Notes on the Floridian Theory of Strongly Semantic Information

George M. Coghill

May 12, 2016

AUCS Research Memorandum: CS2016–02

1 Introduction

In recent times there has been significant interest in ascertaining means by which one may identify (and quantify) the amount of information contained in entities ranging from simple statements [6] to scientific theories [11]. In this area the term "information" has been most frequently been used in the context of syntactic structures in the Mathematical Theory of Communication (MCT) [1]. A more meaningful theory was presented by Bar Hillel and Carnap [2] where the meaning of propositions is also taken into account. However, the truth of the propositions was not seen as being of fundamental importance; hence this may be viewed as a Theory of Weakly Semantic Information (TWSI) [8]. It is well known that TWSI gives rise to a paradox that limits its usefulness (the Bar-Hillel/Carnap paradox, (BCP).

A major contribution to the field has recently been provided by Luciano Floridi in the form of a Theory of Strongly Semantic Information (TSSI) [7, 8]. In this theory a statement only counts as information if it is true (the *Veridicality thesis*). This approach has led Floridi to posit a solution to the BCP.

In this paper I shall review Floridi's TSSI (FTSSI) and present some (fairly minor) criticisms and some modifications to the theory to make it more general and, in particular, applicable to Model Based Reasoning [4, 9, 10]. The paper is structured as follows: in the next section the BCP is summarised. This is followed by a review of FTSSI, then in Section 4 I provide an analysis of the approach identifying a couple of weaknesses and providing a possible solution. Next I return to assess how the generalised solution relates to Floridi's solution to the BCP. The paper ends with some suggestions to explore its potential for assessing the information content of models in the context of Model-based Reasoning and Philosophy of Modelling, with particular reference to Informational (Computational) Philosophy of Science.

2 The Bar-Hillel Carnap Paradox

Floridi [8] provides a detailed analysis of the Bar-Hillel Carnap paradox as a grounding for the TSSI; however, for the purposes of these notes a summary will suffice. As it happens, there is an excellent summary on the website of the Society for the Philosophy of Information [1] that, since I could not surpass it, I include in full here.

> Shannon and Weaver defined information in terms of probability space distribution. Yehoshua Bar-Hillel and Rudolf Carnap developed a related probabilistic approach which sought to do justice to the problem of meaning, which Shannon and Weaver had deliberately set aside. Their approach was based on what is called the inverse relationship principle. According to this principle, the amount of information associated with a proposition is inversely proportional to the probability associated with that proposition. The core idea is that the semantic content of p is measured as the complement of the a priori probability of p,

[1] More commonly referred to by the misnomer "Information Theory" [13].

$$CONT(p) = 1 - P(p)$$

Where CONT is the semantic content of p (p could be a set of sentences, events, situations or possible worlds). Crudely, CONT(p) is a measure of the probability of p not happening, or not being true. This means that the less probable or possible p is, the more semantic information p is assumed to be carrying. Tautologies, like "all ravens are ravens" have to be true. So they are assumed to carry no information at all. Since the probability that all ravens are ravens is 1, P(p) is 1, so CONT(p) is 1-1, i.e. 0. By extension, we might presume that contradictions – statements which describe impossible states or whose probability is 0, such as "Alice is not Alice" – to contain the highest amount of semantic information. Thus we seem to run into what has been called the Bar-Hillel-Carnap paradox: the less likely a statement is, the greater its informational content, until you reach a certain point at which, presumably, the statement contains no information at all since it is [a contradiction].[2] As Bar-Hillel and Carnap state:

> "It might perhaps, at first, seem strange that a self-contradictory sentence, hence one which no ideal receiver would accept, is regarded as carrying with it the most inclusive information. It should, however, be emphasized that semantic information is here not meant as implying truth. A false sentence which happens to say much is thereby highly informative in our sense. Whether the information it carries is true or false, scientifically valuable or not, and so forth, does not concern us. A self-contradictory sentence asserts too much; it is too informative to be true" (Bar-Hillel & Carnap, 1953, p. 229).

3 The Theory of Strongly Semantic Information

Having clearly located the details of BCP Floridi goes on to identify a number of criteria and desiderata for progressing to a solution. We do not have space in this paper to record all the details of the approach, so we will focus on the key points and issues. The interested reader is referred to chapter 5 of Floridi's book [8] for the details.

Floridi borrows from situation logic the term "infon" (symbolised as σ) "to refer to discrete items of factual semantic information qualifiable in principle as true or false, irrespective of their semiotic code and physical implementation". He states three criteria for information equivalence and from these identifies that the best combination (at least to start with) is "an analysis of the quantity of semantic information in σ including a reference to its alethic value. This is TSSI"[3].

After this he identifies three desiderata for a theory of semantic information; it should:

D.1: avoid any counterintuitive inequality comparable to BCP;

[2]The original says "is false", which is not sufficient condition for containing no information.

[3]Note that this includes a notion of equiprobability that is based on classical probability that may need to be revised in the context of modelling. Also, there is an at least tacit aim that qualitative equivalence should eventually be included.

D.2: treat the alethic value of σ not as a supervenient but as a necessary feature of semantic information, relevant to the quantitative analysis;

D.3: extend a quantitative analysis to the whole family of information-related concepts: semantic information vacuity and inaccuracy, informativeness, misinformation(what is ordinarily called 'false information'), disinformation.

