EUJAP | VOL. 3 | No. 1 | 2007
ORIGINAL SCIENTIFIC PAPER
UDK: 165
1:18

IN DEFENCE OF THE VERIDICAL NATURE OF SEMANTIC INFORMATION

LUCIANO FLORIDI

University of Hertfordshire
University of Oxford

ABSTRACT

This paper contributes to the current debate on the nature of semantic information by offering a semantic argument in favour of the veridical thesis according to which *p* counts as information only if *p* is true. In the course of the analysis, the paper reviews some basic principles and requirements for any theory of semantic information.

Key words: Bar-Hillel-Carnap paradox; data; semantic information, veridical thesis

Cominius: "Where is that slave which told me they had beat you to your trenches? Where is he? Call him hither."

Marcius (Coriolanus): "Let him alone; he did inform the truth".

Shakespeare, *Coriolanus* Act I, Scene VI.

Introduction

In recent years, philosophical interest in the concept of information and its logical analysis has been growing steadily.[1] One of the current debates concerns the veridical nature of semantic information. The debate has been triggered by the definition of semantic information as *well-formed, meaningful* and *veridical data*, which I proposed a few years ago (see now Floridi 2004b). Such a definition – according to which semantic information encapsulates truth, exactly as the concept of knowledge does – has attracted some criticism for being too strong.[2]

[1] For an updated overview and guide to the literature see Floridi 2005b.

[2] See for example the discussion in Fetzer 2004, with a

In this paper, I offer an argument, which I hope will be conclusive, in favour of the veridicality thesis. In section one, I shall briefly summarise the proposed definition and then clarify the veridicality thesis and the corresponding objection. In section two, I shall present the argument in favour of the veridicality thesis. This will be divided into five steps. I shall conclude by outlining very briefly some of the main advantages provided by a truth-based understanding of semantic information.

1. The definition, the thesis and the objection

"Information" is one of those crucial concepts whose technical meaning we have not inherited or even adapted from ancient philosophy or theology. It is not a Greek word, and the Latin term happens to have a different meaning, largely unrelated to the way we understand information nowadays. Perhaps it is because of this lack of sedimentation that we have so many different ways of understanding it, depending on the specific area of application and the task or purpose orienting one's analysis. Be that as it may, it is plausible to assume that not all concepts or definitions of "information" are born equal. For the principal sense in which we speak of "information" is in terms of *semantic content* functional for *epistemic purposes*.

By semantic content one is to understand here *meaningful* and *well-formed data*. Strings or patterns of data often constitute sentences in a natural language, but of course they can also give rise to formulae, maps, diagrams, videos or other semiotic constructs in a variety of physical codes, being further determined by the appropriate syntax (well-formedness) and semantics (meaningfulness).

Reference to an epistemic purpose highlights the fact that the semantic content in question works as an interface between

a) a system *A* in a specific state, say Paris and its present role as capital of France, which is variously captured by the relevant data (and here one may then refine the analysis by speaking of the data describing, modelling, representing etc. *A*), and

b) an agent *a*, who can come to know *A* by variously elaborating the relevant data (and here one may then refine the analysis by speaking of *a* acquiring, interpreting, understanding, etc. the data).

This epistemically-oriented sense of "information" is ordinary and familiar. We recognise it immediately when it is translated into propositional attitudes such as "Mary is informed (in the *statal*[3] sense that she has or holds the information) that Paris is the

reply in Floridi 2005a; or the objections moved by Colburn 2000a, Colburn 2000b and Dodig-Crnkovic 2005.

[3] This is the statal condition into which *a* enters, once *a* has acquired the information (*actional* state of being informed) that p. It is the sense in which a witness, for example, is informed (holds the information) that the suspect was with her at the time when the crime was committed. The distinction is standard among grammarians, who speak of passive verbal forms or states as "statal" (e.g. "the door was shut (state) when I last checked it") or

current capital of France". In the rest of this paper, we shall be concerned with this and only this factual and epistemically-oriented concept of semantic information.[4]

In Floridi (2005a), I argued that the definition of information in terms of alethically-neutral *content* – that is, strings of well-formed and meaningful data that can then be additionally qualified as true or untrue (false, for the classicists among us), depending on further evaluations – provides only necessary but insufficient conditions. If *p* is to count as information, *p* must also be *true*. This leads to a refinement of the initial definition into:

> DEF) *p* counts as information if and only if *p* is (constituted by) *well-formed*, *meaningful* and *veridical data*.

