One or Many Concepts of Information?*

Giovanni Sommaruga

University of Fribourg
Department of Informatics DIUF
Bd de Pérolles 90
1700 Fribourg, Switzerland
giovanni.sommaruga@unifr.ch
http://diuf.unifr.ch/tcs/

1 Conceptual Considerations

One of the purposes of this paper is to inquire whether the theoretical (mathematical - logical) study of information can be tracked down to a single concept of information, or whether this study infallibly leads to several distinct concepts (of information). The philosophical interest of this inquiry is, to use an old and slightly infamous concept, self-evident.

1.1 A Few Distinctions to Start with

Let me start by distinguishing 3 kinds of concepts:

a) ordinary language concepts
b) informal theoretical concepts
c) formal theoretical concepts

How does one get from a) to b) or c)? By means of an explication, that is, by a certain way of making the meaning of an ordinary language expression precise (according to Carnap, this way has to satisfy certain well-known criteria).

What is the relationship between b) and c)? There are at least 2 ways to conceive of it:

1. b) is an informal description of a concept embodied in a formal (mathematical) theory. In this case, the relationship between b) and c) is totally unproblematic.
2. c) is to be the formalisation of b). In this case, the relationship is more problematic. (More about it later.)

1.2 Reflections on Informal Theoretical Concepts of Information

How many Concepts of Information? The not entirely unserious question to be answered here is:

(Q) How many informal theoretical concepts of information are there?

* I dedicate this paper to my good friend, the economist, computer scientist and philosopher Ambros Lüthi.

G. Sommaruga (Ed.): Formal Theories of Information, LNCS 5363, pp. 253–267, 2009.

Floridi distinguishes 3 types of answer (‘approaches’) to this question:[1]

$\longrightarrow$ the reductionist answer : 1
$\longrightarrow$ the antireductionist answer : many very different ones
$\longrightarrow$ the nonreductionist answer : here 3 subcases are to be distinguished
$\longrightarrow\longrightarrow$ the centralized answer : essentially 1
$\longrightarrow\longrightarrow$ the multicentered answer : few
$\longrightarrow\longrightarrow$ the completely decentralized answer : quite a few

There is a question analogous to **(Q)** regarding the ordinary language concepts of information, namely:

(Q+) How many ordinary language concepts of information are there?

What do the 3 types of answer to question **(Q)** imply concerning this latter question **(Q+)**? Are there at least as many ordinary language concepts of information as there are informal theoretical explicata? Or is it rather the other way round (are there at most as many)? These questions won’t be followed up in the sequel.

Floridi’s point of view. Floridi’s own point of view is a nonreductionist epistemically centralized type of answer (Floridi 2003b:42). What this point of view amounts to, can well be illustrated by something he calls ‘an informational map’:[2]

Within the network of concepts called ‘information’, the central concept is the one of factual or epistemically oriented (semantic) information. It is defined by the following

Special Def. of (factual) Information **(SDI)**:

σ is an instance of truthful (factual) information iff σ is mwfd + t.

This definition presupposes in turn the

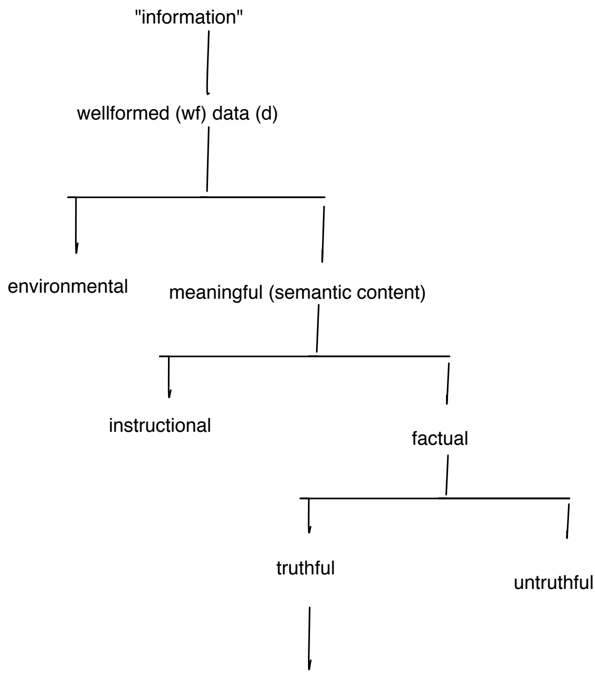
General Def. of Information **(GDI)**:

σ is an instance of information (in the sense of objective semantic content) iff σ is wfd + m.

As a matter of fact, there seems to be a certain ambiguity in Floridi’s point of view:[3] it appears that sometimes semantic information (content) is the central concept, sometimes factual semantic information, and at times it is truthful (factual) semantic information. As will be seen in sect. 3.2, some kind of propensity towards a certain ambiguity is fairly common in this context.