These three desiderata are fundamental. The first two are unproblematic; the third is unproblematic in the abstract, but the way it is unpacked may give rise to some issues (see section 4). And it this unpacking that Floridi proceeds to develop.

The outcomes of this development merit more discussion; and since they have different aspects and sightly different structure, they will be dealt with in the two following subsections.

Throughout the discussion Floridi makes use of an example universe, E, consisting of all possible worlds arising from the conjunction of a set of basic infons. These are made up from two predicates and three constants, which means E has 64 possible worlds (for further details see [8] Pp 111ff).

3.1 Degrees of Inaccuracy

The model universe is maximal and so must contain the true case. Since the infons are conjunctions, any infon, σ, other than that representing the actual state of affairs will be false. However, different infons will contain a greater or lesser number of components that are false, and this gives rise the idea of degree of falsity or inaccuracy. It is straightforward then to create a measure of the distance from the true state of affairs: the ratio of the number of false components to the length of the infon. More formally:

$$-\vartheta(\sigma) = -e(\sigma)/l(\sigma) \tag{1}$$

Here l is the length of the infon (in this case 6), e is the number of erroneous conjuncts in the inforn, and ϑ is the distance of the infon from the true state of affairs (the negative sign refers to the fact that it deals with degrees of error); and spans the range 0 (matching the actual situation) to -1 (the infon contains no truth).

Floridi identifies five conditions for a suitable metric of this type, all or which are straightforwardly unproblematic here. (For details of these see Floridi [8], page 120.)

3.2 Degrees of Vacuity

The other aspect of distance from the actual state of affairs arises when the infon is true by abstraction. The most extreme example of this is a tautology, which includes the actual situation, but is uninformative due to its being true in all circumstances. This would have a distance of +1. In order to fill in the gap between these two extremes Floridi introduced the semi-dual.[4] By this means a set of classes are identified which contain all the infons with the same number of disjunctions. Each member of the class will have the same number of ways of being true given that some components of the infon are false. So

[4]A semi-dual is an infon in which the operators are changed from conjunction to disjunction, but the components are not negated. For example a contradiction is the semi-dual of a tautology.

in this case the distance is the ratio of the number of ways of being true, n, to the size of the universe[5] s^l. More formally:

$$\vartheta(\sigma) = n/s^l \tag{2}$$

3.3 Degrees of Informativeness

Now that a suitable metric has been defined for these two situations Floridi is able to use it to provide a measure of the informativeness, ι, of an infon. Floridi proposes that the distance be viewed as spanning the range [-1, +1], with the actual state of affairs at the origin, the LHS being the degree of inaccuracy (hence the negative value) and the RHS the degree of vacuity.

The fact that the BCP arises, at least in part from the inverse relation between information content and probability in that formulation suggests that a relation that identifies the state of affairs as having maximum informativeness and the two extremæ (tautology and total falsity), zero informativeness.

Floridi proposes a quadratic relation as meeting a number of criteria that he considers mandatory for a measure of informativeness to possess. Most of these are unobjectionable, but a couple appear more problematic and less than optimal (the issues surrounding these will be dealt with in section 4.2 below). The precise formulation used is:

$$\iota(\sigma) = 1 - \vartheta^2(\sigma) \tag{3}$$

This representation (shown in Figure 1) allows him potentially to provide a solution to the BCP and create measures for vacuity and quantities of semantic information.

4 Some possible weaknesses in FTSSI

It is clear that Floridi has taken a major step forward in the understanding of semantic information and the various desirable measures associated therewith. However, there are a couple of weaknesses[6] in the current formulation that limit its applicability. These weaknesses are not all of the same importance with respect to the research programme I am undertaking, and some are not important at all from the perspective of Floridi's original motivation; nonetheless they all merit critical analysis.

4.1 Constraints on the equation for Informativeness

The first issue regards the form of the relation for informativeness.

As noted above, Floridi identified a quadratic as a suitable equation to relate distance and informativeness. He also states that "If possible, the equation should satisfy ... 6 constraints, derived from the five necessary conditions for a satisfactory metric" [8], Pp 124 ff. He believes the quadratic (and no other) equation meets all these constraints. This is true in that at least one of the constraints applies only to quadratics. However

[5]Here s represents the number of values an infon can take (in this case two: {T, F}), and l is the length of the infon

[6]Perhaps "weaknesses" is too strong, but there are some improvements that can be made.

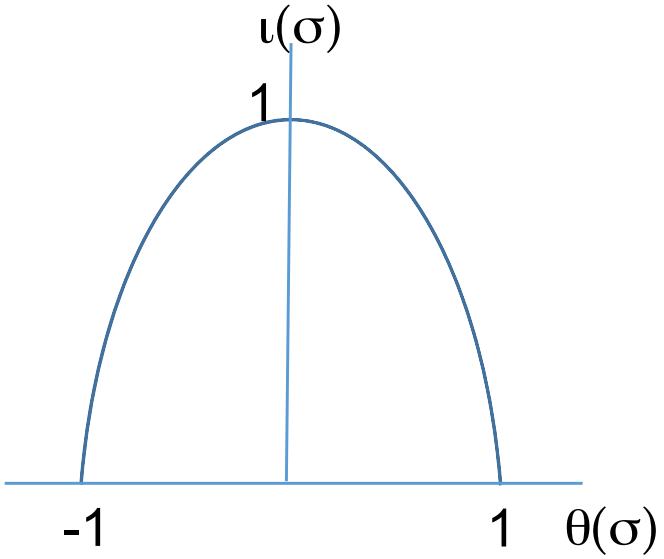

Figure 1: The degree of informativeness (after Floridi [8]). Here, $\iota(\sigma) = 1 - \vartheta^2(\sigma)$

it is not obvious that two of the constraints are appropriate (as one might expect one of those is the one just mentioned) so we shall look at these more closely (the others are straightforwardly correct and will not be dealt with further for the moment).

