Etica & Politica / Ethics & Politics, XIII, 2011, 2, pp. 246–254

Analysis, Clarification and Extension of the Theory of Strongly Semantic Information

Marty J. Wolf
Bemidji State University
Department of Mathematics and Computer Science
mjwolf@bemidjistate.edu

Abstract
This paper analyzes certain technical details of Floridi's Theory of Strongly Semantic Information. It provides a clarification regarding desirable properties of degrees of informativeness functions by rejecting three of Floridi's original constraints and proposing a replacement constraint. Finally, the paper briefly explores the notion of quantities of inaccuracy and shows an analysis that mimics Floridi's analysis of quantities of vacuity.

Keywords
Strongly semantic information, Floridi

1. *Introduction*

The Theory of Strongly Semantic Information (TSSI) is first developed in (Floridi, 2004) and later presented in (Floridi, 2011). This paper provides clarification on certain formal details of TSSI. I begin with an analysis of the example Floridi uses to explain degrees of vacuity and inaccuracy and note its shortcomings with respect to degrees of vacuity. I provide a second example that is better suited to illuminate some of the technical details of TSSI. With the new example in hand, I critique the constraints offered by Floridi for degrees of informativeness functions. I reject three of the constraints and propose a clarifying replacement constraint, along with justification for its inclusion. Finally, I explore the notion of quantities of inaccuracy, an idea suggested by Floridi, but not extensively developed. The clarifications in this paper provide further evidence of the litheness and broad applicability of the TSSI.

2. *Background*

In both of the aforementioned articles on TSSI, Floridi uses an example to help explain and motivate TSSI. The example involves E, a microworld that is described ontologically as the set of all possible states over the conjunction of six boolean variables, $W = \{w_i | 1 \leq i \leq 64\}$. Semantically, the set is describe as the set of messages $\Sigma = \{\sigma_i | 1 \leq i \leq 64\}$. For ease of exposition, I will relabel the boolean variables used by Floridi to x_i for $1 \leq i \leq 6$, where each x_i corresponds to the boolean literal in position i in Table 1 on p. 112 of (Floridi, 2011). Also note that for all i, σ_i and w_i are syntactically identical.

Floridi uses this example to describe and illuminate a number of important concepts in TSSI. Fix $w \in W$ as the ontological state. Let $\sigma \in \Sigma$ be the message that is under consideration. Let $\sigma_w \in \Sigma$ be the message that is syntactically identical to w. Floridi's purpose is to devise and explain a system for analyzing the amount of information in σ, especially when $\sigma \neq \sigma_w$.

Floridi introduces the function f, which measures the degree of discrepancy between w and σ. f maps elements of Σ to real numbers in the range $[-1, 1]$ and is constrained to map σ to $[-1, 0)$ if σ is (estimated to be) false and to $[0, 1]$ if σ is (estimated to be) true. Furthermore, $f(\sigma) = 0$ if and only if σ is true and completely and accurately describes w. For the example, $f(\sigma_w) = 0$. Note that $\vartheta = f(\sigma)$ is a measure of how closely and in which alethic direction σ approximates w. "Intuitively, ϑ indicates the distance of an infon σ from a selected situation w and can be read as the degree of support offered by w to σ" (Floridi, 2011, p. 119). Floridi gives formal descriptions of five conditions that discrepancy functions must meet. The two of interest here are M.4 and M.5. M.4 can be stated as: if σ is (estimated to be) false and not a contradiction then $-1 < f(\sigma) < 0$. M.4 constrains the value describing the amount of inaccuracy contained in a false message. M.5 similarly constrains the value describing the amount of vacuity contained in a true message and can be stated as if σ is (estimated to be) true and not a tautology then $0 < f(\sigma) < 1$.

Presumably, f can depend on w as well, although Floridi does not make this point clear. That is, each $w \in W$ may have its own discrepancy function, f_w, to evaluate the degree of discrepancy between σ and w. Each of these f_w would be subject to the meeting the five conditions for discrepancy functions given by Floridi.