According to the informational map above, factual information is alethically neutral, i.e. neutral w.r.t. truth-values. Floridi points out that this is a point of controversy among information theorists: some take information to be truthful by definition, others take it to be true or false and, at any rate, the truth-value not to be part of the definition of information (Floridi 2003b:45f).

[1] (Floridi 2003b:40f) and (Floridi 2005a:).
[2] Cf. Floridi’s article in this volume, sect. 2.
[3] At least in 2 of his articles, but not in his article (Floridi 2005b).



An informational map (Floridi)

Floridi's thesis If I have understood Floridi's point of view correctly, he upholds something like the following thesis:

(FT) The concept of semantic (factual) truthful information is the most appropriate informal theoretical concept of information.

He appears to justify **(FT)** as follows:

i) This informal theoretical concept explicates or corresponds the best to the most frequent common sense understanding of the word 'information'.
ii) It plays a very important epistemological role, since it provides a necessary condition for knowledge (it is actually sometimes downright confused with knowledge).

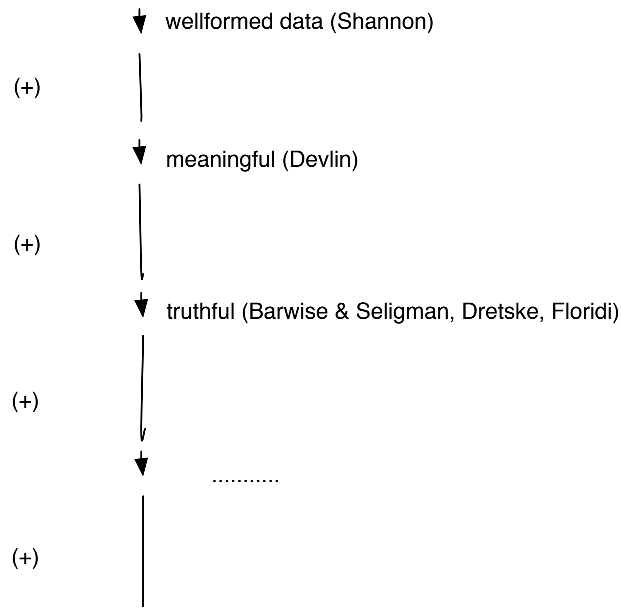
As a consequence of **(FT)**, the concept of semantic (factual) truthful information serves as a standard for the 'measurement' or rather assessment of the relative appropriateness of other informal theoretical concepts of information such as:

Examples

– the algorithmic concept of information (as the size of a computer program necessary to generate certain wellformed data)

- the situation-theoretical concept of information (e.g. as the information content carried by a fact and relative to a constraint)
- the algebraic concept of information (as the information content relative to a question)
- the theoretical concepts of information in the modal or inferential approach (the information content of a sentence as the class of possible states of the universe (possible worlds) which are excluded by the resp. sentence)

These observations suggest another diagram of concepts:



an additive conceptual diagram

As this diagram shows the intension of the informal theoretical concept gets increasingly larger or richer. And **(FT)** allows to assess the degree of deviation of one of these concepts from the most appropriate one.

1.3 Reflections on Formal Theoretical Concepts of Information

The formal theoretical concept of information of statistical IT (information theory) as well as the one of algorithmic IT both simply refer to well-formed data (which needn't be meaningful).[4] What does that mean? The formal theoretical

[4] The term 'well-formed' might actually be misleading and ought then to be replaced by the term 'formed' or rather 'structured' if 'well-formed' should imply 'inductively generated'.

concept of information of strongly semantical IT as well as the one of algebraic IT both refer to truthful meaningful well-formed data. What does that mean?

This indicates - what comes as no surprise - that the just mentioned informal theoretical concepts of information (e.g. as well-formed data or as truthful meaningful well-formed data) are underdetermined: they allow for different formalisations. In other words, different formal theoretical concepts can correspond to one and the same informal theoretical concept of information. And what does that mean?

In order to simplify things, let's make the following assumption:

(A) To every formal theory of information corresponds or in every formal theory of information is embodied a unique formal theoretical concept of information. (This presumably means that the resp. formal theories are what logicians call categorical (or monomorphical).)

Possible relations between two formal theories of information. Let us ask a **(Q)**-type question with regards to formal theoretical concepts of information:

(Q*) How many formal theoretical concepts of information are there?