The two constraints are:

(E.5) a small variation in $\vartheta(\sigma)$ results in a substantial variation in $\iota(\sigma)$.

(E.6) the *marginal information function* (MI) is a linear function.

Taking (E.5) at face value, it is not clear that a quadratic meets this criterion. It is only at the extremæ of the curve that a small variation in ϑ gives rise to a substantial variation in ι; near the origin the converse is the case. However, even if we were to accept E.5 as it stands, a quadratic is not the only curve that meets this criterion: arguably the *arc of a circle* meets it better.

On the other hand with regard to E.5 Floridi states:

> Formula [E.5] is meant to satisfy the requirement according to which, the lower $\iota(\sigma)$ is, the smaller is the possible increase in the relative amount of vacuity or inaccuracy carried by σ. [8] p124

This is not straightforward to interpret with regard to a single criterion. (E.5) refers to ι which is a function of σ and ϑ; and the the measures of inaccuracy and vacuity are different: the former being based on the number of false conjuncts whereas the latter is based on classes of infons. In fact, with the focus here on small change, it is possible to read the explanation as being in conflict with E.5.[7]

Turning now to E.6. The justification provided is:

[7]It is possible that there is a typo in E.5 and it should read "a small variation in $\vartheta(\sigma)$ *does not* result in a substantial variation in $\iota(\sigma)$". Nonetheless, the comments above would still stand, albeit the other way round with the small change only being true near the origin.

> Formula [(E.6)] is justified by the requirement that, a priori, all atomic messages ought to be assigned the same potential degreee of informativeness and therefore, although [21][8] indicates that the graph of the model has a variable gradient, the rate at which $\iota(\sigma)$ changes with respect to change in $\vartheta(\sigma)$ should be assumed to be uniform, continuous and linear. [8] p124 – 125.

There is an intuitive attractiveness about this; however, there are a couple of comments that can be made. Firstly, it is not clear the the relation in question captures this requirement. The positive part of the equation relates informativeness to a set of σ possibilities; and the distance is an abstraction operation which achieves its distance via the increase in the number of disjunctions in the infon (the number of atomic infons remaining constant). Secondly, this has a form analogous to the idea that the whole cannot be greater that the sum of the part: which is, at least, debatable.

4.2 The status of a Contradiction

One of, if not *the*, reasons for the development of TSSI was to provide a solution to the BCP. That being the case, having a clear and coherent means of calculating the informativeness of a contradiction is crucial, and it is this that the relation in Equation 3 seeks to provide.

Floridi [8] identifies a number of criteria related to the measure of degree of inaccuracy; the key one being:

$$(M.3) \quad |\models_{\neg\exists w} \sigma \rightarrow f(\sigma) = -1 \tag{4}$$

That is, "if (it is estimated that) σ is false and conforms to no possible situation, then σ is a contradiction and it is assigned the maximum degree of negative discrepancy" ([8] p 120). This is as it should be and so the calculation of inaccuracy should naturally give this answer for a contradiction.

The next criterion is stated as:

$$(M.4) \quad |\models_{w-\vartheta} \sigma \rightarrow (0 > f(\sigma) > -1) \tag{5}$$

meaning "if (it is estimated that) σ is contingently false, then it is assigned a degree of discrepancy with a value less than 0 but greater than –1 (degrees of semantic inaccuracy)" ([8] p 120).

Unfortunately the example used to illustrate and explain these principles, with the distance metric given in equation 2, appears to be in violation of both these criteria.

The example used is a conjunction of six atomic infons giving a universe, E, of sixty four possible worlds/situations. The distance from the actual world, w, is calculated from equation 2 and the results are shown in Table 1

It can be seen from column 4 of the table that for the class $Inac_6$ the degree of inaccuracy (i.e. the distance from w, ϑ) is equal to -1. Since the only member of $Inac_6$ is a contingent falsehood it is in violation of (M.4), which states that all contingent falsehood must have an inaccuracy strictly greater than -1.

[8] [8] has "[27]"

number of erroneous atomic messages in σ_i	classes of inaccuracy	cardinality of $Inac_i$	degrees of inaccuracy $-\vartheta(\sigma) \in Inac_i$	degrees of informativeness $\iota(\sigma) \in Inac_i$
1	$Inac_1$	6	-1/6	≈ 0.972
2	$Inac_2$	15	-1/3	≈ 0.888
3	$Inac_3$	20	-1/2	$= 0.75$
4	$Inac_4$	14	-2/3	≈ 0.555
5	$Inac_5$	7	-5/6	≈ 0.305
6	$Inac_6$	1	-1	0

Table 1: Classes of inaccuracy in E (from Floridi [8])

On the other hand, if we apply Equation 2 to a contradiction then we will find that the inaccuracy is not equal to -1. This is because any contradiction must contain at least one conjunct that is true, and so the degree of inaccuracy can never be -1 (in violation of (M.3))! In fact for an infon containing n conjuncts the number of true conjuncts will lie in the range from 1 to $\frac{n}{2}$, and so ϑ will lie in the range $\frac{1}{n}$ to 0.5.