The veridical thesis embedded in DEF corresponds to the one characterising the definition of "knowledge". Taking advantage of this parallelism, one may rely on the ordinary apparatus of modal logic (Chellas 1980) to formalise "*a* is informed that *p*" as $I_a p$, and hence formulate the veridicality (of semantic information) thesis (VT) in terms of the so-called veridicality axiom (also known as *T*, *M* or *K2*) ($\Box \varphi \supset \varphi$) thus:

$$\text{VT)}\ I_a p \supset p$$

The intended interpretation of VT is this: *a* is informed that *p* only if *p* is true, where, for present purposes, "true" is suitable for a Tarskian treatment.

VT associates information logic[5] (*IL*) to epistemic logics (*EL*) based on the normal modal logics *KT*, *S4* or *S5*. It differentiates both *IL* and *EL* from doxastic logics (*DL*) based on *KD*, *KD4* and *KD45*, since, of course, no *DL* satisfies the veridicality axiom.

The problem concerning the veridical definition of information can now be phrased more accurately. What might be objected is not that

a) there are other, non-epistemically-oriented concepts of information whose definitions do not satisfy VT;

for this is trivially true and uninteresting. Nor that

"actional" (e.g. "but I don't know when the door was shut (act)"). Here, we are interested only in the statal sense of "is informed". This is related to cognitive issues and to the logical analysis of an agent's "possession" of a belief or a piece of knowledge.

[4] There are many plausible contexts in which a stipulation ("let the value of x = 3" or "suppose we discover the bones of a unicorn"), an invitation ("you are cordially invited to the college party"), an order ("close the window!"), an instruction ("to open the box turn the key"), a game move ("1.e2-e4 c7-c5" at the beginning of a chess game) may be correctly qualified as kinds of information understood as semantic content. These and other similar, non-epistemically oriented meanings of "information" (e.g. to refer to a music file or to a digital painting) are not discussed in this paper, where semantic information is taken to have a declarative or factual value i.e. it is suppose to be correctly qualifiable alethically.

[5] In Floridi (2006) I have shown that information logic can be modelled in terms of an interpretation of the relation "*a* is informed that *p*" based on the normal modal logic *B*.

b) *IL* cannot be formalised in such a way as to satisfy VT;

for this is merely false, see Floridi (2006). Rather, the objection is that

c) when analysing the semantic, factual, epistemically-oriented concept of information, an *IL* that satisfies VT might be inadequate because too strong.

In other words, one may contend that "*a* is informed (has or holds the information) that *p*" is more like "*a* believes that *p*" rather than "*a* knows that *p*", and hence that the veridicality thesis should be dropped.

In Floridi (2004b), I showed that VT is pivotal in order to solve the so-called Bar-Hillel-Carnap Paradox (more on this in the next section); in Floridi (2005a), I have argued that all the main reasons to support an alethically-neutral interpretation of semantic information are flawed, while there are several good reasons to assume VT. Finally, in Floridi (2006), I have proved that information logic may allow truth-encapsulation (i.e. may satisfy VT) without facing epistemic collapse (i.e. merely morphing into another epistemic logic). I shall not rehearse these results here, since the new task of the present paper is to seek to persuade those still unconvinced that there are very reasonable principles that, if endorsed, force one to include VT in the definition of semantic information.

2. A semantic argument in favour of veridicality thesis

The semantic argument that I wish to offer is based on four elementary principles and three basic requirements. Any satisfactory understanding of semantic information should implement the former and try to satisfy the latter, if possible. For the sake of simplicity, henceforth I shall speak of propositions instead of alethically neutral, well-formed and meaningful data. I shall also assume a two-values logic in which bivalence applies. All this will only simplify the argument and make no difference to its cogency. Moreover, in order to formulate the argument more precisely the following vocabulary will be used:

$D = \{p_1,\ldots p_n\}$; D is a (possibly empty) domain of propositions.

