought to be connected via the channel if they are components of the same instance in **C**, so the instances of **C** act as connections between components. Thus, in the radar example a particular screen will be connected to a particular plane if they belong to the same monitoring situation.

On the other hand, suppose that every instance in **C** verifies some relation between types that happen to be counterparts of types from **A** and **B**. Then such relation captures a *constraint* on how the parts of the system are correlated. In the radar example, the theory of the core classification **MonitSit** will include constraints such as `PLAINMOVINGNORTH`  $\Rightarrow$  `DOTGOINGUP`. This regularity of monitoring situations, which act as

connections between radar screen-shots and planes, reveals a way in which radar screens and monitored planes correlate with each other. All this affords the following version of information transfer.

**Channel-enabled signalling:** Suppose that

$$f: \mathbf{A} \rightarrow \mathbf{C} \text{ and } g: \mathbf{B} \rightarrow \mathbf{C}.$$

Then instance  $a$  being of type  $T$  in  $\mathbf{A}$  indicates that instance  $b$  is of type  $T'$  in  $\mathbf{C}$  if  $a$  and  $b$  are connected by an instance from  $\mathbf{C}$  and the relation  $f^{\wedge}(T) \Rightarrow g^{\wedge}(T')$  between the counterpart interpreted types is satisfied by all instances of  $\mathbf{C}$ .

Now, for each classification  $\mathbf{A}$ , the collection

$$L_{\mathbf{A}} = \{T \Rightarrow T' \mid \text{every instance of } \mathbf{A} \text{ of type } T \text{ is also of type } T'\}$$

formed by all the *global constraints* of the classification can be thought of as a logic that is intrinsic to  $\mathbf{A}$ . Then a distributed system consisting of various classifications and infomorphisms will have a logic of constraints attached to each part of it,<sup>[4]</sup> and more sophisticated questions about information flow within the system can be formulated.

For example, suppose an infomorphism  $f: \mathbf{A} \rightarrow \mathbf{C}$  is part of the distributed system under study. Then  $f$  naturally transforms each global constraint  $T \Rightarrow T'$  of  $L_{\mathbf{A}}$  into  $f^{\wedge}(T) \Rightarrow f^{\wedge}(T')$ , which can always be shown to be an element of  $L_{\mathbf{C}}$ . This means that one can reason within  $\mathbf{A}$  and then *reliably* draw conclusions about  $\mathbf{C}$ . On the other hand, it can be shown that using preimages under  $f^{\wedge}$  in order to translate global constraints of  $\mathbf{C}$  does *not* always guarantee the result to be always a global constraint of  $\mathbf{A}$ . It is then desirable to identify extra conditions under which the reliability of the inverse translation can be guaranteed, or at least improved. In a sense,

these questions are qualitatively close to the concerns Shannon originally had about noise and reliability.

Another issue one may want to model is reasoning about a system from the perspective of an agent that has only *partial* knowledge about the parts of a system. For a running example, think of a plane controller who has only worked with ACME monitors and knows nothing about electronics. The logic such an agent might use to reason about part  $\mathbf{A}$  of a system (actually part **Screens** in the case of the controller) will in general consist of some constraints that may not even be global, but satisfied only by some subset of instances (the ACME monitors). The agent's logic may be *incomplete* in the sense that it might miss some of the global constraints of the classification (like the ones involving inner components of the monitor). The agent's logic may also be *unsound*, in the sense that there might be instances out of the awareness of the agent (say monitors of unfamiliar brands) that falsify some of the agent's constraints (which do hold of all ACME monitors). As before, when part  $\mathbf{A}$  of a system is linked to part  $\mathbf{C}$ , one also wants to know whether the linking infomorphism allows to translate an agent's logic from  $\mathbf{A}$  while preserving good properties it might have (such as not having exceptions for its constraints).

For an extensive development of the theory of channels sketched here, plus several explorations towards applications, see Barwise and Seligman (1997). See van Benthem (2000) for a study of conditions under which constraint satisfiability is preserved under infomorphisms, and Allo (2009) for an application of this framework to an analysis of the distinction between cognitive *states* and cognitive *commodities*. Finally, it must be mentioned that the notion of classification has been around for some years now in the literature, having been independently studied and introduced under names such as Chu spaces (Pratt 1999) or Formal Contexts (Ganter and Wille 1999).

### 3. Information as Code

For information to be computed, it must be handled by the computational mechanism in question, and for such a handling to take place, the information must be *encoded*. *Information as Code* is a stance that takes this encoding-condition very seriously. The result is the development of fine-grained models of information flow that turn on the syntactic properties of the encoding itself.

To see how this is so, consider again cases involving information flow via observations. Such observations are informative because we are not omniscient in the normal, God-like sense of the term. We have to go and observe that the cat is on the mat precisely because we are not automatically aware of every fact in the universe. Inferences work in an analogous manner. Deductions are informative for us precisely because we are not *logically* omniscient. We have to reason about matters, sometimes at great length, because we are not automatically aware of the logical consequences of the body of information that we are reasoning with.

To come full circle—reasoning explicitly with information requires handling it, where in this case such handling is cognitive act. Hence the information in question is encoded in some manner, hence Information as Code underpins the development of fine-grained models of information flow that turn on the syntactic properties of the encoding itself, as well as the properties of the actions that underpin the various information-processing contexts involved.

Such information-processing contexts are not restricted to explicit acts of inferential reasoning by human agents, but include *automated reasoning* and *theorem proving*, as well as machine-based computational procedures in general. Approaches to modelling the properties of these latter

information-processing scenarios fall under *algorithmic information theory*.

In section 3.1, we will explore a major approach to modelling the properties of information-processing within the information as code framework via *categorical information theory*. In section 3.2, we will examine the more general approach to modelling information as code of which categorical information theory is an instance, the modelling of information as code via *substructural logics*. In section 3.3 we will lay out the details of several other notable examples of logics of information flow motivated by the information as code approach.

#### 3.1 Categorical Information Theory

*Categorical information theory* is a theory of fine-grained information flow whose models are based upon those specified by the categorical grammars underpinned by the Lambek Calculi, due originally to Lambek (1958, 1961). The motivation for categorical information theory is to provide a logical framework for modelling the properties of the very cognitive procedures that underpin deductive reasoning.

The conceptual origin of categorical information theory is found in van Benthem (1995: 186):

[I]t turns out that, in particular, the Lambek Calculus itself permits of procedural re-interpretation, and thus, categorial calculi may turn out to describe cognitive procedures just as much as the syntactic or semantic structures which provided their original motivation.

The motivation for categorical information theory is to model the cognitive procedures constituting deductive reasoning. Consider as an analogy the following example. You arrive home from IKEA with an unassembled

table that is still flat-packed in its box. Now the question here is this, do you have your table? Well, there is a sense in which you do, and a sense in which you do not. You have your table in the sense that you have all of the pieces required to construct or generate the table, but this is not to say that you have the table in the sense that you are able to *use* it. That is, you do not have the table in any useful form, you have merely pieces of a table. Indeed, getting these table-pieces into their useful form, namely a table, may be a long and arduous process...

The analogy between the table-example above and deductive reasoning is this. It is often said that the information encoded by (or “contained in” or “expressed by”) the conclusion of a deductive argument is encoded by the premises. So, when you possess the information encoded by the premises of some instance of deductive reasoning, do you possess the information encoded by the conclusion? Just as with the table-pieces, you do not possess the information encoded by the conclusion in any useful form, not until you have put the “information-pieces” constituting the premises together in the correct manner. To be sure, when you possess the information-pieces encoded by the premises, you possess all of the information required for the construction or generation of the information encoded by the conclusion. As with the table-pieces however, getting the information encoded by the conclusion from the information encoded by the premises may be a long and arduous process. This information-generation via deductive inference may be thought of also as the movement of information from implicit to explicit storage in the mind of the reasoning agent, and it is the cognitive procedures facilitating this storage transfer that motivate categorial information theory.

Categorial information theory is a theory of dynamic information processing based on the *merge/fusion* ( $\otimes$ ) and *typed function* ( $\rightarrow$ ,  $\leftarrow$ ) operations from categorial grammar. The conceptual motivation is to understand the information in the mind of an agent as the agent reasons

deductively to be a database in much the same way as a natural language lexicon is a database, Sequoiah-Grayson (2013). In this case, a *grammar* will be understood as a set of processing constraints so imposed as to guarantee information flow, or well-formed strings as outputs. Recent research on proofs as *events* from a very similar conceptual starting point may be found in Stefaneas and Vandoulakis (forthcoming).

Categorial information theory is strongly algebraic in flavour. Fusion ‘ $\otimes$ ’ corresponds to the binary composition operator ‘.’, and ‘ $\vdash$ ’ to the partial order ‘ $\leq$ ’ (see Dunn 1993). The merge and function operations are related to each other via the familiar *residuation conditions*:

$$(5) A \otimes B \vdash C \text{ iff } B \vdash A \rightarrow C$$

$$(6) A \otimes B \vdash C \text{ iff } A \vdash C \leftarrow B$$

In general, applications for directional function application will be restricted to algebraic analyses of grammatical structures, where commuted lexical items will result in non-well-formed strings.