In the light of both these issues we will need to find a more appropriate metric to represent the inaccuracy of an infon. I address this in section 5.2.

4.3 The Information Space

It has already been noted that the distance metric used on the LHS and RHS of the diagram representing degrees of informativeness (Figure 1) are not the same: different criteria are used to calculate the distance.[9] On the LHS it is a straightforward ratio, whereas on the RHS it is a class measure. As such it seems a little strange that Floridi would use the same axis to represent both.

Another issue with this representation arises if we wish to use the informativeness relation to represent the relationship between different information statements, or shifts between infons (e.g. by means of abstraction). As an illustration consider the situation where the infon (length six as in the original example) contains one false atomic statement. In that case the informativeness would be -0.972. If we now abstract the infon, by means of a single disjunction, it becomes true with an informativeness of +0.953. This is discontinuous in the most basic sense: the single operation makes the informativeness jump from negative to positive without passing through zero. And this jumping back and forth wold continue as more falsehoods and disjunctions are gradually introduced to the infon. This suggests that it is worth exploring to see if there is a representation and relationship that better captures what Floridi is aiming for here.

To conclude this section: these issues lead one to believe that there are perhaps alternative ways to calculate the informativeness of an infon that may be better able to meet (most of) the criteria Floridi identifies without the ensuing problems. We will explore and expound some suggestions for this in the following section.

[9]As such these may also be in violation of some of the criteria for variables at a particular Level of Abstraction. While the RHS and LHS may have the same Type, they cannot represent the same *Observable* according to Floridi's definitions.

5 A more general representation of TSSI

In the previous section we looked at a number of issues with the current version of the Floridian Theory of Strongly Semantic Information (FTSSI). In this section we will explore some ways in which these might be overcome and lead to a more general version of the theory; in particular one that could utilised in assessing the information content of computational (and possibly other) models of scientific systems/theories.

5.1 The Relation between Inaccuracy and Vacuity

As noted in Section 4.3 the LHS and RHS of the information space do not measure the same thing. In addition they represent contradictory conditions (Truth and Falsehood). The relation was presented in a spatial/ geometrical form, which suggests that it may be dealt with in a manner common to geometrical relations.

One key aspect of representing relations that are incompatible spatially is that their dot product should be zero. This is achieved by making the axes orthogonal. This is a general approach; but since there may be some feeling that true and false can be realised on a single axis since it is similar to moving from Right to Left relative to a point located at an origin we can consider how this is handled in quantum theory [14]. There in, for example, *spin* the concepts of left and right, up and down, and in and out are mapped to a conceptual multidimensional space where "left" is orthogonal to "right" etc.[10]. This representation has the form of a 'unit circle' which is useful because it suggests that information can be represented by complex numbers of various forms, and is that bit closer to the way 'information' (data) is represented in MCT.

With respect to the relation between error and vacuity on the information space, this representation will permit the smooth movement through the space as the vacuity or error changes without introducing any discontinuity.

The relation is shown schematically in Figure 2. Here we have to now include two versions of the distance, $\vartheta(\sigma)$: $\vartheta_F(\sigma)$ for the inaccuracy dimension, and $\vartheta_T(\sigma)$ for the vacuity dimension.[11]

5.2 How to handle a Contradiction

We argued above (in Section 4.2) that the measure of inaccuracy in [8] failed to yield the informativeness of a contradiction as zero, as required by (M.3). The question is, how do we better represent the inaccuracy so that it meets all the constraints, in particular (M.3) and (M.4). It appears to me that there are a few potential solutions: 1) To use the cardinality of the error, 2) to sum the cardinality of the error, 3) make use of the semi dual with respect to disjunction, and 4) make use of the semi-dual with respect to erroneous components.

The first of these is given in column 3 of Table 1 from which it can be seen that it is not monotonic, and therefore not suitable. The second is monotonic, but again does not

[10]This analogy to quantum physics will, I hope, prove useful in the depiction of model spaces and the relation between models, but is beyond the scope of these notes.

[11]The form of the space shown in Figure 2 is polar, the construction of this and its relation to Floridi's original metrics is detailed in Appendix A.

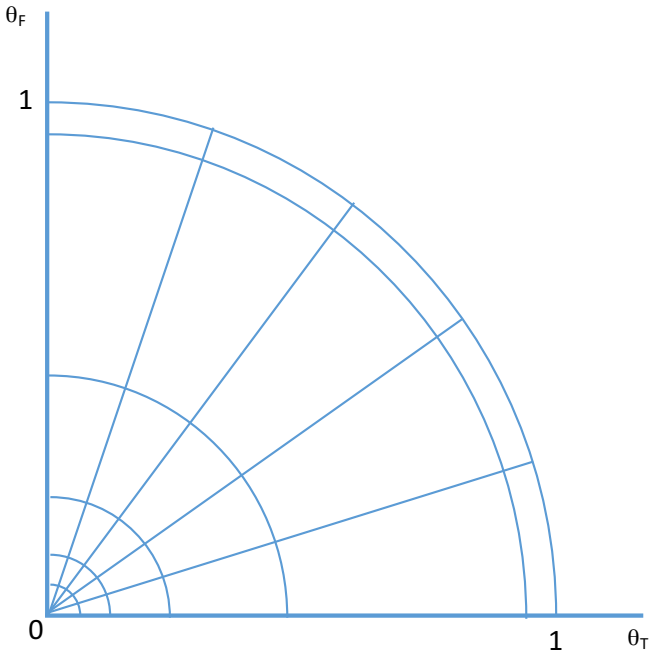

Figure 2: The distance metric space

yield the desired result for a contradiction (i.e. its distance is not maximal).