Floridi, presumably as a matter of notational convenience, allows ϑ to act as a function. That is $\vartheta = \vartheta(\sigma) = f(\sigma)$. I adopt this convention here as well.

He goes on to describe how to partition the set $\Sigma - \{\sigma_w\}$ into six equivalence classes with respect to inaccuracy. Simply, the method counts the number of literals in σ_i that differ from those in w. Those messages with the same number of differences fall into the same inaccuracy equivalence class. It is important to note that each σ_i is in one and only one of the six classes.

3. *Analysis of the Classes of Vacuity Example*

Floridi's method of determining which $\sigma_i \in \Sigma$ are at the same level of vacuity is not based directly on the messages themselves, but on properties held by semiduals of the conjunction. In this approach, a fixed number, say m, of conjunctions in σ_i are replaced with disjunctions, and if the resulting message is consistent with w, then σ_i is placed in the same vacuity class as all other elements of Σ that are consistent with w when m conjunctions are replaced with disjunctions.

This approach has a number of objectional features. The first has to do with the arbitrariness of choice with respect to which conjunctions are replaced. Presumably, Floridi's method replaces the first m conjunctions based on the ordering of the boolean literals. Typically, any such ordering of the literals is arbitrary. In fact, when all of the operators are conjunctions, it is customary to consider a boolean expression as a set of literals. Thus, rather than one fixed and immutable way to pick m conjunctions, there are $\binom{5}{m}$ ways to choose, none of which have any sort of inherent priority over the others. Further, the approach of privileging the initial conjunctions seems to contradict Floridi's later statement that "all atomic messages ought to be assigned the same potential degree of informativeness" (Floridi, 2011, p. 124).

This objection can be overcome by a small modification of Floridi's approach and adopting his call for equal treatment of the literals (atomic messages) in each message. For each m, apply the semidual construction to all $\binom{5}{m}$ possibilities and then union the resulting messages into a single vacuity class. This approach yields classes of vacuity with the following cardinalities: $|Vac_1| = 63, |Vac_2| = 62, |Vac_3| = 42, |Vac_4| = 22, |Vac_5| = 7$.

While this approach does away with favoring any particular literal over any other, it still suffers. Neither Floridi's approach nor the proposed modification determine a unique ϑ for each σ_i. The most straightforward way to see this is by considering the cardinalities of the vacuity sets. In the example above Vac_1 has 63 elements out of a possible 63. Vac_2 has 62 elements. Clearly, at least 62 elements are contained in at least two Vac_i's. In fact, $Vac_i \subsetneq Vac_j$ for $i > j$ in this example. (It is unclear whether this property holds for Floridi's method.) This observation raises the question of just how vacuous a given message is since it is contained in multiple sets.

A simple response to this concern is to modify the method slightly. Note that as i increases the cardinality of Vac_i decreases. By determining Vac_5 first, the other sets can easily be determined. Using the methodology above determine Vac_4, but exclude from Vac_4 any σ_i already present in Vac_5. Proceeding in this manner and never placing a σ_i in a vacuity class when it is already contained in a (higher numbered) vacuity class results in true equivalence classes.

There is one more small problem with this method (and presumably with Floridi's method as well, but the details are unclear). To see this problem, consider a specific w, say $w_{45} = \bar{x}_1 \wedge x_2 \wedge \bar{x}_3 \wedge x_4 \wedge x_5 \wedge x_6$.

Its literal-by-literal negation, $\sigma_{21} = x_1 \wedge \bar{x}_2 \wedge x_3 \wedge \bar{x}_4 \wedge \bar{x}_5 \wedge \bar{x}_6$ is not classified by the method. As with σ_w, this message can be placed in its own equivalence class, but doing so raises a deeper question. What does it mean when the literal-by-literal negation a boolean expression is deemed to be true and vacuous, regardless of the method used to determine this classification?