Despite its algebraic nature, the operations can be given their evaluation conditions via “informationalised” Kripke frames (Kripke 1963, 1965). An information frame (Restall 1994)  $\mathbf{F}$  is a triple  $\langle S, \sqsubseteq, \bullet \rangle$ .  $S$  is a set of information states  $x, y, z, \dots$ .  $\sqsubseteq$  is a partial order of informational development/inclusion such that  $x \sqsubseteq y$  is taken to mean that the information carried by  $y$  is a development of the information carried by  $x$ , and  $\bullet$  is an operation for combining information states. In other words, we have a domain with a combination operation. The operation of information combination and the partial order of information inclusion interrelate as follows:

$$(7) x \sqsubseteq y \text{ iff } x \bullet y \sqsubseteq y$$

Reading  $x \Vdash A$  as *state  $x$  carries information of type  $A$* , we have it that:

- (8)  $x \Vdash A \otimes B$  iff for some  $y, z, \in \mathbf{F}$  s.t.  $y \bullet z \sqsubseteq x$ ,  $y \Vdash A$  and  $z \Vdash B$ .
- (9)  $x \Vdash A \rightarrow B$  iff for all  $y, z \in \mathbf{F}$  s.t.  $x \bullet y \sqsubseteq z$ , if  $y \Vdash A$  then  $z \Vdash B$ .
- (10)  $x \Vdash B \leftarrow A$  iff for all  $y, z \in \mathbf{F}$  s.t.  $y \bullet x \sqsubseteq z$ , if  $y \Vdash A$  then  $z \Vdash B$ .

At the syntactic level, we read  $X \vdash A$  as *processing on  $X$  generates information of type  $A$* . In this case we are understanding  $\vdash$  as an information processing mechanism as suggested by Wansing (1993: 16), such that  $\vdash$  encodes not just the output of an information processing procedure, but the properties of the procedure itself. Just what this processing consists of will depend on the processing constraints that we set up on our database. These processing constraints will be imposed in order to guarantee an output from the processing itself, or to put this another way, in order to preserve information flow. Such processing constraints are fixed by the presence or absence of various *structural rules*, and structural rules are the business of *substructural logics*.

### 3.2 Substructural logics and information flow

Categorical information theory is precipitated by giving the Lambek calculi an informational semantics. At a suitable level of abstraction, the Lambek calculi is seen to be a highly expressive *substructural logic*. Unsurprisingly, by giving an informational semantics for substructural logics in general, we get a hierarchy of logics that exemplify the information as code approach. This hierarchy is organised by expressive power, with the expressive power of the logics in question being captured by the presence of various *structural rules*.

A structural rule is of the following general form:

$$(11) X \Leftarrow Y$$

We may read (11) as *any information generated by processing on  $X$  is generated by processing on  $Y$  also*. Hence the long-form of (11) is as follows:

$$(12) \frac{X \vdash A}{Y \vdash A}$$

Hence  $X$  is a structured body of information, or “data structure” as Gabbay (1996: 423) puts it, where the actual *arrangement* of the information plays a crucial role. The structural rules will fix the structure of the information encoded by  $X$ , and as such impact upon the granularity of the information being processed.

Consider Weakening, the most familiar of the structural rules (followed by its corresponding frame condition:

$$\begin{aligned} \text{(Weakening)} \quad & A \Leftarrow A \otimes B \\ & x \bullet y \sqsubseteq z \rightarrow x \sqsubseteq z \end{aligned}$$

With Weakening present, we loose track of which pieces of information were actually used in an inference. This is precisely why it is that the rejection of Weakening is the mark of relevant logics, where the preservation of bodies of information relevant to the derivation of the conclusion is the motivation. By rejecting Weakening, we highlight a certain type of informational *taxonomy*, in the sense that we know *which* bodies of information were used. To preserve more structural detail than simply which bodies of information were used, we need to consider rejecting further structural rules.

Suppose that we want to record not only which pieces of information were used in an inference, but also how often they were used. In this case we

would reject Contraction:

$$\begin{aligned} \text{(Contraction)} \quad & A \otimes A \Leftarrow A \\ & x \bullet x \sqsubseteq x \end{aligned}$$

Contraction allows the multiple use, without restriction, of a piece of information. So if keeping a record of the “informational cost” of the execution of some information processing is a concern, Contraction will be rejected. The rejection of Contraction is the mark of linear logics, which were designed for modelling just such processing costs (see Troelstra 1992).

If we wish to preserve the *order* of use of pieces of information, then we will reject the structural rule of Commutation:

$$\begin{aligned} \text{(Commutation)} \quad & A \otimes B \Leftarrow B \otimes A \\ & x \bullet y \sqsubseteq z \rightarrow y \bullet x \sqsubseteq z \end{aligned}$$

Information-order will be of particular concern in temporal settings (consider action-composition) and natural language semantics (Lambek 1958), where non-commuting logics first appeared. Commutation comes also in a more familiar strong form:

$$\begin{aligned} \text{(Commutation (strong))} \quad & (A \otimes B) \otimes D \Leftarrow (A \otimes D) \otimes B \\ & \exists u(x \bullet z \sqsubseteq u \wedge u \bullet y \sqsubseteq w) \rightarrow \exists u(x \bullet y \sqsubseteq u \wedge \\ & \quad u \bullet z \sqsubseteq w) \end{aligned}$$

The strong form of Commutation results from its combination with the structural rule of Association:<sup>[5]</sup>

$$\begin{aligned} \text{(Association)} \quad & A \otimes (B \otimes C) \Leftarrow (A \otimes B) \otimes C \\ & \exists u(x \bullet y \sqsubseteq u \wedge u \bullet z \sqsubseteq w) \rightarrow \exists u(y \bullet z \sqsubseteq u \wedge x \bullet u \sqsubseteq w) \end{aligned}$$

Rejecting Association will preserve the precise fine-grained properties of the combination of pieces of information. Non-associative logics were introduced originally to capture the combinatorial properties of language syntax (see Lambek 1961).

In the presence of Commutation, a double implication pair ( $\rightarrow$ ,  $\leftarrow$ ) collapses into single implication  $\rightarrow$ . In the presence of all of the structural rules, fusion,  $\otimes$ , collapses into Boolean conjunction,  $\wedge$ . In this case, the residuation conditions outlined in (5) and (6) collapse into a mono-directional function.

The choice of which structural rules to retain obviously depends on just what informational phenomena is being modelled, so there is a strong *pluralism* at work. By rejecting Weakening say, we are speaking of *which* data were relevant to the process, but are saying nothing about its multiplicity (in which case we would reject Contraction), its order (in which case we would reject Commutation), or the actual patterns of use (in which case we would reject Association). By allowing Association, Commutation, and Contraction, we have the taxonomy locked down. We might not know the order or multiplicity of the data that were used, but we do know what types, and exactly what types, were relevant to the successful processing. The canonical contemporary exposition of such an information-based interpretation of propositional relevant logic is Mares (2004). Such an interpretation allows for an elegant treatment of the contradictions encoded by relevant logics. By distinguishing between *truth conditions* and *information conditions*, we allow for an interpretation of  $x \Vdash A \wedge \neg A$  as *x carries the information that A and not A*. For an exploration of the distinction between truth-conditions and information-conditions within *quantified* relevant logic, see Mares (2009).

At such a stage, things are still fairly *static*. By shifting our attention from static bodies of information, to the manipulation *of* these bodies, we will



reject structural rules beyond Weakening, arriving ultimately at categorical information theory, as it is encoded by the very weakest substructural logics. Hence the weaker we go, the more “procedural” the flavour of the logics involved. From a dynamic/procedural perspective, linear logics might be thought of as a “half way point” between static classical logic, and fully procedural categorical information theory. For a detailed exposition of the relationship between linear logic and other formal frameworks in the context of modelling information flow, see Abramsky (2008).

### 3.3 Related Approaches

The information as code approach is a very natural perspective on information flow, hence there are a number of related frameworks that exemplify it.

One such approach to analysing information as code is to carry out such an analysis in terms of the computational complexity of various propositional logics. Such an approach may propose a hierarchy of propositional logics that are all decidable in polynomial time, with this hierarchy being structured by the increasing computational resources required for the proofs in the various logics. D'Agostino and Floridi (2009) carry out just such an analysis, with their central claim being that this hierarchy may be used to represent the increasing levels of informativeness of propositional deductive reasoning.

Gabbay's (1993, 1996) framework of *labelled deductive systems* exemplifies the information as code approach in manner very similar to the informationalised substructural logics of section 3.1. An item of data is given as a pair of the form  $x : A$ , where  $A$  is a piece of declarative information, and  $x$  is a label for  $A$ .  $x$  is a representation of information that is needed operate on or alter the information encoded by  $A$ . Suppose that

we have also the data-pair  $y : A \rightarrow B$ . We may apply  $x$  to  $y$ , resulting in the data-pair  $x + y : B$ . In this case, a database is a configuration of labelled formulas, or data-pairs (Gabbay 1993: 72). The labels and their corresponding application operation are organised by an algebra, and the properties of this algebra will impose constraints on the applications operation. Different constraints, of “meta-conditions” as Gabbay calls them (Gabbay 1993: 77), will correspond to different logics. For example, if we were to ignore the labels, then we would have classical logic, if we were to accept only the derivations which used all of the labelled assumptions, then we would have relevance logic, and if we accepted only the derivations which used the labelled assumptions exactly once, then we would have linear logic. Labels are behaving very much like possible worlds here, and the short step from possible worlds to information states makes it obvious how it is that the meta-conditions on labels may be captured by structural rules.