The third possibility is based on analogy to the semi-dual method applied to the vacuity dimension by Floridi [8]. In that case all the conjuncts are true and the distance is based on successive introduction of disjunctions (from 1 to $l-1$) and identifying the number of ways the components have to change for the infon to become false. In the present case the infon is false and the error is based on the successive introduction of false components. The suggestion is to identify how many disjunctions are required to make the infon true. Unfortunately, this does not solve the problem of a contradiction and has the same drawbacks as the original solution (namely that a the contingent infon with all element false will have an error distance of -1 and a contradiction won't.)

That leaves the final possibility. Here the analogy with the vacuity case lies in accepting that the components are successively false; but we keep the connectives as conjunctions and seek to identify how many changes are required to make the infon true: but in a slightly different way to that suggested by Floridi.

The error distance is defined by Floridi as e/l, where e it the number of erroneous elements in the infon, and l is the length of the infon. This assumes that the only relevant factor is simply the *number* of erroneous elements (because the length is the only thing that is seen as important). However, if we define the distance with respect to the total universe then the maximum length will be 64 (rather than 6). In that case the error should be measured as the number of steps needed to remove the error when it is viewed as a number. In that case the distance to 0 (the actual state of affairs) will be $2^e - 1$. This will make the error and vacuity measures more similar.

A (quite weak) analogy with physical measurement is: if we have 110 as the state we

are in, then how far is this from 000. If we treat it as "the number of wrong elements" then the distances is 2/3. However, if we treat it as a (decimal) number, or say that it has units (e.g metres) then it is clear that the steps to 000 have to go through all the intermediate values (109, 108, ...), and so the distance is 110 units. In addition, if we take the number of wrong elements then 110 and 011 are the same, but obviously very different as a numerical distance.

Of particular importance is the fact that with this measure a contradiction always has the maximum distance, i.e. 64/64 in Floridi's example; and a contingently false infon will always have a distance (ϑ) greater than –1, as required by (M.3) and (M.4)!

The implementation of this also plays to the idea of semi-dual. On the vacuity side the connectives are the number of disjunction and we see how many elements have to be changed before the infon becomes false, In this case it the connectives remain as conjunctions and we see how many elements have to change before it becomes true. The similarity and symmetry should be obvious.

5.3 Constraints on the equation for Informativeness

Now that we have addressed the issues surrounding the relationship between inaccuracy and vacuity, and the informativeness of a contradiction we can return to the constraints (E.5) and (E.6) that affect the putative form of the equation for *informativeness*.

As noted in section 4.1 (E.5) is possibly better met by a circle. Now that we have identified that the unit circle captures the essential relation between inaccuracy and vacuity this seems even more plausible. Of course we now have to represent the relation between informativeness (ι) and distance (ϑ) in three dimensions as an octant of a sphere(as shown in Figure 3).

On the other hand it is still not clear what purpose (E.6) serves. The comments about the marginal informational function may be true in some respects, but they do not appear to be general enough. In fact in Floridi's exposition it would only apply to the inaccuracy side of the space; on the other side the distance is based on sets and subsets, so the fact that each individual infon contains the same information is not relevant. Now, if we accept that the distance for both inaccuracy and vacuity are of similar form, then (E.6) adds nothing. In addition, Floridi's approach makes the whole equal to the sum of the parts, but if the focus is on the relata then it is entirely possible that the whole, from an informational standpoint, is greater than the sum of the parts.[12]

Therefore I propose that (E.6) be dropped as a criterion/constraint.

5.4 The extended Information Space

We are now in a position to put these things together and propose an improved version of TSSI.

Since in this version the dimensions for inaccuracy and vacuity are orthogonal the informativeness, ι of an infon must be a function of both ϑ_T and ϑ_F. In order to permit the representation of information as a complex quantity, and to keep the relation to MCT

[12]A possible example is simple harnonic motion, where, for example using an inductor and capacitor, one can generate oscillatory motion when no such behaviour is manifested by either component individually.

as close as possible for purposes of analysis, we have proposed that a circular relation best fits the bill. In this case that results in the the following formula for $\iota(\sigma)$:

$$\iota(\sigma) = (1 - \vartheta_T^2(\sigma) - \vartheta_F^2(\sigma))^{\frac{1}{2}} \tag{6}$$

This is one form of the equation for a sphere. Since we are only interested in that portion where the ϑs are positive, it is an octant of a sphere as depicted in Figure 3.