φ, ψ = propositional variables ranging over *D* (for the sake of simplicity I shall occasionally leave implicit the universal quantification when it is obvious).

$S = \{i_1,\ldots i_n\}$; *S* is a (possibly empty) domain of instances of information.

$t(\varphi)$ = φ is contingently true.

$f(\varphi)$ = φ is contingently false.

$t/f(\varphi)$ = φ is contingently true or false.

$T(\varphi)$ = φ is a tautology.

$C(\varphi)$ = φ is a contradiction.

$H(\varphi)$ = primary[6] informative content of φ.

$P(x)$ = probability of x.

Independently of how members of *S* are defined, the four principles are (for *x* ranging over *S*)[7]:

P.1 $\forall x H(x) \geq 0$; *principle of the non-negative nature of information*: no instance of information can have negative, primary informative content.

P.2 $\forall x \forall y((x \neq y) \supset (H(x \cup y) = H(x) + H(y)))$; *additive principle*: for any two different instances of information, their overall informative content is equal to the sum of their informative contents.

P.3 $\forall \varphi(P(\varphi) = 1) \supset (H(\varphi) = 0))$; *inverse relationship principle*: any proposition whose probability is 1 has no informative content.

P.4 $\forall \varphi\ (H(\varphi) = 0) \supset \sim (\varphi \in \mathrm{S}))$; any proposition with no informative content fails to qualify as information.

A brief comment is in order. These four principles are uncontroversial and fairly standard assumptions in information theory and in the philosophy of information (Bar-Hillel and Carnap 1953, 242 ff.; Dretske 1981; Barwise and Seligman 1997; Van Der Lubbe 1997, pp. 10-11). P.1 and P.2 concern *S* and the cumulative nature of informative contents. P.3 and P.4 concern *D* and the relation between information and probability.

We are now ready to consider the general strategy of the argument. Its form is indirect and basically reverses the steps that would be taken in a "slippery slope" reasoning. We shall begin by assuming that opponents of the veridical nature of information are correct. We shall then see that this is too permissive: too many items slip in. We shall then make the definition progressively tighter, until only the items that we wish to include in the definition of information are actually captured, and all the counterintuitive consequences are avoided. At that stage, we shall realize that we have endorsed the veridicality thesis itself.

2.1. First step: too much information

Suppose we equate *S* to *D*, that is, let us assume that we accept the position according to which all propositions, independently of their truth value, already count as instances of information. An elementary consequence of P.3 and P.4 is that:

---

6 "Primary" here means that we are not interested in secondary, derivative or meta-information. Hence the primary informative content of "Paris is the capital of France" is just the information about which city is the capital of which European state, although, of course, it could be used, e.g. by a spy as "code" for a completely different sort of message, or to check someone's knowledge of English and so forth.

7 Note that, since we are assuming the possibility of empty sets, existentially (instead of universally) quantified formulae would be false.

i) $T(\varphi) \supset (P(\varphi) = 1)$;

ii) $(P(\varphi) = 1) \supset \sim(\varphi \in S)$; therefore

iii) $T(\varphi) \supset \sim(\varphi \in S)$.

Tautologies are not instances of information. Intuitively, no one can inform you about the outcome of a tossed coin by telling you that "it is either head or tail", since this much you know already. Note that tautologies represent a limit case since, following P.1 and P.4, they may be represented as instances of information devoid of any informativeness.

This initial and fairly weak constraint on the extension of the concept of semantic information is accepted by virtually all theories of semantic information. Even if we restrict our analysis to propositional information, our initial equation $D = S$ is too permissive and needs to be revised.

2.2. Second step: excluding tautologies

Following the previous analysis and P.1-P.4, we may adopt a weak alethic restriction on the extension of the concept of information, namely:

$$\forall\varphi((T(\varphi) \supset (H(\varphi) = 0)) \supset \sim(\varphi \in S)) \qquad [1]$$

Unfortunately, even if we implement [1], we still have that

i) $(P(\varphi) < 1) \supset (\varphi \in S)$;

ii) $C(\varphi) \supset (P(\varphi) = 0)$; therefore

iii) $C(\varphi) \supset (\varphi \in S)$.