Artemov's (2008) framework of *justification logic* shares many surface similarities with Gabbay's system of labelled deduction. The logic is composed of *justification assertions* of the form  $x : A$ , read as  *$x$  is a justification for  $A$* . Justifications themselves are evidential bases of varying sorts that will vary depending on the context. They might be mathematical proofs, sets of causes or counterfactuals, or something else that fulfils the justificatory role. What it means for  $x$  to justify  $A$  is not analysed directly in justification logic. Rather, attempts are made to characterise the justification relation  $x : A$  itself, via various operations and their axioms. The application operation, ‘.’ mimics the application operation ‘+’ from labelled deduction, or the fusion ‘ $\otimes$ ’ operation from categorical information theory. In justification logic, the symbol ‘+’ is reserved for the representation of joint evidence. Hence ‘ $x + y$ ’ is read as ‘*the joint evidence of  $x$  and  $y$* ’. Application and join are characterised in justification logic by the following axioms respectively:

$$(13) \ x : (A \rightarrow B) \rightarrow (y : A \rightarrow (x . y) : B)$$

$$(14) \ x : A \rightarrow (x + y) : A, \text{ and } x : A \rightarrow (y + x) : A$$

The latter axiom characterises the monotonicity of joint evidential bases. Apart from the commutativity of  $+$ , the structural properties of the justification operations are currently unexplored, although the potential for such an exploration is exciting. Justification logic is used to analyse notoriously difficult epistemic problems such as the Gettier cases and more. If we take our epistemology to be informationalised, then the constitution of evidential bases as information states places justification logics within the information as code approach in a straightforward manner. For further details, see Artemov and Fitting (2012).

Zalta's work on object theory (Zalta 1983, 1993) provides a different way to analyse informational content—understood as propositional content—and its structure. Motivated by metaphysical considerations, object theory starts by proposing a theory of objects and relations (usually formulated in a second order quantified modal language). This theory can then be used to define and characterise states of affairs, propositions, situations, possible worlds, and other related notions. The resulting picture is one where all these things have internal structure, their algebraic properties are axiomatized, and one can therefore reason about them in a classical proof-theoretical way.

A philosophical point touched by this approach concerns the link between the propositional content (information) expressed by sentences and the idea of predication. Relevant to this entry is Zalta's (1993) development of a version of situation theory that follows this approach, and where a key element is the usage of two forms of predication. Briefly, the formula ' $Px$ ' corresponds to the usual form of predication by exemplification (as in "Obama is American"), while ' $xP$ ' corresponds to predication via

*encoding*. Abstract objects are then defined to be (essentially) encodings of properties, in combinations which might not even be made factual. These provisions enable the existence of information about abstract, possible, or fictional entities. For details on the tradition to which object theory belongs see Textor (2012), McGrath (2012), and King (2012).

## 4. Connections Between the Approaches

While the three approaches discussed above (range, correlations, code) differ in that they emphasise different informational themes, the underlying notion they aim to clarify is the same (information). It is then natural to find that the similarities and synergies between the approaches invite the exploration of ways to combine them. Each one of the next subsections illustrates how one could bring together two out of the three approaches. Section 4.1 exemplifies the interface between the info-as-range and info-as-correlation views. Sections 4.2 and 4.3 do the same with the other two pairs of combinations, namely code and correlations, and code and ranges.

### 4.1 Ranges and correlations

A central intuition in the information-as-range view is the correspondence that exists between information at hand (where this can be qualified in various ways) and the range of possibilities which are compatible with such information. On the other hand, a key feature of the correlational approach to information is its reliance on a structured information system formed by components that are systematically connected. In general, many properties of a structured system will actually be *local* properties, in that they are determined by only some of the components (the fact that there is a dot moving upwards in a radar can be determined only by looking at the screen, even if this behaviour is correlated with the motion of a remote plane, which is another component of the system). If one has access to

information pertaining to only a few of the many components of a system, a natural notion of range of possibilities arises, consisting of all the possible global configurations of the system that are compatible with such local information. This subsection expands on this particular way to link the two approaches, but as it will be noted at the end, this is not the only one and the search for other ways lies ahead as an open area of inquiry.

Formally, the link between ranges and correlations described above may be approached by using a *restricted* product state space as a model of the architecture of the system (van Benthem and Martinez 2008). The basic structures are *constraint models*, versions of which have been around in the literature for some years (for example Fagin et al. 1995 in the study of epistemic logic, and Ghidini and Giunchiglia 2001 in the study of context dependent reasoning). Constraint models have the form

$$\mathcal{M} = \langle \text{Comp}, \text{States}, C, \text{Pred} \rangle.$$

Here, the basic component spaces are indexed by *Comp*, the states of each component are taken from *States* (with different components using maybe only a few of the elements of *States*), and the global states of the system are global valuations, that is, functions that assign a state to each basic component *Comp*. Not all such functions are allowed, only those in *C*. Finally, *Pred* is a labelled family of predicates (sets of global states).

To see how this fits with the information-as-correlation view, consider again the example of planes being monitored by radars. As before, each monitoring situation will be modelled as having only two parts, now indexed by the members of  $\text{Comp} = \{\text{screen}, \text{plane}\}$ . The actual instances of screening situations could then correspond to global states, which in this case with only two components can be thought of as pairs  $(s, b)$  where  $s$  is a particular screen and  $b$  a particular plane. Hence, global states connect instances of parts, so representing instances of a whole system.

But then a crucial restriction comes into play, because not all screens are connected with all planes, only with those belonging to the same monitoring situation. The set *C* selects only such permissible pairs, thus playing a role similar to that of a channel in section 2.3. Finally, *Pred* classifies global states into types, similar to the classification relations of section 2.3.

As we said before, some properties of systems are local properties, with only some of the components of the systems being relevant in determining whether it holds or not. That a monitoring situation is one where the plane is moving North depends only on the plane, not on the screen. In general, if a property is completely determined by subset of components  $\mathbf{x}$  then, in what concerns that property, any two global states that agree on  $\mathbf{x}$  should be indistinguishable. In fact, each such  $\mathbf{x}$  induces an equivalence relation of local property determination so that for every two global states  $\mathbf{s}, \mathbf{t}$ :

$\mathbf{s} \sim_{\mathbf{x}} \mathbf{t}$  if and only if the values of  $\mathbf{s}$  and  $\mathbf{t}$  at each one of the components in  $\mathbf{x}$  are the same.

In this way one gets not only a conceptual but also a formal link to the information-as-range approach, because constraint models can then be used to interpret a basic modal language with atomic formulas of the form  $P$ —where  $P$  is one of the labels of predicates in *Pred*—and with complex formulas of the form  $\neg\phi$ ,  $\phi \vee \psi$ ,  $U\phi$ , and  $\Box_{\mathbf{x}}\phi$ , where  $\mathbf{x}$  is a partial tuple of components and  $U$  is the universal modality. More concretely, given a constraint model  $\mathcal{M}$  and a global state  $s$ , the crucial satisfaction conditions are given by:

- $\mathcal{M}, \mathbf{s} \models P$  if and only if  $\mathbf{s} \in P$
- $\mathcal{M}, \mathbf{s} \models U\phi$  if and only if  $\mathcal{M}, \mathbf{t} \models \phi$  for all  $\mathbf{t}$
- $\mathcal{M}, \mathbf{s} \models \Box_{\mathbf{x}}\phi$  if and only if  $\mathcal{M}, \mathbf{t} \models \phi$  for all  $\mathbf{t} \sim_{\mathbf{x}} \mathbf{s}$

The resulting logic is axiomatised by the fusion of  $S_5$  modal logics for the universal modality  $U$  and each one of the  $\Box_x$  modalities, plus the addition of axioms of the form  $U\phi \rightarrow \Box_x\phi$ , and  $\Box_x\phi \rightarrow \Box_y\phi$  whenever  $\sim_y \subseteq \sim_x$ .

The information-as-range research agenda includes other topics, such as agency and the dynamics of information update, which can in principle be incorporated to the constraint models setting (van Benthem and Martinez 2008). For example, in the case of agency, to the architectural structure of a state system captured by a constraint model one could add epistemic accessibility relations for a group of agents  $\mathcal{A}$ , so to obtain *epistemic constraint models* of the form

$$\mathcal{M} = \langle \text{Comp}, \text{States}, C, \text{Pred}, \{\approx_a\}_{a \in \mathcal{A}} \rangle.$$

where  $\approx_a$  is the equivalence accessibility relation of agent  $a$ . Here one could refine the planes and radar example above by adding some agents, say the controller and the pilot. By relying only on the controls each agent can see, the controller will not be able to distinguish states that agree on the direction of the plane but that differ, say, on the meteorological conditions around the plane. Those states will be related by the controller's relation in the model, but not by the pilot's relation. In principle, this merge of modal epistemic models and constraint models allows one to study, in a single setting, aspects of both the information-as-range and information-as-correlation points of view. The corresponding logical language for epistemic constraint models is the same as for basic constraint models, expanded with the  $K_i$  modal operators, one per agent. The logic is the fusion of the constraint logic from above and a  $S_5$  logic per each agent  $a$ .

As mentioned earlier, constraint models exploit a particular kind of link that exists between information-as-correlation and information-as-range, a link that arises from the locality of properties in a structured system and

yields models (constraint models) that are formally closer to the ones used in the information-as-range approach. But other links between the approaches may be found that come from addressing other kind of questions, and their resulting formal models may end up being closer to situation theoretic-ones. For example, one may ask: how could a relation of epistemic accessibility arise from a setting of epistemic states in which agents may have *incomplete* information about an intended range of such states? This is not the case with the epistemic constraint models described above, where agents have complete information about what holds true of all the epistemically accessible worlds. One way to address this question (Barwise 1997) is to consider a fixed classification  $A$  the instances of which are the epistemic states, and with a local logic per agent attached to each state. For some states these local logic may be incomplete (see section 2.3), so agents may not have information about everything that holds true of the intended range of states. Then, roughly, the states accessible from a given state  $s$  and agent  $a$  will be those whose properties (types) do not contradict the local logic of  $a$  in  $s$ . With these epistemic relations in place, the expanded classification  $A$  can be used to interpret a basic modal language.

## 4.2 Code and correlations

Logical frameworks that crossover information as code and information as correlation get their most explicit representation in work that does just this —model the crossover between the two frameworks. Restall (1994) and Mares (1996) give independent proofs of the representability of Barwise's information as correlation channel-theoretic framework within the information as code approach as exemplified by substructural logics framework. In this section we will trace the motivations and the main details of the proof, before demonstrating the connection with *category theory*.