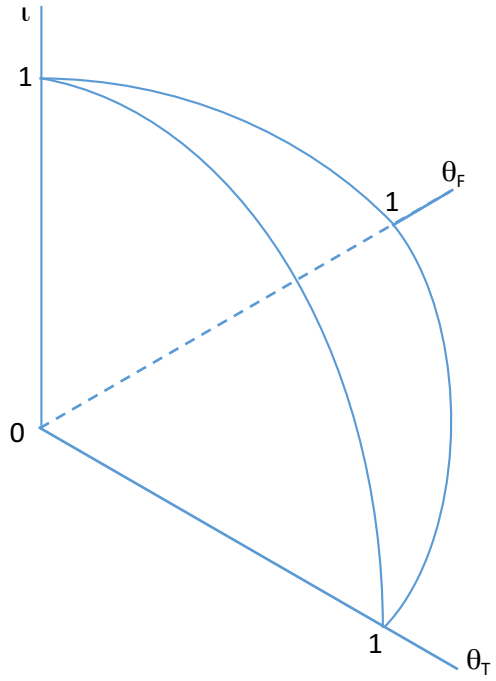

Figure 3: The informativeness of an infon

Following through on the steps followed by Floridi [8], the marginal information function over this space are the derivatives of the informativeness. In this case there are two partial derivatives of Equation 6:

$$\frac{\partial \iota}{\partial \vartheta_T} = \frac{\partial}{\partial \vartheta_T}\left[(1 - \vartheta_T^2(\sigma) - \vartheta_F^2(\sigma))^{\frac{1}{2}}\right]$$

$$= \frac{-\vartheta_T(\sigma)}{(1 - \vartheta_T^2(\sigma) - \vartheta_F^2(\sigma))^{\frac{1}{2}}} = \frac{-\vartheta_T(\sigma)}{\iota(\sigma)}$$

Similarly,

$$\frac{\partial \iota}{\partial \vartheta_F} = \frac{-\vartheta_F(\sigma)}{(1 - \vartheta_T^2(\sigma) - \vartheta_F^2(\sigma))^{\frac{1}{2}}} = \frac{-\vartheta_F(\sigma)}{\iota(\sigma)}$$

It can be seen that in both these cases the result is in the form of a *Tangent*, as one would expect.[13]

The Information Content, $\mathcal{I}$, of an infon was defined by Floridi [8] as the integral of the informativeness. However, he only used the vacuity side of the relation to calculate $\mathcal{I}$. This would suggest that the inaccuracy contributed nothing to $\mathcal{I}$ despite having non-zero informativeness. It seems obvious that the information content of the correct infon will have contributions from both inaccuracy and vacuity. This is easily calculated for the information space as given in Figure 3: it is simply the volume of the octant.

Since the volume of a sphere is: $\frac{4}{3}\pi r^2$, the Information Content (the volume of the octant with $r = 1$) is:

$$\mathcal{I}(\sigma) = \frac{1}{8} \cdot \frac{4}{3}\pi = \frac{\pi}{6}$$

The final calculation is that of the information content of a vacuous (or inaccurate) statement. In [8], for a value of a for ϑ, it is the area from a to 1 ([8], p126f); again omitting any involvement from the inaccuracy part. For our information space, we can identify a possible volume, as depicted in Figure 4.

In this case we exploit the relation between the informativeness of the infon, and the number of false components and disjunctions it contains. For the example in [8], if there are n false components then the infon must have n disjunctions in order to guarantee its truth. This is captured by the *contour lines* in Figure 2. The information content for any given distance is the volume of the octant remaining when the "plug" represented by the relevant contour is removed.[14] The expression for $\mathcal{I}$ is then:

$$\begin{aligned}\mathcal{I}(\sigma) &= \frac{\pi}{6} - \frac{\pi}{24}(h^3 - 3a^2(h-2)) \\ &= \frac{\pi}{6}\left[1 - \frac{1}{4}(h^3 - 3a^2(h-2))\right]\end{aligned}$$

where $h = 1 - \sqrt{(1-a^2)}$ (as shown in Figure 4).

This expression makes normalisation particularly easy: the normalised information content is simply that part of the expression inside the square brackets.[15]

[13] If it turns out to be the case that something like (E.6) is in fact required, then this is the form the constraint should take: that the marginal information function is a *Tangent*.

[14] The volume of the "plug" is one quarter of a cylinder plus a quarter of a spherical cap. There are standard formulæ for these.

[15] While no definite claim is made here as to whether this approach will extend to other possible infons, in particular computational models, it is reasonable to expect that there would be some relation between the degree of inaccuracy of a model and the level of abstraction required to make it true. This is something to be explored further. The expression used is fairly simple whereas the general case would be captured by the solution to the following double integral:

$$\mathcal{I}(\sigma) = \int\int \sqrt{1 - \vartheta_T^2 - \vartheta_F^2}\, d\vartheta_T d\vartheta_F$$

The full expression for the solution to this, as calculated via Mathematica, is:

$$\mathcal{I}(\sigma) = \frac{1}{6}\left(2\vartheta_T\vartheta_F\sqrt{1-\vartheta_T^2-\vartheta_F^2} - \vartheta_F(-3+\vartheta_F^2)tan^{-1}\left[\frac{\vartheta_T}{\sqrt{1-\vartheta_T^2-\vartheta_F^2}}\right] - \vartheta_T(-3+\vartheta_T^2)tan^{-1}\left[\frac{\vartheta_F}{\sqrt{1-\vartheta_T^2-\vartheta_F^2}}\right]\right)$$