This is what I have defined as the Bar-Hillel-Carnap Paradox, according to which "a self-contradictory sentence, hence one which no ideal receiver would accept, is regarded as carrying with it the most inclusive information" (Bar-Hillel & Carnap 1953, p. 229). Since contradictions are most unlikely, to the point of having zero probability, they are very informative; indeed they are the most informative propositions. Counterintuitively, you may be receiving an increasing amount of information about the outcome of an event by receiving an increasingly unlikely message but, strictly speaking, the most unlikely message is a contradictory one. We need to exclude this possibility. Again, our position is too permissive and needs to be revised.

2.3. Third step: excluding contradictions

The temptation, in taking the next step, would be to impose a straightforward, and very tight, veridicality constraint: something needs to be true to count as information;

this is why contradictions do not count. Yet this would be tantamount to endorsing the veridicality thesis. Our unconvinced opponent might resist the temptation by including, in the original set of propositions, only those that are contingently true or false, and then argue that these and only these qualify as information, independently of their contingent truth values. Here is the new restriction, revised:

$$\forall\varphi((T(\varphi) \vee C(\varphi)) \supset \sim(\varphi \in S)) \quad [2]$$

This seems to be the model of information that most opponents of VT have in mind (see for example Dodig-Crnkovic [2005]). They may accept that tautological and contradictory propositions do not qualify as information because they are in principle not informative, but they are still bent on arguing that contingently false propositions should count as information because they could be informatively useful (e.g. heuristically) or counterfactually informative about what could be (or have been) the case, although not about what is the case.

Intuitive as this might seem to some, it is still an untenable position, since it denies the possibility of erasing (in the sense of loosing) information *syntactically*, that is, by generating inconsistencies (the same φ is affirmed and denied, independently of its truth value or semantic interpretation). Consider that from P1-P.4 and [2] it follows that

$$\forall\varphi\forall\psi((\varphi \neq \psi \wedge t/f(\varphi) \wedge t/f(\psi)) \supset (0 < H(\varphi) < H(\varphi \cup \psi) > H(\psi) > 0)) \quad [3]$$

Formula [3] says that, if you take any two, different, contingent propositions, then the union of their informative contents is always greater that the informative content of each of them considered separately. Now [3] might seem reasonable, until one realizes that it forces one to endorse the following, highly counterintuitive conclusion: by accumulating any contingent propositions, we are always enlarging our stock of information, independently of whether the propositions in question are mutually inconsistent, thus generating a contradictory repository of information. More formally, we are endorsing:

$$H(\textstyle\bigcup_1^n \varphi) < H(\bigcup_1^{n+1} \varphi) \quad [4]$$

Formula [4] is utterly implausible. Although by definition (see P.2) our interpretation is meant to support only zero-order Markov chains, [4] generates sets that are, monotonically, increasingly informative, despite the random choice of the members. In simple terms, according to [4], the more propositions (of a contingent nature) are uploaded in a database, the more informative the latter becomes. This is obviously false, at least because one may "diagonalise" the uploaded propositions in such a way that every progressively odd-numbered proposition uploaded is the negation of every even-numbered one previously uploaded.

A way of making [4] less unpalatable is to add the clause that φ may also range over informationally equivalent propositions (imagine "John drives the car" and "The car is driven by John") and tautologies. In this case one would obtain:

$$H(\bigcup_1^n \varphi) \leq H(\bigcup_1^{n+1} \varphi) \quad [5]$$

Yet even in [5], informative contents cannot decrease in time unless data are physically damaged or erased. In fact, according to [5], it is still almost impossible not to increase informative contents by adding a random choice of contingent propositions. This is just too good to be true. We need to take a further step.

2.4. Fourth step: excluding inconsistencies

The fact that the new interpretation turns out to be so counterintuitive does not prove that it is logically unacceptable, but it does show that it is at least in need of substantial improvements if it has any hope of becoming reasonable. The problem with it is that [2] is still insufficient, so that the ensuing analysis of what may count as information is too inflated, even if one adopts [5]. Our model of information should satisfy the following two requirements.