The basic steps are these—if we understand information channels to be information states of a special sort, namely the sort of information state that carries information of conditional types, then there is an obvious meeting point between information as correlation as exemplified by channel theory, and information as code as exemplified by informationalised substructural logics. The intermediate step is to reveal the connection between channel semantics for conditional types, and the frame semantics for conditionals given by relevance logics.

Starting with the channel theoretic analysis of conditionals, as noted already, the running motivation behind Barwise's channel-theoretic framework is that information flow is underpinned by an information channel. Barwise understood conditionals as *constraints* in the sense that  $A \rightarrow B$  is a constraint from  $A$  to  $B$  in the sense of  $A \Rightarrow B$  from section 2.2 above. In the information that  $A$  is combined with the information encoded by the constraint, then the result or output is the information that  $B$ .

The information that  $A$  and that  $B$  is carried by the situations  $s_1, s_2, \dots$  and the information encoded by the constraint is carried by an information channel  $c$ . Given this, Barwise's evaluation condition for a constraint is as follows (the condition is given here in Barwise's notation from his later work on conditionals, although in earlier writings such conditions appeared in the notation given in section 2.2 above):

$$(15) \quad c \models A \rightarrow B \text{ iff for all } s_1, s_2, \text{ if } s_1 \xrightarrow{c} s_2 \text{ and } s_1 \models A, \text{ then } s_2 \models B$$

$s_1 \xrightarrow{c} s_2$  is read as

the information carried by the channel  $c$ , when combined with the information carried by the situation  $s_1$ , results in the information carried by the situation  $s_2$ .

Obviously enough, this is very close in spirit to (9) in the section on information as code above.

As noted above, the intermediate step concerns the ternary relation  $R$  from the early semantics for relevance logic. The semantic clause for the conditional from relevance logic is:

$$(16) \quad x \Vdash A \rightarrow B \text{ iff for all } y, z \in \mathbf{F} \text{ s.t. } Rxyz, \text{ if } y \Vdash A \text{ then } z \Vdash B$$

$Rxyz$  is, by itself, simply an abstract mathematical entity. One way or reading it, the way that became popular in relevance logic circles, is

$Rxyz$  iff the result of combining  $x$  with  $y$  is true at  $z$ .

Given that the points of evaluation in relevance logics were understood originally as impossible situations (since they may be both inconsistent and incomplete), the main conceptual move was to understand channels to be special types of situations. The full proofs may be found in Restall (1994) and Mares (1996), and these demonstrate that the expressive power of Barwise's system may be captured by the frame semantics of relevance logic. What it is that such “combining” of  $x$  and  $y$  amounts to depends on, of course, which structural rules are operating on the frame in question. As explained in the previous section above, the choice of which rules to include will depend on the properties of the phenomena being modelled.

The final step required for locating the meeting point between information as code and information as correlation is as follows. Contemporary approaches to relevance and other substructural logics understand the points of evaluation (impossible situations) to be information states. There is certainly no constraint on information that it be complete or consistent, so the expressibility of impossible situations is not sacrificed. Such an informational reading (Paoli 2002; Restall 2000; Mares 2004) lends itself to multiple applications of various substructural frameworks, and also

does away with the ontological baggage brought by questions like “what are impossible situations?” in the “What are possible worlds?” spirit. An information-state reading of  $Rxyz$  will be something like

the result of combining the information carried by  $x$  and  $y$  generates the informations carried by  $z$ .

Making this explicit results in  $Rxyz$  being written down as  $x \bullet y \sqsubseteq z$ , in which case (15) is, via (16), equivalent to (9).

An important structural rule for the composition operation on information channels, that is, on information states that carry information of conditional types, is that it is associative. What this means is that:

$$(17) \quad z \xrightarrow{x \bullet (y \bullet v)} w = z \xrightarrow{(x \bullet y) \bullet v} w$$

Where  $z \Vdash A$  and  $w \Vdash D$ , this will be the case for all  $x, y, v$  s.t.  $x \Vdash A \rightarrow B$ ,  $y \Vdash B \rightarrow C$ ,  $v \Vdash C \rightarrow D$ . This is just the first step required to demonstrate that channel theory, and its underlying substructural logic, form a *category*.

Category theory is an extremely powerful tool in its own right. For a thorough introduction see Awodey (2006). For more work on the relationship between various substructural logics and channel theory, see Restall (1994a, 1997, 2006). Further category-theoretic work on information flow may be found in Goguen (2004—see Other Internet Resources). Recent important work on category-theoretic frameworks for information flow that extend to *quantifiable/probabilistic* frameworks is due to Seligman (2009). Perhaps the most in depth treatment of information flow in category theoretic terms is to be found in the work of Samson Abramsky, and an excellent overview may be found in his “Information, Processes, and Games” (2008).

### 4.3 Code and ranges

Logical frameworks that model information as code and range along with information about encoding have been developed by Velázquez-Quesada (2009), Liu (2009), Jago (2006), and others. The key element to all of these approaches is the introduction of some syntactic code to the conceptual architecture of the information as range approach.

Taking Velázquez-Quesada (2009) as a working example, start with a *modal-access model*  $M = \langle S, R, V, Y, Z \rangle$  where  $\langle S, R, V \rangle$  is a Kripke Model,  $Y$  is the *access set function* and  $Z$  is the *rule set function* s.t. (where  $I$  is the set of classical propositional language based on a set of atomic propositions):

- $Y : W \rightarrow \wp(I)$  assigns a set of formulas of  $I$  to each  $x \in S$ .
- $Z : W \rightarrow \wp(R)$  assigns a set of rules based on  $I$  to each  $x \in S$ .

A modal-access model is a member of the class of modal access models **MA** iff it satisfies truth for formulas and truth preservation for rules. **MA<sub>k</sub>** models are those **MA** models such that  $R$  is an equivalence relation.

From here, inference is represented as a modal operation adding the rule's conclusion to the access set of information states of the of the agent such that the agent can access both the rule and its premises. Where  $Y(x)$  is the access set at  $x$ , and  $Z(x)$  is the rule set at  $x$ :

- **Inference on knowledge:** Where  $M = \langle S, R, V, Y, Z \rangle \in \mathbf{MA}_k$ , and  $\sigma$  is a rule,  $M_k\sigma = \langle S, R, V, Y', Z \rangle$  differs from  $M$  in  $Y'$ , given by  $Y'(x) := Y(x) \cup \{\text{conc}(\sigma)\}$  if  $\text{prem}(\sigma) \subseteq Y(x)$  and  $\sigma \in Z(x)$ , and by  $Y'(x) := Y(x)$  otherwise.

The dynamic logic for inference on knowledge then incorporates the ability to represent “*there is a knowledge inference with  $\sigma$  after which  $\phi$* ”

*holds*” (Velázquez-Quesada 2009). It is in just this sense that such modal information theoretical approaches model the outputs of inferential processes, as opposed to the properties of the inferential processes that generate such outputs (see the section on *categorical information theory* for models of such dynamic properties).

Jago (2009) proposes a rather different approach based upon the elimination of worlds *considered* possible by the agent as the agent reasons deductively. Such epistemic (doxastic) possibilities structure an epistemic (doxastic) space under bounded rationality. The connection with information as code is that the modal space is individuated syntactically, with the worlds corresponding to possible results of step-wise rule-governed inferences. The connection with information as range is that the rules that the agent does or does not have access to will impact upon the range of discrimination for the agent. For example, if the agent's epistemic-base contains two worlds, a  $\neg\phi$  world and a  $\phi \vee \psi$  world say, then can refine their epistemic base only if they have access to the disjunctive syllogism rule.

A subtle but important contribution of Jago's is the following: the modal space in question will contain only those epistemic options which are not *obviously* impossible. However, what is or is not obviously impossible will vary from both agent to agent, as well as for a single agent over time as that agent refines its logical acumen. This being the case, the modal space in question has *fuzzy* boundaries.

## 5. Special topics

There is a varied list of special topics pertaining to the logical approach to information. This section briefly illustrates just a couple of them, which are important regardless of the particular stance one takes (information as range, as correlation, as code). The first topic is the issue of informational

equivalence: when are two structures in the logical approach one is using indistinguishable in terms of the information they are meant to encode, convey, or carry? The second topic concerns the various ways in which the idea of negative information can be understood conceptually, and properly dealt with formally.

### 5.1 Information Structures and Equivalence

Every logical approach to information comes with its own kind of information structures. Depending on the particular stance and the aspect of information to be stressed, these structures may stand for informational states, structured syntactic representations, pieces of information understood as commodities, or global structures made up from local interrelated informational states or stages. Under which conditions can two informational structures be considered to be informationally equivalent?

Addressing this question brings out the need to have it clear at which level of granularity one is testing for equivalence. Take for example the classical extensional notion of logical equivalence. This is a coarse equivalence, in that informationally different claims such as *2 is even* and *2 is prime* cannot be distinguished, as their extensions will coincide. Take instead an equivalence given by identity at the level of representations (say syntactic equality). This is on the contrary too fine-grained in some cases: to a bilingual speaker, the information that the shop is closed would be equally conveyed by a sign saying “Closed” as by a sign saying “*Geschlossen*”, even if the two words are different.

An intermediate notion of equivalence that has proved central to the range, correlational, and code takes on information is the relation of bisimulation between structures. A bisimulation relation between two graphs  $G$  and  $H$  (where both the arrows and nodes of the graphs are labelled) is a binary

relation  $R$  between the nodes of the graphs with the property that whenever a node  $g$  of  $G$  is related to a node  $h$  of  $H$ , then:

1.  $g$  and  $h$  have the same labels, and
2. For every relation label  $L$  and every  $L$ -child of  $g'$  of  $g$ , there must be a  $L$ -child  $h'$  of  $h$  such that  $h$  and  $h'$  are related by  $R$ . A similar condition must hold for every  $L$ -child of  $h$ .