To answer **(Q*)** let's distinguish the following ways 2 formal theories of information can relate to each other:

Let 2 formal theories be called totally coextensional if they have the same extension, i.e. they have the same (set of) theorems; let them be called partially coextensional, if they share some of their extension, but not the whole one; let them be non-coextensional if their extensions are disjoint. Let 2 formal theories be compatible if the union of their extensions is consistent; o.w. they are called incompatible. The following 6 relations are possible, where case 0 is completely uninteresting (and therefore 'zero' – it is plain inconsistency):

	totally coextensional	partially coextensional	non-coextensional
incompatible	0	2	4
compatible	1	3	5

2 Application of This Conceptual Apparatus to Theories of Information

Applying this conceptual apparatus to the 2 mathematical theories of statistical and algorithmic IT, one can ask the question: Which category do these 2 theories belong to (or in which type of relation do they stand)?

Likewise one can ask the questions: Which category do the 2 formal theories of algebraic IT[5] and (strongly) semantical IT[6] belong to? or Which category do

[5] Cf. Kohlas' and Schneuwly's contribution to this volume.

[6] E.g. in the form of van Rooij's logic of questions and answers, cf. his article in this volume.

the 3 formal theories of (strongly) semantical IT, situation-theoretical IT and Barwise-Seligman IT belong to?[7]

2.1 A Digression on Facts and 'Wishful Thinking'

So far there has been talk about questions and facts. Lets turn for a moment to 'wishful thinking':

Suppose we formulate the following thesis:

(T) The 'essential' meaning of 'information' as truthful meaningful wellformed data is captured by the algebraic IT which is totally coextensional with the (strongly) semantical, (veridical) situation-theoretical and Barwise-Seligman ITs (and possibly other ITs).[8]

(T) is some kind of an information-theoretical analogue to Church's Thesis **(CT)**. Its handicap or one of its handicaps is that the informal theoretical concept of information as truthful meaningful wellformed data is still much vaguer than the analogous one of effectively computable function in **(CT)**. A consequence of this is that **(T)** might not or won't be suitable for use in informal arguments and proofs in information theory the way **(CT)** is in recursion theory.

What is the appeal, the wishfulness (or rather desirability) of such a thesis **(T)**?

The appeal is first and foremost a philosophical one: **(T)** would most likely show that the various equivalent (or totally coextensional) formalisations are appropriate formalisations of the informal theoretical concept of information as truthful meaningful wellformed data. It would mean that there is one formal theoretical concept only corresponding to the just mentioned informal theoretical concept and that the former is clearly a sharpening of the latter. To thesis **(T)** applies what Chaitin wrote about a useful theory: it is a compression of data. And, as he added: comprehension is compression (and at least often, I'd say, vice versa). **(T)** would equally point in the direction of a unification of 'information theory' by connecting similar or related ideas and approaches which have to a great extent been discovered or developed independently from each other and which undoubtedly all have to do with information as truthful meaningful wellformed data.

Needless to add, 'totally coextensional' doesn't mean 'cointensional', and that is why the different totally coextensional theories of information would still reflect or highlight different aspects or features of the informal theoretical concept (as is the case with **(CT)**).

[7] This is a bit of a loose way of writing: The term 'situation-theoretical IT' can refer to Barwise and Perry's original veridical version or to Devlin's and Barwise and Perry's later alethically neutral version of this IT.

[8] J. Seligman drew my attention to the point that there is a descriptive and a normative understanding of **(T)**. Its primary understanding here is descriptive. Should there be or show up a formal IT which doesn't fit in with thesis **(T)**, its understanding might become normative (in the sense of leading to the question: What is wrong with this new formal IT that it doesn't fit thesis **(T)**?).

2.2 Digression Continued

Suppose thesis **(T)** is wishful but not true; i.e. it can be falsified by proving that 2 of the relevant formal theories of information belong to one of the categories 2-5 rather than to 1.

This could mean lots of things. It could e.g. mean that there is a family of meanings of 'information' and one could invoke Wittgenstein's term of family-resemblance (rather than e.g. claim downright ambiguity of the word 'information'). Family-resemblance of a concept means that the concept consists 'of a complex network of overlapping and criss-crossing similarities, just as the different members of a family resemble each other in different respects' (but there isn't one common single feature) (Glock 1996:121). Invoking family-resemblance could be done for good reasons: According to Wittgenstein some branches of a family-resemblance concept allow for an analytical definition, or alternatively, for a formalisation. Another reason is, that even the informal theoretical term of information (as meaningful well-formed data) might refer, just like Wittgenstein's own examples of family-resemblance concepts 'language' and 'proposition', to a variety of different but related phenomena.(cf. Glock 1996:120-124)

2.3 Another Digression on Facts and 'Wishful Thinking'

Even the situation of formal concepts of information as wellformed data (wd) is of considerable interest: The original concept is no doubt the one of Shannon entropy $H(X)$ of the random variable X, i.e. the amount of uncertainty and information of the scheme of choice presented by a random variable X (where a random variable presents a probabilistic scheme of choice where one of its possible values $x_i \in S = \{x_1, \ldots, x_m\}$ is chosen with probability $P_X(x_i)$ ($\sum_{i=1}^{m} P_X(x_i) = 1$ and $0 \leq P_X(x_i) \leq 1$, for $i = 1, \ldots, m$)):

$$H(X) = -\sum_{i=1}^{m} P_X(x_i) \log_2 P_X(x_i) = E\left[-\log_2 P_X(X)\right]$$