One might argue that TSSI itself is weakened by its inability to account for something as simple as boolean logic. However, this example is not strong enough to support such a statement. The difficulty of this example is the example itself. It has been asked to undertake a task it does not, nor should not, have the strength to perform. At the very least, it is counter-intuitive to ascribe a positive alethic value to elements of $\Sigma - \{\sigma_w\}$ when they all are false. It is better to recognize that they all are false and of varying degrees of inaccuracy. They should not be judged to be true and ascribed varying degrees of vacuity. There is no vacuity carried by such a message from Σ. The vacuity is carried by some other expression, one obtained by starting with σ and applying some set of operations to it. The result is an entirely different boolean expression. It may make sense to add these expressions to the model, but the result of that action is a different model requiring a separate and different analysis.

4. *A Second Example*

For this example, fix a microworld where at any one time there are between one and fifty people at the pub. The lower limit of one is there to simplify the exposition of the example and the upper limit may be there for reasons having to do with the amount of space in the pub or safety regulations. Thus there are 50 possible ontological states, $W = \{w_1, w_2, \dots w_{50}\}$. Consider 50 possible semantic messages $\Sigma = \{\sigma_1, \sigma_2, \dots \sigma_{50}\}$. Each of these messages is to be interpreted as σ_i means there are i people in the pub. Note that none of the statements are either tautologies or contradictions.

Let us assume that there are exactly ten people at the pub. Thus, our ontological state is w_{10}. This provides an initial partition of Σ into three sets: the (presumably true) and vacuous statements $\Sigma_V = \{\sigma_i | 1 \leq i \leq 9\}$, the (presumably false) and inaccurate statements $\Sigma_I = \{\sigma_i | 11 \leq i \leq 50\}$, and $\{\sigma_{10}\}$.

Next we construct our degree of discrepancy function.

$$f(\sigma_i) = \begin{cases} 1 - i/10 & \text{for } 1 \leq i \leq 10 \\ 10/i - 1 & \text{for } i > 10 \end{cases}$$

This function trivially satisfies M.1. M.2 and M.3 of Floridi's requirements for discrepancy functions. M.4 is satisfied since false statements, statements that there are more people at the pub than there really are, are all given negative values. Furthermore, the larger i is, the closer $f(\sigma_i)$ is to -1. M.5 is satisfied since true, but vacuous statements, are given positive discrepancy values. This makes sense because, for example, σ_4 is true when w_{10} is the ontological state. Note that when there are ten people in the pub, four people are indeed there, but so are six additional people and this information is not captured by σ_4.

In this example, with this degree of discrepancy function, each element in Σ_V is in a vacuity class by itself and each element in Σ_I is in an inaccuracy class by itself. One might consider alternate discrepancy functions which would give different discrepancy classes if there was a need or motivation to have vacuity or inaccuracy classes that are not singleton sets.

5. *Clarification of Degrees of Informativeness*

Floridi introduces a degree of informativeness function $\iota(\sigma) = 1 - \vartheta^2(\sigma)$ to capture an essential property of statements. While it is clear that the choice of this function is arbitrary, it is then used to justify six constraints (E.1-E.6) for ι identified by Floridi as desirable for degree of informativeness functions. E.1 normalizes informativeness by specifying that the maximum informativeness of any message is 1. E.3 captures the intuitive notion that contradictions and tautologies are not informative, i.e. their degree of informativeness is 0. This is not only desirable, but necessary for the theory. However, E.3 only makes sense in the presence of E.4,[1] which says that if σ is not a contradiction or tautology, then its informativeness must be positive. Thus, E.1, E.3, and E.4 taken together give upper and lower bounds on the how informative a single message can be. Collectively they force any measure of informativeness to be normalized to $[0, 1]$ much like ϑ is normalized to the range $[-1, 1]$.

The remaining constraints, E.2, E.5 and E.6, are all questionable in their desirability. In what follows, I will argue that E.2, E.5 and E.6 should be rejected and conclude that the degree of informativeness function should cap-

[1] Note that there is a typographical error in the presentation of E.4 in both (Floridi, 2004) and (Floridi, 2011). It should read $((0 < \vartheta(\sigma) < +1) \vee (0 > \vartheta(\sigma) > -1)) \rightarrow 0 < \iota(\sigma) < 1$.

ture the nuances of the informativeness of the messages that are contained in the world that is being modeled.