A simple example would be the relation between the following two graphs (empty set of labels) that relates the point  $x$  with  $a$  and the point  $y$  with the points  $b, c, d$ .

$$x \rightarrow y \quad a \rightarrow b \rightarrow c \rightarrow d$$

Bisimulation is a central notion from the information-as-range perspective because the Kripke models of section 1 are precisely labelled graphs. It is a classical result of modal logic that if two states of two models are related by a bisimulation, then the states will satisfy exactly the same modal formulas, and in addition a first order property of states is definable in the basic modal language if and only if the property is preserved under bisimulation.

In the correlational stance, in situation theory, bisimulation turns out to be the right notion in determining whether two infons that might look structurally different actually are the same as pieces of information. For example, the analysis of Liar-like claims leads to infons that are nested in themselves, such as

$$\sigma = \langle \langle \text{True, what} : \sigma, 0 \rangle \rangle.$$

One can naturally depict the structure of  $\sigma$  as a labelled graph, which will be bisimilar to the graph associated with the apparently different infon

$$\psi = \langle \langle \text{True, what} : \langle \langle \text{True, what} : \psi, 0 \rangle \rangle, 0 \rangle \rangle.$$

The notion of bisimulation appeared independently in computer science, so it is no surprise that it also shows in issues closest to the information-as-code approach, with its focus on representation and computation. Several versions of bisimulation have been applied to classes of automata to determine when two of them are behaviourally equivalent, and data encodings such as

$$L = \langle 0, L \rangle \text{ and } L = \langle 0, \langle 0, L \rangle \rangle,$$

both of which represent the same object (an infinite list of zeroes) can be identified as such by noticing that the graphs that depict the structure of these two expressions are bisimilar. See Aczel (1988), Barwise and Moss (1996), and Moss (2009) for more information about bisimulation and its connections with modal logic, data structures, and coalgebras.

## 5.2 Negative information

This entry has focused mostly on *positive information*. Formally speaking, *negative information* is simply the extension-via-negation of the positive fragment of any logic built around information-states. Different negation-types will constrain the behaviour of negative information in various ways. Informally, negative information may be thought of variously as what is canonically expressed with sentential negation, process exclusion (both propositional and sub-propositional) and more. Even when we restrict ourselves to a single conceptual notion, there may be vigorous philosophical debate as to which formal construction best captures the notion in question. In this section, we run through several formal analyses of negative information, we examine some of the philosophical debates surrounding the suitability of various formal constructions with respect to particular applications, and examine the related topic of failure of



information flow in the situation-theoretic sense, which may give rise to misinformation or lack of information in particular settings.

Non-constructive intuitionistic negation, is aimed towards accounting for negative information in the context of information flow via observation. For more details on this point, see the subsection intuitionistic logics and Beth and Kripke models in the supplementary document: Abstract Approaches to Information Structure.

Working with the frames from section 3.1, non-constructive intuitionistic negation is defined in terms of the constructive implication, (21), which is combined with bottom,  $\mathbf{0}$ , which holds nowhere, as specified by its frame condition:

$$(18) \ x \Vdash \mathbf{0} \text{ for no } x \in \mathbf{F}$$

Hence intuitionistic negation is defined as follows:

$$(19) \ -A := A \rightarrow \mathbf{0}$$

Hence the frame condition for  $-A$  is as follows:

$$(20) \ x \Vdash -A [A \rightarrow \mathbf{0}] \text{ iff for all } y \in \mathbf{F}, \text{ s.t. } x \sqsubseteq y, \text{ if } y \Vdash A \text{ then } y \Vdash \mathbf{0}$$

(20) states that if  $x$  carries the information that  $-A$ , then there no state  $y$  such that  $y$  is an informational development of  $x$  where  $y$  carries the information that  $A$ .

The definition of  $-A$  in terms of  $A \rightarrow \mathbf{0}$  throws up an asymmetry between positive and negative information. In an information model  $-A$  holds at  $x \in \mathbf{F}$  iff  $A$  does not hold at any  $y \in \mathbf{F}$  such that  $x \sqsubseteq y$ . Whilst the verification of  $A$  at  $x \in \mathbf{F}$  only involves checking  $x$ , verifying  $-A$  at  $x \in \mathbf{F}$  involves checking all  $y \in \mathbf{F}$  such that  $x \sqsubseteq y$ . According to Gurevich (1977)

and Wansing (1993), this asymmetry means that intuitionistic logic does not provide an adequate treatment of negative information, since, unlike the verification of  $A$ , there is no way of verifying  $-A$  “on the spot” so to speak. Gurevich and Wansing's objection to this asymmetry is a critical response to Grzegorzczuk (1964). For arguments in support of Grzegorzczuk's asymmetry between positive and negative information, see Sequoiah-Grayson (2009). A fully constructive negation that allows for falsification “on the spot” is known also as *Nelson Negation* on account of it being embedded within Nelson's constructive systems (Nelson 1949, 1959). For a contemporary development of these constructive systems, see section 2.4.1 of Wansing (1993).

In a static logic setting, negation is, at the very least, used to rule out truth (if not to express explicit falsity). In a dynamic setting, negation will be used to rule out particular *processes*. For a development negative information as process exclusion in the context of categorial information theory see Sequoiah-Grayson (2013). This idea has its origins in the Dynamic Predicate Logic of Groenendijk and Stokhof (1991), in particular with their development of negative information via negation as *test-failure*. For an exploration between the conceptions of negative information as process exclusion and test-failure, see Sequoiah-Grayson (2010).

In any logic for negation as process-exclusion, the process-exclusion will be *non-directional* if the logic in question is commutative. Directional process-exclusion will result when we remove the structural rule of commutation. For a discussion of the relationship between the formalisation of directional process exclusion as commutation-failure along with symmetry-failure on compatibility and incompatibility relations on information states, see Sequoiah-Grayson (2011). For an extended discussion of negative information in the context of categorial grammars, see Buszkowski (1995).

## 6. Conclusion

There is a bi-directional relation between logic and information. On the one hand, information underlies the intuitive understanding of standard logical notions such as inference (commonly thought of as the process that turns implicit information into explicit one) and computation. On the other hand, logic provides a formal framework for the study of the notion of information.

The logical study of information reveals and studies some of the most fundamental *qualitative* aspects of information. Different stances on information naturally highlight some of these aspects more than others. Thus, the information-as-range stance most naturally highlights *agency* and the *dynamics of information* in settings with multiple agents that can *interact* with each other. The *aboutness of information* (information is always about something) is a central theme in the information-as-correlation stance, as it is the fact that *information is situated* (what information is carried by two instances of the same signal may have different informational content). And the topic of *encoding* of information and its *processing* (as in the case of formal inference) is at the core of the information-as-code stance. None of these qualitative aspects of information is exclusive to just one of the stances, even if some stress some topics more than others, and in fact some themes such as the *structure of information and its relation with information content* are equally pertinent regardless of the stance. The view of information studied in this entry thus differs from other important formal frameworks that study *quantitative* notions of information. For example, Shannon's statistical theory of information is concerned with things such as optimizing the amount of data that can be transmitted via a noisy channel, and the Kolmogorov's complexity theory quantifies the informational complexity of a string as the length of the shortest program that outputs it when executed by a fixed universal Turing machine.

The contributions of the logical approach to information include fruitful reinterpretations of known logical systems (such as epistemic logic or relevance logic), and also new systems that result from elaborating on these so to capture further informational aspects, or from combining aspects of two different stances, as the constraint systems of section 4. New frameworks (situation theory in the 80s) have also resulted from exploring from scratch what sorts of inference—maybe new and non-classical—one should allow in order to model certain aspects of information.

Looking for *interfaces* between the three stances is a nascent direction of inquiry, discussed here in section 4. A complementary issue is whether the stances can be *unified*. There are several formal frameworks that, beyond serving as potential settings for exploring the issue of unification, are abstract mathematical theories of information in their own right. Each of these goes well beyond the scope of this entry, so they are just listed next.

- Domain Theory (Abramsky and Jung 1994): it has been used to study the processes of unraveling or “improvement” of informational states in terms of partial orderings of information states that naturally arise across the stances.
- Point-free topology: it has deep connections with computer science and it can actually be motivated as a logic of information (Vickers 1996).
- Chu Spaces (Pratt 1999): in category theory they are presented as generalizations of topologies. The immediate link with things discussed in this entry is that the classifications used in situation theory are simply Chu spaces, discovered independently and with different aims.
- Coalgebra: another branch of category theory that has also been presented as the “mathematics of sets and observations” (Jacobs 2012, Other Internet Resources). This framework has strong links

with many notions discussed in this entry, in particular modal logic (section 1) and bisimulation (section 5.1).

- Probability Theory: it is clearly at the center of abstract *quantitative* approaches to information. Various versions of the inverse relationship principle that lead to measures of *semantic* information (see section 1.3 and Floridi 2013) descend from the version used by Shannon (1950): in a communication setting via noisy channels, the less expected a received message is, the more informative it is.

The logical study of information resembles in spirit other more traditional endeavours, such as the logical study of the concept of truth or computation: in all these cases the object of logical study plays a central role in the intuitive understanding of logic itself. The three perspectives on qualitative information presented in this entry (ranges, correlations, and code) portrait the diverse state of the art in this field, where many directions of research lay open, both on the way of looking for unifying or interfacing settings for the different stances, and of deepening the understanding of the main qualitative features of information (dynamics, aboutness, encoding, interaction, etc.) within each stance.

Finally, interested readers may wish to pursue the topics in the supplementary document

#### Abstract Approaches to Information Structure

which covers the topics intuitionistic logic, Beth and Kripke models, and algebraic and other approaches to modal information theory and related areas.