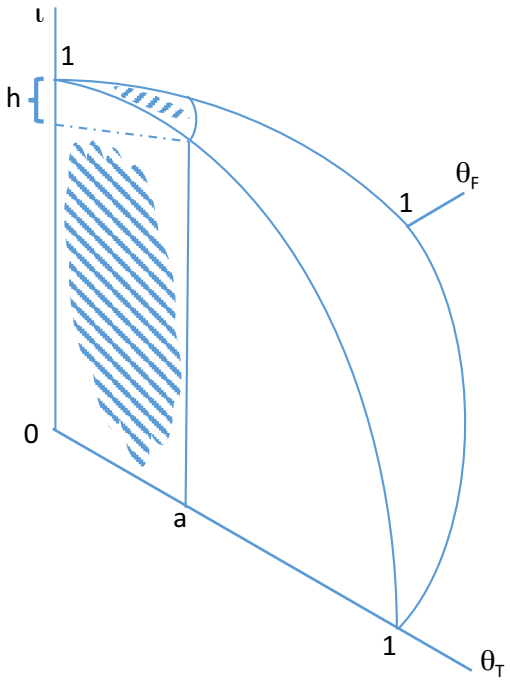

Figure 4: The degree of informativeness of an inaccurate and vacuous infon

6 BCP Revisited

So how does FTSSI solve, or relate to, BCP? Floridi highlights the fact that although in TSSI a contradiction does not have any informativeness and $CONT(\sigma) \neq \iota(\sigma)$, this is not sufficient. In addition Floridi points out that a key difference between TWSI and TSSI is that in the latter information is bound by the veridicality thesis (ie the infon must be true in order to count as information) whereas in the former there is no such restriction (ie "information" may be either true or false). So from the perspective of FTSSI what is being talked about in TWSI, and hence the context in which BCP arises, is *uninterpreted data*, not information.

Since the modifications to FTSSI described in this paper do not contradict, but rather enhance, Floridi's arguments we simply direct the reader to Floridi's writings on the subject for further details ([8], Pp 127ff).

In TWSI the issue arises in the context of the relation between information content

$$-\frac{1}{6}\left(iLog\left[\frac{6i(-1+\vartheta_T^2+\vartheta_F+i\vartheta_T\sqrt{1-\vartheta_T^2-\vartheta_F^2})}{\vartheta_T^2(-1+\vartheta_F)}\right]-iLog\left[\frac{-6i\vartheta_T^2+6i(1+\vartheta_F)+6\vartheta_T\sqrt{1-\vartheta_T^2-\vartheta_F^2}}{\vartheta_T^2(1+\vartheta_F)}\right]\right)$$

The result of using this integral to calculate the volume of the full octant yields:

$$\mathcal{I}(\sigma)=\frac{1}{12}(2\pi-i(-4+Log(16)))$$

The Real part of which is $\frac{\pi}{6}$, the actual volume. This suggests that, although the formula is cumbersome and the calculations tedious, we can use the real part of the result for any particular values of ϑ_T and ϑ_F to get the answer we require.

and probability. In TWSI probability refers to the *a priori* likelihood of the infon being true (which is a measure of the possible worlds it excludes) and Floridi's arguments follow therefrom. On the other hand, it is not without reason that we have proposed that the relations in the information space be circles. This allows one to view information as a complex number, and from there it is a short step to the notions of quantum probability as utiilised in the cognitive science/mathematical psychology of Busemeyer *et al* [3].

From that viewpoint the projections to the subspaces are probability amplitudes (and their squares are then probabilities). This raises the question: "Probabilities of what?" To answer that one must look at the terms *informed*, *misinformed*[16], and *uninformed* and associate them with the dimensions ι, ϑ_F, and ϑ_T respectively. Then taking the lead from quantum theory, where the square of the projection is the probability that a particular spin (down say) is 1, one may say that the value of the square of the projection onto the ϑ_T dimension is the probability that one is completely uninformed, say.

Then if one lets $P(i)$, $P(u)$, and $P(m)$ be the probability of being informed, uninformed, or misinformed respectively, we have:

$P(i) = \iota^2(\sigma)$, $P(u) = \vartheta_T^2(\sigma)$, and $P(m) = \vartheta_F^2(\sigma)$; and from the geometry of the situation one can write:

$$P(i) + P(u) + P(m) = 1$$

or

$$P(i) = 1 - P(u) - P(m)$$

That is, the probability of one's being informed is related to the probability of one's being uninformed and misinformed.

7 Conclusion

As noted above the Floridi's solution to the BCP was a major step forward and so it is to be hoped that the modifications and extension proposed here make it even more useful in the general context of the Philosophy of Information. However, the main motivation for the analysis is the overall context of my interest in Computational Philosophy of Science (which in line with one of Floridi's suggestions should take a more Informational turn).

While the solution proposed here is of general relevance and applicability, a major interest is to provide a means of identifying the information content of models a model and from that build a means of assessing the similarity of models [15] that includes the level of abstraction of model [5]. The fact that one now has a representation that permits one to view the Informativeness as a complex number enables one to explore the possibility of achieving these aims by means of "quantum geometry" measures [12]. This is my ongoing research activity in the domain.

[16]For the purposes of this discussion one assumes that there is no malicious intent and so *disinformation* is ignored

Acknowledgement

Thank you to Dr David Lusseau who gave helpful comments on an earlier draft.

References

[1] Patrick Allo and et. al. The philosophy of information: a simple introduction, 2014.

[2] Y. Bar-Hillel and R. Carnap. An outline of a theory of semantic information. In Y. Bar-Hillel, editor, *Language and Information*, pages 221–274. Addison-Wesley, 1964.

[3] Jerome R. Busemeyer and Peter D. Bruza. *Quantum Models of Cognition and Decision*. CUP, 2012.

[4] George M. Coghill. A feasibility study on model-based diagnosis for an ammonia washer system. Technical report, Centre of Intelligent Systems, Department of Computer Science, University of Wales, Aberystwyth, 2000.