> R.1 informative contents can decrease syntactically, without necessarily being damaged or erased physically.

In symbols, we have the following consequence:

$$\Diamond(H(\bigcup_1^n \varphi) > H(\bigcup_1^{n+1} \varphi)) \quad [6]$$

R.1 and [6] indicate that, by adding a new proposition, the result could be H(input) ≥ H(output), as we ordinarily assume. Imagine receiving first the proposition that p and then the proposition that $\sim p$. If you unable to assess which message is reliable, the new propositions $p \vee \sim p$ has no informative content.

The second requirement is:

> R.2 an information repository is unlikely to be increased by adding any contingent proposition; that is, the probability that, by adding any contingent φ to a information depository D, the informative content of D might increase becomes lower than (or at best equal to) the probability that it might be equal to what it was before, the larger the repository is.

R.2 further qualifies R.1 entropically: *ceteris paribus* (e.g. given the same amount of computational and intellectual resources involved in the production of informative contents), it is reasonable to assume that informative contents are comparatively more likely to decrease or remain unmodified (depending on how strongly R.2 is interpreted) than to increase. It is often the case that enlarging an information repository becomes increasingly expensive and difficult in terms of availability of management of resources. In symbols we have that, for n progressively larger:

$$P(H(\bigcup_1^n \varphi) \geq H(\bigcup_1^{n+1} \varphi)) > P(H(\bigcup_1^n \varphi) < H(\bigcup_1^{n+1} \varphi)) \quad [7]$$

$$P(H(\bigcup_1^n \varphi) \geq H(\bigcup_1^{n+1} \varphi)) = P(H(\bigcup_1^n \varphi) < H(\bigcup_1^{n+1} \varphi)) \quad [8]$$

Note that the two formulae [7] and [8] need to be written separately because [7] represents a more radical version of R.2 than [8].

The implementation of R.1\[6] and at least R.2\[8] invites the elaboration of a new step, in which tautological and contradictory propositions have no informative content, and hence fail to qualify as information, while care needs to be exercised not to introduce inconsistency in our information repository. Here is the new set of restrictions, revised:

$$\begin{array}{l} \forall\varphi((T(\varphi) \vee C(\varphi)) \supset \sim(\varphi \in S)) \sim C(\bigcap_1^x \varphi) \\ \Diamond(H(\bigcup_1^n \varphi) \leq H(\bigcup_1^{n+1} \varphi)) \end{array} \quad [9]$$

As one would expect, now informative content can decrease syntactically and, if it increases, it does so much less easily than before. Have we finally reached a reliable model of semantic information? Not yet. Consider R.1 and R.2 once again. They specify that adding contingent propositions may lead to a decrease in the informative content obtained, but they do not specify that this might happen only for syntactic reasons (inconsistencies). Information content might decrease also semantically. In other words, the model of information implemented by [9] satisfies R.1/R.2 only partially, because it cannot fully account for the ordinary phenomenon of semantic loss of informative content. This is a serious shortcoming.

Imagine that the last extant manuscript of an ancient work tells us that "Sextus Empiricus died in 201 AD, when Simplicius went to Rome". Suppose this is true. This informative content could be lost if the manuscript is burnt (physical loss of information), if it is badly copied so that the letters/words are irreversibly shuffled or the names swapped (syntactical loss of information), but also if some false statement is added or if the meaning is changed. However, according to [9], no loss of informative content would occur if the copyist were to write "Sextus Empiricus was alive in 201 AD, when Simplicius went to Alexandria". Quantitatively, this may be true, but semantically it seems unacceptable. The former sentence would count as information, if true, the latter would not, if false. This is our third requirement:

> R.3 informative content can be lost both physically, syntactically and semantically.

Information loss can occur by negation, by falsification, by making propositions satisfiable by all possible worlds (the upper limit represented by tautologies) or by making propositions inconsistent. In symbols:

$$\Diamond(H(\cup_1^n \varphi) > H(\cup_1^{n+1} \varphi)) \text{ physically, syntactically and semantically} \qquad [10]$$

R.3 and [10] motivate our last step.