## Bibliography

Abramsky, S., 2008, “Information, Processes, and Games”, in Adriaans and van Benthem 2008, 483–550.

- Abramsky, S. and A. Jung, 1994, “Domain Theory”, in *Handbook of Logic in Computer Science*, S. Abramsky, D. Gabbat, and T. S. E. Maibaum (eds.), Oxford: Oxford University Press, pp. 1–168.
- Aczel, P., 1988, *Non-Well-Founded Sets*, (CSLI Lecture Notes 14), Stanford: CSLI Publications.
- Adriaans, P. and J. F. A. K. van Benthem (eds.), 2008, *Philosophy of Information* volume 8 of *Handbook of the Philosophy of Science*, Amsterdam: North Holland.
- Allo, P., 2009, “Reasoning about Data and Information”, *Synthese*, 167: 231–249.
- Artemov, S., 2008, “The Logic of Justification”, *Review of Symbolic Logic*, 1(4): 477–513.
- Artemov, S. and M. Fitting, 2012, “Justification Logic”, *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2012/entries/logic-justification/>
- Awodey, S., 2006, *Category Theory*, Oxford Logic Guides 49, Oxford: Clarendon Press.
- Baltag, A., B. Coecke, and M. Sadrzadeh, 2007, “Epistemic Actions as Resources”, *Journal of Logic and Computation*, 17(3): 555–585.
- Baltag A., H. P. van Ditmarsch, and L. S. Moss, 2008, “Epistemic Logic and Information Update”, in Adriaans and van Benthem, 2008, p. 361–456.
- Baltag, A. and S. Smets, 2008, “A Qualitative Theory of Dynamic Interactive Belief Revision”, in *Logic and the Foundation of Game and Decision Theory* (LOFT7), G. Bonanno, W. van der Hoek, and M. Wooldridge (Eds.), Volume 3 of *Texts in Logic and Games*, pp. 13–60. Amsterdam: Amsterdam University Press.
- Bar-Hillel, Y. and R. Carnap, 1952, “An Outline of a Theory of Semantic Information”, Technical Report No. 247, Research Laboratory of

- Electronics, MIT. Reprinted in *Language and Information: Selected Essays on their Theory and Application*, Y. Bar-Hillel, Addison-Wesley Series in Logic, Israel: Jerusalem Academic Press and Addison-Wesley, 1964, pp. 221–74.
- Barwise, J., 1988, “Three views of common knowledge”, *TARK '88 Proceedings of the 2nd conference on Theoretical aspects of reasoning about knowledge*, San Francisco: Morgan Kaufmann, p. 365–379.
- , 1993, “Constraints, Channels, and the Flow of Information”, in *Situation Theory and its Applications*, 3, (CSLI Lecture Notes 37), Aczel et al. (eds.), Stanford: CSLI Publications.
- , 1997, “Information and Impossibilities”, *Notre Dame Journal of Formal Logic* 38(4): 488–515.
- Barwise, J. and J. Etchemendy, 1987, *The Liar*, Oxford: Oxford University Press.
- Barwise, J. and L. Moss, 1996, *Vicious Circles*, (CSLI Lecture Notes 60), Stanford: CSLI Publication.
- Barwise, J. and J. Perry, 1983, *Situations and Attitudes*, Cambridge, MA: MIT Press.
- , 1985, “Shifting Situations and Shaken Attitudes”, *Linguistics and Philosophy*, 8: 105–161.
- Barwise J. and J. Seligman, 1997, *Information Flow: The Logic of Distributed Systems*, Cambridge Tracts in Theoretical Computer Science 44, New York: Cambridge University Press.
- Beal, J. C. and G. Restall, 2006, *Logical Pluralism*, Oxford: Clarendon Press.
- van Benthem, J., 1995, *Languages in Action: Categories, Lambdas, and Dynamic Logic*, Cambridge, MA: MIT Press.
- , 2000, “Information Transfer Across Chu Spaces”, *Logic Journal of the IGPL*, 8(6): 719–731.
- , 2003, “Logic and Dynamics of Information”, *Minds and Machines*,

- 13(4): 503–519.
- , 2004, “Dynamic Logic for Belief Revision”, *Journal of Applied Non-Classical Logics*, 14(2): 129–155.
- , 2009, “The Information in Intuitionistic Logic”, *Synthese*, 167: 251–270.
- , 2010, “Categorical versus Modal Information Theory”, *Linguistic Analysis*, 36: 533
- , 2011, *Logical Dynamics of Information and Interaction*, Cambridge: Cambridge University Press.
- van Benthem, J., J. van Eijck, and B. Kooi, 2006, “Logics of Communication and Change”, *Information and Computation*, 204(11): 1620–1662.
- van Benthem, J. and M. Martinez, 2008, “The stories of logic and information”. in *Philosophy of Information*, in Adriaans and van Benthem, 2008, p. 217–280 .
- Beth, E. W., 1955, “Semantic Entailment and Formal Derivability”, *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings of the Section of Sciences*, 18: 309–342).
- , 1956, “Semantic Construction of Intuitionistic Logic”, *Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings of the Section of Sciences*, 19: 357–388.
- Blackburn, P., M. de Rijke, and Y. Venema, 2001, *Modal Logic*, Cambridge tracts in theoretical computer science 53, Cambridge: Cambridge University Press.
- Buszkowski, W., 1995, “Categorical Grammars with Negative Information”, in *Negation, A Notion in Focus*, H. Wansing (ed.), Berlin: Gruyter, pp. 107–126.
- D'Agostino, M. and L. Floridi, 2009, “The Enduring Scandal of Deduction”, *Synthese* 167: 317–315.
- Devlin, K., 1991, *Logic and Information*, Cambridge: Cambridge University Press.

- van Ditmarsh, H., W. van der Hoek, and B. Kooi, 2008, *Dynamic Epistemic Logic*, The Netherlands: Springer.
- Dretske, F., 1981, *Knowledge and the flow of information*, Cambridge: Cambridge University Press.
- Dunn, J. M., 1993, “Partial Gaggles Applied to Logics with Restricted Structural Rules”, in *Substructural Logics*, P. Schroeder-Heister and K. Dosen (eds.), Oxford: Oxford Science Publications, Clarendon Press, 63–108.
- Duzi, M., 2010, “The Paradox of Inference and the Non-Triviality of Analytic Information”, *Journal of Philosophical Logic* 38(5): 473–510.
- Duzi, M., B. Jespersen, and P. Materna, 2010, *Procedural Semantics for Hyperintensional Logic (Foundations and Applications of TIL)*, Logic, Epistemology, and the Unity of Science, vol. 17, Springer.
- Dyckhoff, R. and M. Sadrzadeh, 2010, “A Cut-Free Sequent Calculus for Algebraic Dynamic Epistemic Logic”, Computer Science Research Report, University of Oxford, CS-RR-10-11.
- van Eijck, J. and A. Visser, 2012, “Dynamic Semantics”, *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2012/entries/dynamic-semantics/>.
- Fagin R., J. Halpern, and M. Vardi, 1995, “Reasoning about Knowledge”. Cambridge, MA: MIT Press.
- Floridi, L., 2004, “Outline of a Theory of Strongly Semantic Information”, *Minds and Machines*, 14(2): 197–22.
- , 2006, “The Logic of Being Informed”, *Logique et Analyse*, 49(196): 433–460.
- , 2013, “Semantic Conceptions of Information”, *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL =

- <https://plato.stanford.edu/archives/spr2013/entries/information-semantic/>.
- Gabbay, D. M., 1993, “Labelled Deductive Systems: A Position Paper”, in *Logic Colloquium '90: ASL Summer Meeting in Helsinki*, J. Oikkonen and J. Vaananen (eds), Berlin: Springer-Verlag, 66–88.
- , 1996, *Labelled Deductive Systems: Volume 1*, Oxford Logic Guides 35, New York: Oxford University Press.
- Ganter, B. and R. Wille, 1999, *Formal Concept Analysis, Foundations and Applications*. LNCS 3626, Springer.
- Ghidini, C. and F. Giunchiglia, 2001, “Local Model Semantics, or Contextual Reasoning = Locality + Compatibility”. *Artificial Intelligence*, 127: 221–259.
- Groenendijk, J. and M. Stokoff, 1991, “Dynamic Predicate Logic”, *Linguistics and Philosophy*, 14: 33–100.
- Gurevich, Y., 1977, “Intuitionistic Logic with Strong Negation”, *Studia Logica*, 36: 49–59.
- Grzegorzczak, A., 1964, “A Philosophically Plausible Interpretation of Intuitionistic Logic”, *Indagationes Mathematicae*, 26: 596–601.
- Hintikka, J., 1970, “Surface Information and Depth Information”, in *Information and Inference*, J. Hintikka and P. Suppes (eds.), Dordrecht: Reidel, 263–97.
- , 1973, *Logic, Language Games, and Information*, Oxford: Clarendon Press.
- , 2007, *Socratic Epistemology: Explorations of Knowledge—Seeking by Questions*, Cambridge: Cambridge University Press.
- Israel, D. and J. Perry, 1990, “What is information?”, in *Information, Language and Cognition*, P. Hanson, (ed.), University of British Columbia.
- , 1991, “Information and Architecture, in *Situation Theory and its Applications* Vol 2, J. Barwise, J.M. Gawron, G. Plotkin, and S. Tutiya, (eds.), Stanford: CSLI Publications.