[5] George M. Coghill. On model-based systems and qualitative reasoning with reference to the philosophy of information. In Luciano Floridi, editor, *5th International Workshop on the Philosophy of Information*, 25–27 March 2013.

[6] Fred I. Dretske. *Information and the Flow of Knowledge*. CSLI, 1999.

[7] Luciano Floridi. Outline of a theory of strongly semantic information. *Minds and Machines*, 14(2):197–222, 2004.

[8] Luciano Floridi. *The Philosophy of Information*. OUP, 2011.

[9] Kenneth D. Forbus. Qualitative physics: Past Present and Future. pages 1–39. Morgan Kaufmann Publishers, Inc, 1990.

[10] Roy R. Leitch, Qiang Shen, George M. Coghill, and Michael J. Chantler. Choosing the right model. *IEE Proc.- control Theory Appl.*, 146(5):435 – 449, 1999.

[11] Karl Popper. *The Logic of Scientific Discovery*. Routledge, 1959.

[12] Emmanuel M. Pothos, Jerome R. Busemeyer, and Jennifer S. Trueblood. A quantum geometric model of similarity. *Psychological Review*, 120(3):679–696, 2013.

[13] C. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, 1949.

[14] Leonard Susskind and Art Friedman. *Quantum Mechanics, the Theoretical Minimum*. Penguin, 2014.

[15] Michael Weisberg. *Simulation and Similarity: Using Models to Understand the World*. OUP, 2013.

A Creation of the Information Space

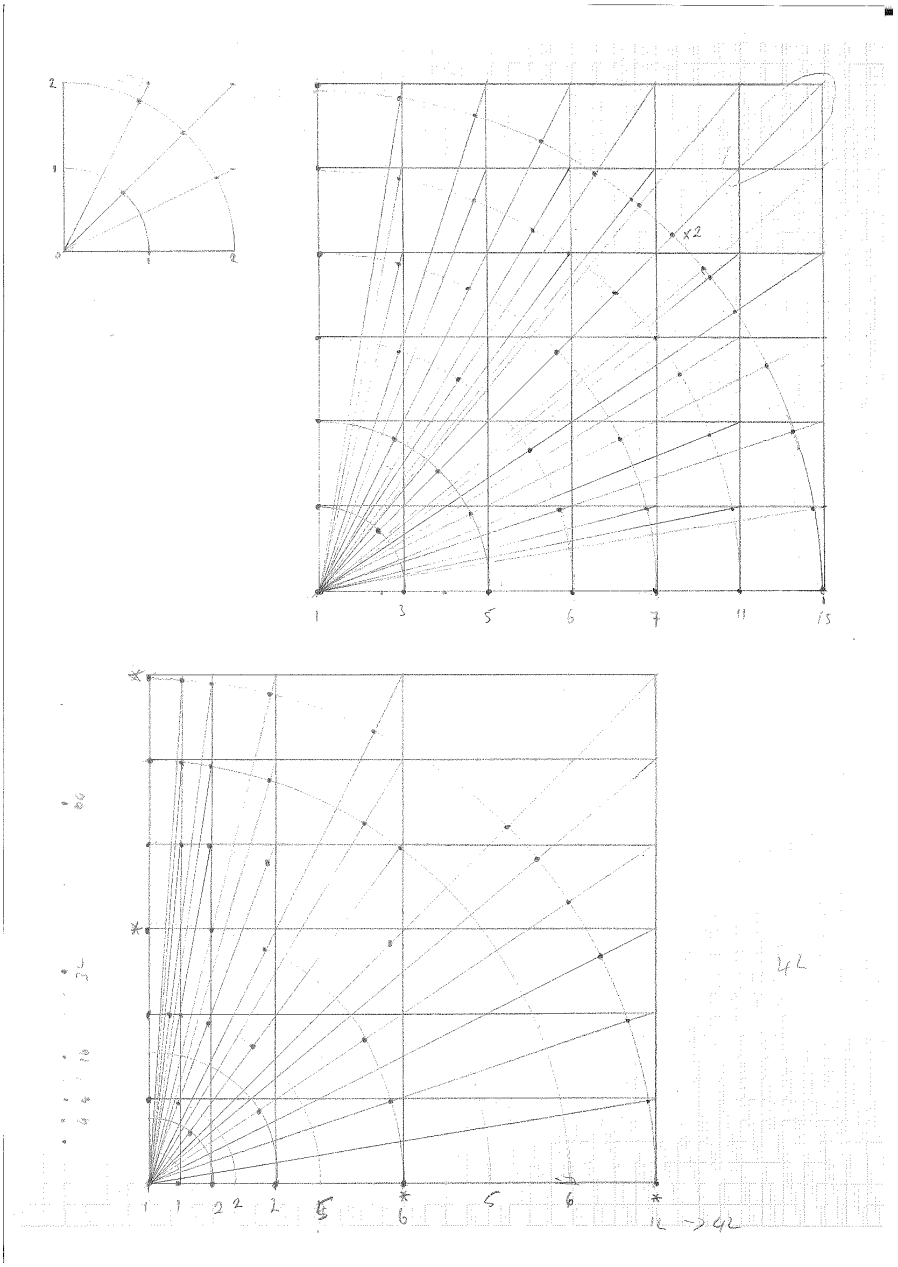

Figure 5: The construction of the infon distance relation