E.2 calls for $\int_a^b \iota(\sigma)dx$ to be a proper integral. Floridi argues that this is desirable to simplify calculations. While at first blush this seems to make sense, there are other functions that are not integrable, such as step functions, that are easy to calculate. Also, as we shall see below, it may be desirable to have $\iota(\sigma)$ be discontinuous at 0, i.e., when $\iota(\sigma) = 1$. While this makes the calculation more complex, it is not impossible.

E.5 calls for "a small variation in $\vartheta(\sigma)$ to result in a substantial variation in $\iota(\sigma)$" (Floridi, 2011, p. 124). This constraint does not take into account the fact that the average slope from $(0, 1)$ to $(1, 0)$ on any curve is -1. The impact of this constraint is that if E.5 is to hold on some part of ι then there must be some other part of ι that does not have the property called for by E.5. Thus, E.5 cannot be universally true for any degree of informativeness function.

E.6 calls for "the marginal information function (MI) [to be] a linear function" (Floridi, 2011, p. 124). Floridi justifies this by appealing to the notion that "all atomic messages ought to be assigned the same potential degree of informativeness" (Floridi, 2011, p. 124). While this justification makes sense for the example from propositional logic used to motivate TSSI, it is not clear that other information systems have the same sort of underlying uniformity. Indeed, the pub example presented in section 4 does not carry this sort of uniformity. Another concern is that if the atomic messages (literals) are uniformly informative, a MI function that is constant more accurately represents the model. Furthermore, this implies that an ι that is linear, rather than quadratic, would be better support the underlying information system.

The pub example bolsters the case for non-quadratic degree of informativeness functions. Note that in the pub example the loss of informativeness in going from σ_9 to σ_8 is no different than loss experienced in going from σ_7 to σ_6. In each case, the change is constant (much like atomic messages in Floridi's microworld E). In general for any two true semantic messages reporting one less person in the pub the loss of informativeness is the same constant. The responsibility of the MI function is to accurately capture this notion. Given that the marginal information is a constant and the fact that the MI function is the first derivative of the degree of informativeness function, ι ought to be linear in this case.

In information systems where "first steps can be expected to bring a comparatively greater loss of informativeness" (Floridi, 2011, p. 124), it is better if ι is discontinuous at 0. For example the function

$$\iota(x) = \begin{cases} (x+1)^3 & \text{for } -1 \leq x \leq 0 \\ -(x-1)^3 & \text{for } 0 \leq x \leq 1 \end{cases}$$

has that first step away from 0 drop off rather sharply in each direction. That first error is significant. It brings "a comparatively greater loss of informativeness ... than following ones" (Floridi, 2011, p. 124). Note that with this function, after significant errors have occurred, say in the range from 0.8 to 1 (also from -0.8 to -1), there is very little change in informativeness from one error to the next. This compounding effect can be enhanced with larger powers for $x+1$ and $-(x-1)$. These sorts of informativeness functions effectively describe non-brittle information systems.

Note that the discontinuity at 0 is not damaging, other than it introduces a slight complication in terms of computation. Around $\vartheta = 0$, analysis must be done carefully.

This example suggests another possibility for the degree of informativeness function–that it need not be the same function for both the inaccurate statements and the vacuous statements. In Floridi's example the curves are mirror images of each other, suggesting that instances of inaccurate messages have the same degree of informativeness as vacuous messages. There is no *a priori* reason to assume this to be the case for all, or even typical, information systems.

The pub example makes the case for a linear degree of informativeness function for the vacuous statements. Analyzing inaccurate statements and their informativeness one sees a precipitous drop in informativeness as i gets larger. That is, σ_{11} is a little informative, but σ_{35} is hardly informative at all. A strong case can be made for using $\iota(\vartheta) = (\vartheta + 1)^3$ for $-1 \leq \vartheta < 0$ for the degree of informativeness function for inaccurate statements in the pub example. In fact, a polynomial of higher degree may be an even better candidate. Regardless, the point is that the informativeness of true messages can be (and often is) different than the informativeness of false messages in an information system.