- Jago, M., 2006, “Rule-based and Resource-Bounded: A New Look at Epistemic Logic”, in *Proceedings on the Workshop on Logics for Resource—Bounded Agents, as part of ESSLLI 2006*, T. Agotnes and N. Alechina (eds.), pp. 63–77, 2006.
- , 2009, “Logical Information and Epistemic Space”, *Synthese*, 167: 327–341.
- King, J. C., 2012, “Structured Propositions”, *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2012/entries/propositions-structured/>.
- Kratzer, A., 2011, “Situations in Natural Language Semantics”, *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2011/entries/situations-semantics/>.
- Kripke, S. A., 1963, “Semantical Analysis of Modal Logic”, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9: 67–96.
- , 1965, “Semantical Analysis of Intuitionistic Logic I”, in *Formal Systems and Recursive Functions*, J. Crossley and M. Dummett (eds.), Amsterdam: North Holland, 92–129.
- Lambek, J., 1958, “The Mathematics of Sentence Structure”, *American Mathematical Monthly*, 65: 154–170.
- , 1961, On the Calculus of Syntactic Types, in *Structure of Language and its Mathematical Aspects*, R. Jakobson (ed.), Providence: American Mathematical Society, 166–178.
- Lewis, D., 1969, *Convention: A Philosophical Study*, Cambridge: Harvard University Press.
- Liu, F., 2009, “Diversity of Agents and Their Interaction”, *Journal of Logic, Language, and Information*, 18(1): 23–53.
- Mares, E., 1996, “Relevant Logic and the Theory of Information”, *Synthese* 109: 345–370.





- , 2004, *Relevant Logic: A Philosophical Interpretation*, Cambridge: Cambridge University Press.
- , 2009, “General Information in Relevant Logic”, *Synthese*, 167: 343–362.
- McGrath, M., 2012, “Propositions”, *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2012/entries/propositions/>.
- Moss, L. S., 2009, “Non-wellfounded Set Theory”, *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2009/entries/nonwellfounded-set-theory/>.
- Moss, L. and J. Seligman, 1996, “Situation Theory”, in *Handbook of Logic and Language*, J. van Benthem and A. ter Meulen, (eds.), Amsterdam: Elsevier.
- Nelson, D., 1949, “Constructible Falsity”, *Journal of Philosophical Logic*, 14: 16–26.
- , 1959, “Negation and the Separation of Concepts in Constructive Systems”, in *Constructivity in Mathematics*, A. Heyting (ed.), Amsterdam: North-Holland, 208–255.
- Pacuit, E., 2011, “Logics of Informational Attitudes and Informative Actions”, *Journal of the Council of Indian Philosophy*, Vol. XXVII, No. 2.
- Paoli, F., 2002, *Substructural Logics: A Primer*, Kluwer Academic Publishers.
- Pratt, V.R., 1999, “Chu Spaces”, *Notes for School on Category Theory and Applications*, University of Coimbra, Portugal, July 13–17.
- Primiero, G., 2006, “An Epistemic Constructive Definition of Information”, *Logique et Analyse*, 50(200): 391–416.
- , 2008, *Information and Knowledge: A Constructive Type-Theoretical Approach*, Logic, Epistemology, and the Unity of Science Series, vol.

- 10, Springer.
- Restall, G., 1994, "Information Flow and Relevant Logics", in *Logic, Language, and Computation*, Jerry Seligman and Dag Westeråhl (eds.), Stanford: CSLI Publications, 1995, 139–160.
- , 1994a, "A Useful Substructural Logic", *Bulletin of the Interest Group in Pure and Applied Logics*, 2: 137–148.
- , 1997, "Ways Things Can't Be", *Notre Dame Journal of Formal Logic*, 38: 583–596.
- , 2000, *Substructural Logics, and Introduction*, London: Routledge.
- , 2006, "Logics, Situations, and Channels", *Journal of Cognitive Science*, 6: 125–150.
- Sadrzadeh, M., 2009, "Ockham's Razor for Reasoning about Information Flow", *Synthese*, 167: 391–408.
- Segerberg, K., 1998, "Irrevocable Belief Revision in Dynamic Doxastic Logic", *Notre Dame Journal of Formal Logic*, 39(3): 287–306.
- Seligman, J., 1990, "Perspectives in Situation Theory, in *Situation Theory and its Applications*, Vol 1, R. Cooper, K. Mukai, and J. Perry, (eds.), Stanford: CSLI Publications.
- , 2009, "Channels: From Logic to Probability", in *Formal Theories of Information: From Shannon to Semantic Information Theory and General Concepts of Information*, G. Sommaruga (ed.), LNCS 5363, Berlin: Springer Verlag, 193–233.
- Sequoiah-Grayson, S., 2007, "The Metaphilosophy of Information", *Minds and Machines*, 17: 331–44.
- , 2008, "The Scandal of Deduction: Hintikka on the Information Yield of Deductive Inferences", *Journal of Philosophical Logic*, 37: 67–94.
- , 2009, "Dynamic Negation and Negative Information", *Review of Symbolic Logic* 2(1): 233–248.
- , 2010, "Lambek Calculi with 0 and Test-Failure in DPL", *Linguistic Analysis*, 36: 517–532.
- , 2011, "Non-Symmetric (In)Compatibility Relations and Non-

- Commuting Types", *The Logica Yearbook 2010*, Michael Peliš and Vit Punčochář (eds.) London: College Publications.
- , 2013, "Epistemic Closure and Commuting, Nonassociating Residuated Structures", *Synthese*, 190(1): 113–128
- Shannon, C. E., 1948, "A Mathematical Theory of Communication", *Bell System Technical Journal* 27: 379–423 and 623–656.
- , 1950, "The Lattice Theory of Information", in *Report of Proceedings, Symposium on Information Theory, London, Sept., 1950*, Institute of Radio Engineers, Transactions on Information Theory, No. 1 (February, 1953), pp. 105–107. Reprinted in *Claude Elwood Shannon Collected Papers*, N. J. A. Sloan and A. D. Wyner (eds.), Los Alamos, CA: IEEE Computer Science Press, 1993.
- Stefaneas P. and Vandoulakis, I. M., forthcoming, "Proofs as Spatio-temporal Processes", in *Proceedings of the 14th Congress of Logic, Methodology and Philosophy of Science*, P. E. Bour, G. Heinzmann, W. Hodges and P. Schroeder-Heister (eds.). To appear in *Philosophia Scientiae*, 19(1).
- Textor, M., 2012, "States of Affairs", *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2012/entries/states-of-affairs/>>.
- Troelstra, A. S., 1992, *Lecture Notes on Linear Logic* (CSLI Lecture Notes 29), Stanford: CSLI Publications.
- Vanderschraaf, P. and G. Sillari, 2009, "Common Knowledge", *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2009/entries/common-knowledge/>>.
- Velázquez-Quesada, F. R., 2009, "Dynamic Logics for Explicit and Implicit Information", in Xiangdong He and John F. Horty and Eric Pacuit (eds.), *Logic, Rationality, and Interaction: Second*

- International Workshop, LORI 2009*, Chongqing, China, October 8–11, 2009, Berlin: Springer, pp. 325–326.
- Vickers, S., 1996, *Topology via Logic*, Cambridge: Cambridge University Press.
- Wansing, H., 1993, *The Logic of Information Structures*, (Lecture Notes in Artificial Intelligence no. 681, Subseries of Lecture Notes in Computer Science), Berlin: Springer-Verlag.
- Zalta, Edward N., 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel.
- , 1993, “Twenty-Five Basic Theorems in Situation and World Theory”, *Journal of Philosophical Logic*, 22(4): 385–428.

## Academic Tools

-  How to cite this entry.
-  Preview the PDF version of this entry at the Friends of the SEP Society.
-  Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
-  Enhanced bibliography for this entry at PhilPapers, with links to its database.

## Other Internet Resources

- Goguen, J., 2004, “Information Integration in Institutions,” online manuscript.
- Jacobs, B., 2012 *Introduction to Coalgebra. Towards Mathematics of States and Observations*, online manuscript, Version 2.0.

## Related Entries

common knowledge | information | information: semantic conceptions of | logic: dynamic epistemic | propositions: structured | set theory: non-wellfounded | situations: in natural language semantics | states of affairs

## Acknowledgments

The authors would like to extend their thanks to the Editors of the Stanford Encyclopaedia of Philosophy, as well as to Johan van Benthem, Olivier Roy, and Eric Pacuit. Their assistance and advice has been invaluable.

## Abstract Approaches to Information Structure

The material in the section constitutes some major instances very abstract approaches to information structure. The abstractness means that many connections with topics from the previous sections can be made, with no one connection being exclusive enough to warrant the material be placed in any one section as apposed to some other.

## Intuitionistic logic, Beth and Kripke models

Consider again the waiter-at-the-cafe example from section 1. We noted that it was the verbal interaction between several agents that facilitated the information flow that enabled the logical reasoning to be undertaken. Along with verbal interactions, there was another type of epistemic action crucial for execution of the dynamic reasoning procedure undertaken by the waiter, *observation*. Simply put, the waiter observed the coffees in his possession (and presumable the ticket with the table number and so forth), and these observations facilitated information flow between the waiter and the waiter's physical environment, and this information flow too played a necessary role in enabling the waiter's actions.



Beth models for constructive logic are needed when modelling the information update from observations in order that the primeness of disjunction be circumvented (see the following paragraph for an example of why it is that we should want to circumvent it). A disjunction operation is said to be *prime* when it obeys the conditions specified by (23) below. To see why this is a problem, consider the standard Kripke models for constructive logic:

(21)  $x \Vdash A \rightarrow B$  iff for all  $y \in \mathbf{F}$ , s.t.  $x \sqsubseteq y$ , if  $y \Vdash A$  then  $y \Vdash B$

(22)  $x \Vdash A \wedge B$  iff  $x \Vdash A$  and  $x \Vdash B$

(23)  $x \Vdash A \vee B$  iff  $x \Vdash A$  or  $x \Vdash B$

Take the states of  $S$  to be stages of a scientific research project. Now consider (21). (21) makes sense on this interpretation. It means simply that  $A \rightarrow B$  has been established at a particular stage  $x$  in a research project iff at any future stage  $y$  in our project, if  $\alpha$  has established  $A$ , then  $\alpha$  has established  $B$ . (22) makes sense under this interpretation, however (23) does not. (23) does not make sense because knowledge is not *prime* in the sense of (23). Clearly,  $K_\alpha(A \vee B)$  does *not* imply that *either*  $K_\alpha A$  or  $K_\alpha B$  (see Beal and Restall 2006). In order to take care of this, we need a frame construction with different disjunction conditions.