2.5. Last step: only contingently true propositions count as information

Our last step consists now in revising the alethic constraint thus:

$$\forall\varphi((\varphi \in S) \supset t(\varphi)) \qquad [11]$$

According to [11], informative content can easily decrease (one merely need to generate an inconsistency or a falsehood), when it increases it does so even more slowly than in the previous model, and it can now be lost semantically, not only physically and syntactically, as one would reasonably expect from a correct model of informative content dynamics. But [11] is just another way of formulating the veridical thesis. And this shows that Def is correct.

Conclusion

In this paper, I have offered a semantic argument in favour of the veridical interpretation of information. In the course of the analysis, the paper has provided a review of some basic principles and requirements for any theory of semantic information. The main thesis supported has been that semantic information encapsulates truth, and hence that false information fails to qualify as information at all. The expression "false information" is to be equated to expressions such as "false policeman" (not a policeman at all) or "false passage" (not a passage at all), *not* to "false teeth" (still teeth, though artificial) or "false impression" (still an impression, though unreliable).

Two important advantages of this approach have been the solution of the Bar-Hillel-Carnap Paradox in Floridi (2004b) and the development of a logic of information in Floridi (2006). A result of conceptual interest, which has been left to a second stage of the research (but see Floridi (2004a), is the analysis of the standard definition of knowledge as true justified belief, in light of a "continuum" hypothesis that knowledge encapsulates truth because it encapsulates factual semantic information.[8]

[8] I wish to thank Patrick Allo and Paul Oldfield for their useful comments and criticisms on previous drafts of this paper, and Gordana Dodig-Crnkovic for having made her paper available to me before its publication. As usual, they are responsible only for the improvements and not for any remaining mistakes.

REFERENCES

Bar-Hillel Y. & R. Carnap (1953), "An Outline of a Theory of Semantic Information" repr. in Bar-Hillel 1964, pp. 221-274

Barwise J. & J. Seligman (1997), *Information Flow: The Logic of Distributed Systems,* Cambridge: Cambridge University Press

Chellas B. F. (1980), *Modal Logic : An Introduction,* Cambridge: Cambridge University Press

Colburn T. R. (2000a), "Information, Thought, and Knowledge", *Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics,* pp. 467-471

Colburn T.R. (2000b), *Philosophy and Computer Science,* Armonk, N.Y.: M.E. Sharpe

Dodig-Crnkovic G. (2005), "System Modeling and Information Semantics", *Proceedings of the Fifth Promote IT Conference*, Borlänge, Sweden, ed. by J. Bubenko, O. Eriksson, H. Fernlund & M. Lind, Studentlitteratur: Lund

Dretske F. I. (1981), *Knowledge and the Flow of Information,* Oxford: Blackwell, repr. in 1999 Stanford, CA: CSLI Publications

Fetzer J. H. (2004), "Information, Misinformation, and Disinformation", *Minds and Machines*, 14, pp. 223-229

Floridi L. (2004a), "On the Logical Unsolvability of the Gettier Problem", *Synthese*, 142, pp. 61-79

Floridi L. (2004b), "Outline of a Theory of Strongly Semantic Information", *Minds and Machines*, 14, pp. 197-222

Floridi L. (2005a), "Is Information Meaningful Data?" *Philosophy and Phenomenological Research*, 70, pp. 351-370

Floridi L. (2005b), "Semantic Conceptions of Information", *Stanford Encyclopedia of Philosophy,* ed. by E. N. Zalta , <http://plato.stanford.edu/entries/information-semantic/>

Floridi L. (2006), "The Logic of Being Informed", *Logique et Analyse*, 49, pp. 433-460

Van der Lubbe J.C.A. (1997), *Information Theory,* Cambridge: Cambridge University Press

Accepted: June 06, 2007
Received: January 06, 2007

Department of Philosophy
University of Hertfordshire and IEG
University of Oxford
St Cross College OX1 3LZ Oxford, UK
luciano.floridi@philosophy.ox.ac.uk