The primary observation of this section is that the six constraints on the degree of informativeness function are largely conveniences, rather than constraints. E.1, E.3 and E.4 can be retained, although it is important to bear in mind that they are mathematical and intuitive conveniences and should be readily dismissed should they become cumbersome as TSSI does not require them. E.2 can be done away with as it is overly restrictive in its intended purpose. E.5 is impossible to achieve as stated and must be rejected. E.6 should also be rejected since rather than being a desirable constraint, it describes a property of some specific information systems, but not others. Given that the theory of strongly semantic information supposes complete knowledge of the information system, these three constraints can be replaced with the following constraint:

> (E.7) The degree of informativeness function ought to accurately reflect the nature of informativeness of the underlying messages in the information system being modeled.

6. *An Extension to Quantities of Inaccuracy*

Floridi does not take up the notion of the quantity of inaccuracy, although he does suggest its existence (Floridi, 2011, p. 124). It seems that addressing it is less straightforward than addressing quantities of vacuity, since only true messages carry semantic information. Here I briefly extend Floridi's analysis of quantities of inaccuracy to describe the difference in the amount of vacuous (mis)information between any two vacuous messages. The structure of the analysis parallels the structure of Floridi's analysis of the quantities of vacuity.

Note that Floridi defines the quantity of semantic information in a vacuous σ on a normalized basis. He defines the maximum amount of semantic information as the area under the curve ι on the interval [0,1] and then divides that total amount among the true, but vacuous messages.

The first difficulty is determining what the area under the curve ι on the interval $[-1, 0)$ represents, call it μ. Because each of the messages under consideration is false, μ cannot represent an amount of semantic information as it did on the non-negative interval. One possibility is that μ represents the maximum amount of misinformation. Having a maximum possible amount of misinformation is reasonable for any closed information system under such analysis. All of the false statements are classified. Each has been tagged with a fixed degree of inaccuracy. Since both are finite, it is not unreasonable to expect that the amount of misinformation in the system is finite.

Continuing with the parallel analysis, the amount of misinformation in σ is the difference between μ and the amount of inaccuracy in σ. For ease of exposition, let us assume that ι is integrable on [-1,0). Then let $\nu = \int_{\vartheta}^{0} \iota(\sigma)dx$. Now ν is the amount of misinformation in σ. If σ has a low degree of discrepancy from w, then the amount of misinformation is small. As $\vartheta(\sigma)$ moves closer to -1, the discrepancy increases, as does the amount of misinformation.

7. *Conclusions*

In this paper I analyzed some of the technical details of Floridi's Theory of Strongly Semantic Information. This analysis yields insight into some of the nuances of the theory, as well as a rejection of what Floridi identifies as three desirable constraints on degree of informativeness functions. The analysis suggests another constraint that closely ties the informativeness function to the properties of the information system under study. This new constraint increases the robustness of TSSI.

The analysis also provides a minor extension of TSSI into misinformation. Given the pervasiveness of misinformation and its interplay with information in the world in which we live, a real test of the strength of TSSI would be the development of a complete extension that thoroughly accounts for misinformation and its interplay with information. For example, it ought to make sense to compare the total amount of misinformation μ in a system to the total amount of information α (Floridi, 2011, p. 126) in a system. As just a minor illustration using the pub example, it seems that there is potential for a lot more misinformation than information–at least for w_{10}. Maybe the relative values of μ and α ought to reflect this. Perhaps the two ought not be comparable. As TSSI is refined to take into account experiences with information and misinformation, its value as an analytical tool will increase.

References

Floridi, L. 2004. Outline of a theory of strongly semantic information, *Minds and Machines*, 14(2), 197-222.

Floridi, L. 2011. The Philosophy of Information. Oxford: Oxford University Press.