The original (and arguably most straightforward) way to take care of this is via *Beth frames* (Beth 1955, 1956). The conceptual motivation behind Beth's disjunction condition is straightforward. Beth's disjunction condition states that  $x \Vdash A \vee B$  iff either  $A$  or  $B$  *will* be established at some stage, where this stage is an informational development of  $x$ . In other words, if  $K_\alpha(A \vee B)$ , then  $\alpha$  will establish  $A$  or establish  $B$  at a future stage of the research project. Note that this assumes logical omniscience on behalf of  $\alpha$ . To get off the ground formally, we need to introduce (information-)chains and (information-)bars:

- **Chains:** Take an information frame  $\mathbf{F} := \langle S, \sqsubseteq \rangle$ . A subset  $s$  of this partial order  $\sqsubseteq$  will be *totally ordered* iff for all  $x, y \in s$ , either  $x \sqsubseteq y$  or  $y \sqsubseteq x$ .  $s$  will be *maximally totally ordered* iff  $s$  is totally ordered, and there is no  $y \in \mathbf{F}$  such that  $y \notin s$  and  $s \cup \{y\}$  is a total order also.  $s$  is a **chain** when  $s$  is a maximal total order.
- **Bars:** A subset  $s' \subset S$  **bars** an information state  $x$  iff every chain  $s$  through  $x$  intersects  $s'$ .

The Beth model for disjunction (where  $\llbracket \_ \rrbracket$  is an evaluation function) is as follows:

(24)  $x \Vdash A \vee B$  iff  $\llbracket A \rrbracket \cup \llbracket B \rrbracket$  bars  $x$

That is, for any possible course of research through  $x$  (i.e., on any chain through  $x$ ) either  $A$  will be established or  $B$  will be established. The other connectives remain the same, with the caveat the propositions are closed under barring. That is, we get  $x \in \llbracket A \rrbracket$  whenever we have it for some  $x$  that  $\llbracket A \rrbracket$  bars  $x$ . This makes sense on our interpretation, since if  $\alpha$  knows that  $A$  will be established at *any* stage of research extending the present stage, then  $\alpha$  will have established  $A$  at the present stage.

Considerable work has been done in order to reveal the connections between modal information theory and constructive logic systems by van Benthem (2009). van Benthem begins by taking an informationalised model condition for the universal modality to be the following:

(25)  $x \Vdash \Box \phi$  iff  $\phi$  is true at all future states  $y \sqsupseteq x$

From here van Benthem may define constructive implication with the informationalised universal modal operator as  $\Box (\phi \rightarrow \psi)$ . Atomic propositions are upwardly preserved due to the addition of the atomic

persistence law  $p \rightarrow \Box p$ . The existential modal operator  $\Diamond$  is not persistent however, as facts higher up the information ordering may negate any proposition falling within the scope of  $\Diamond$ . Hence there is no constructive counterpart to the existential modality, and “Modulo a few technicalities, the intuitionistic language is the ‘persistent fragment’ of the modal one...”, van Benthem (2009: 257). In this case, we have a single constructive logic/modal logic framework that contains multiple views of informational processes.

### Algebraic and Other Approaches to Modal Information Theory and Related Areas.

Sadrzadeh (2009) develops algebraically an expressively rich modal logic for multi-agent information flow that rejects the necessity of positive introspection, negative introspection, and the T-axiom:

$$(26) K_\alpha \phi \Rightarrow K_\alpha K_\alpha \phi$$

$$(27) \neg K_\alpha \phi \Rightarrow K_\alpha \neg K_\alpha \phi$$

$$(28) K_\alpha \phi \Rightarrow \phi$$

Although the rejection of the *T*-axiom seems odd at first, owing to the factive nature of knowledge, Sadrzadeh is not working with knowledge and belief, but with their weaker relatives, *appearance* (things appear to  $\alpha$  to be thus and so) and *information* ( $\alpha$  receives the possible false information) from which knowledge and belief may be gotten respectively if certain extra conditions are imposed.

Sadrzadeh's motivation is to identify the minimal logical structures which allow for a compositional analysis of the epistemic actions (observations and announcements) which underpin the information flow at work. This is accomplished by the actions being understood as functions that take an agent (or groups of agents or parts of groups of agents) from one state (or

certain states) of information to another (or others). Hence action composition is understood as an instance of function composition, and the algebraic semantics is lifted to the powerful compositional framework provided by *category theory* (see section 4.2 for further connections between category theory and the situation theory expounded in section 2). Sadrzadeh's work is extended in Dyckhoff and Sadrzadeh (2010). Related work on an algebraic analysis of information update via epistemic actions, where such epistemic actions are understood as *resources* may be found in Baltag, Coecke, and Sadrzadeh (2007). For still further connections between modal logical and category theoretical notions such as coalgebras, fixpoint theory, and bisimulation with information flow. For an extension of the logical approach to the analysis of information flow via the logical apparatus of channel theory, see Seligman (2009).

A related modal informational framework for *being informed* is developed by Floridi (2006). Like Sadrzadeh and Dyckhoff, Floridi rejects the introspection axioms, however he retains an informational version of the *T*-axiom. This is due to Floridi working with his veridical notion of strongly semantic information, expounded in section 1.3. Floridi's full logic of being informed is the modal logic KTB, which, reading  $I_\alpha \phi$  as  $\alpha$  is informed that *A*, takes as informationalised axioms the following:

$$(29) I\phi \rightarrow \phi$$

$$(30) I(\phi \rightarrow \psi) \rightarrow (I\phi \rightarrow I\psi)$$

$$(31) NI\phi \rightarrow \phi$$

$$(32) \vdash \phi, \vdash \phi \rightarrow \psi \Rightarrow \vdash \psi$$

$$(33) \vdash \phi \Rightarrow \vdash I\phi$$

*U* is the dual existential modality to *I*, that is,  $N_\alpha \phi := \neg I_\alpha \neg \phi$ , read as  $\alpha$  is uninformed, or for all  $\alpha$ 's information, it is possible that  $\phi$ .<sup>[1]</sup> (Note that

$NI\phi \rightarrow \phi$  is the derivable-in-KTB dual to the usual KTB axiom:  $\phi \rightarrow NI\phi$ .)

Floridi's logic of being informed is a static system. It models the results of cognitive procedures rather than the properties of the cognitive procedures that bring these results about. However, Floridi is aware of the connections between KTB and constructive logics, and Primiero (2006, 2008) extends Floridi's work into a type-theoretical constructive framework for *becoming* informed.

Another type-theoretical approach is Duzi's (2010) type-theoretical approach. Duzi distinguishes between what she calls *empirical* and *analytic* information. As one would expect, empirical information is that which is gotten *a posteriori*, and analytic information is that which is gotten *a priori*. Logically true sentences and logical inferences convey the same empirical information, but their analytic information may differ. Analytic information is procedural information in the sense that it is individuated via the constructive procedures which take the reasoner to the truths in question. Duzi's work is extended extensively in Duzi, Jespersen, and Materna 2010.

## Notes to Logic and Information

1. The issue turns on modelling *inferential information*—the agent possessing the reasoning skills required to move information from implicit to explicit storage. Models for inferential information for non-ideal agents will be explored in sections below, but for now we note that reasoning is just one type of *informational action*—an action-type dedicated to “transporting” information. Another type of informational action is that of an *announcement*.

2. Although a serious comparison lies outside the scope of this entry, DEL has deep connections with dynamic semantics, and the point above is very close in spirit to the insight that “the utterance of a sentence brings us from a certain state of information to another one”, from the Dynamic Predicate Logic (DPL) of Groenendijk and Stokhof (1991). For further information, see van Eijck and Visser (2012).

3. Readers familiar with model theory may have noticed that the ‘part-of’ relation described here is a generalization of the notion of theory interpretation —just consider the case when the instances of the two classifications  $\mathbf{A}$ ,  $\mathbf{C}$  are the models of two first order theories  $T_A$ ,  $T_B$ , respectively, and the types are the sentences of the language of the corresponding theory.”

4. In general, the definition of  $L_A$  can be made more sophisticated, allowing for constraints of a more elaborated form than the simple two-type ones used in this discussion.

5. In other words, the Strong Commutation rule is derived. Strong commutation follows from commutation and association: From  $(A \otimes B) \otimes D$  we can associate to get  $A \otimes (B \otimes D)$ , then commute to get  $A \otimes (D \otimes B)$ , associating again to get  $(A \otimes D) \otimes B$ . In non-associating logics with commutation, commutation is restricted to left-right swaps within the parameters of a given bracketing. In this case, from  $(A \otimes B) \otimes D$  we could get only  $(B \otimes A) \otimes D$ ,  $D \otimes (A \otimes B)$ , and  $D \otimes (B \otimes A)$ . Since non-associative logics with commutation are rare, commutation appears usually in its stronger, derived form. In non-associating logics with commutation, commutation is restricted to left-right swaps within the parameters of a given bracketing.

## Notes to Supplement: Abstract Approaches to Information Structure

1. Floridi uses  $U$  instead of  $N$ , but we have changed the notations here so as to guard against confusion with regard to the universal modality in situation semantics.

Copyright © 2016 by the authors

Maricarmen Martinez and Sebastian Sequoiah-Grayson