Cover & Thomas note that in a certain sense, $\log_2 \frac{1}{P_X(x_i)}$ is the descriptive complexity of the event $X = x_i$, since $\lceil \log_2 \frac{1}{P_X(x_i)} \rceil$ is the number of bits required to describe x_i by a Shannon code. They observe that the descriptive complexity of such an object x_i depends on the probability distribution P_X, it is thus a relative concept (Cover and Thomas 1991:144).

Another famous concept is of course the one of Kolmogorov complexity (plain or prefix) of an object x ($x \in \{0,1\}^*$), also called the algorithmic complexity of x by Cover & Thomas, which is the shortest binary computer program p which describes or rather produces x by means of a universal computer $\mathcal{U}$:

$$K_{\mathcal{U}}(x) = \min\{l(p) | \mathcal{U}(p) = x\}$$

And the conditional Kolmogorov complexity $K_{\mathcal{U}}(x|y)$ is defined by

$$K_{\mathcal{U}}(x|y) = \min\{l(p) | \mathcal{U}(p, y) = x\}$$

where y is taken to be the length $l(x)$ of x (Cover and Thomas 1991:147ff).[9]

Cover & Thomas observe that in the definition of Kolmogorov complexity of an object x, a probability distribution P plays no role whatsoever; it is thus an absolute concept (unlike the one of descriptive complexity). Moreover, the definition of Kolmogorov complexity is also computer independent - this provides for the universality of Kolmogorov complexity (Cover and Thomas 1991:144).

Now, there is the following remarkable relationship between the entropy $H(X)$ of the random variable X and the expectation of the (conditional) Kolmogorov complexity $K_{\mathcal{U}}(x^n|n)$ of an object x^n: the expectation of $K_{\mathcal{U}}(x^n|n)$ differs from $H(X)$ merely by a constant c_P depending on the probability distribution P.

Let x^n be a string of length n of elements $x_i \in S$; and let $P_X(x^n) = \prod_{i=1}^{n} P_X(x_i)$. Then there exists a constant c_P s.t.[10]

$$H(X) = \frac{1}{n}\sum_{i=1}^{n} P_X(x^n) K_{\mathcal{U}}(x^n|n) - c_P$$

In other terms, the expected value of Kolmogorov complexity of a random sequence tends to its Shannon entropy (Cover and Thomas 1991:154f).

As a matter of fact, there exists another universal complexity measure, i.e. another formal concept of information as wellformed data, which serves as sort of an intermediate between Kolmogorov complexity and Shannon entropy, namely the universal probability of an object (a string) x: on the one hand universal probability is essentially equivalent to Kolmogorov complexity, on the other hand it exhibits the same basic form as Shannon entropy.

Cover & Thomas define the universal probability Pr of an object (string) x as the probability that a computer program p randomly drawn as a sequence of fair coin flips $p_1, p_2, \ldots$ produces x by means of a universal computer $\mathcal{U}$:

$$Pr(\mathcal{U}(p) = x) = \sum p : \mathcal{U}(p) = x2^{-l(p)} = P_{\mathcal{U}}(x)$$

They observe that the concept of universal probability is essentially determined by the one of Kolmogorov complexity: There exists a constant c, independent of x, s.t. for all strings x:

$$P_{\mathcal{U}}(x) = c \cdot 2^{-K_{\mathcal{U}}(x)} \qquad \text{or} \qquad K_{\mathcal{U}}(x) = \log_2 \frac{1}{P_{\mathcal{U}}(x)} + c'$$

Cover & Thomas show that Pr is – just as Kolmogorov complexity – a universal probability distribution and it is thus likewise an absolute concept (i.e. independent of any particular probability distribution P) (Thomas and Cover 1991:160ff, 169f).

[9] It is somewhat confusing that Cover & Thomas use K to denote plain Kolmogorov complexity (which Li & Vitanyi denote by C), while Li & Vitanyi use the same letter to denote prefix Kolmogorov complexity. Cf. (Li and Vitanyi 1997:194)

[10] There actually exists a constant c'_P s.t. $c_P = \frac{|S| \log_2 n + c'_P}{n}$.

From all of this, the following conclusions can be drawn:
First, even though the concepts of Shannon entropy and Kolmogorov complexity have been introduced and developed in different contexts (the theory of communication and algorithmic complexity theory resp.) and for different purposes, it is a fortunate coincidence that they both are equivalent up to an additive constant (reflecting the choice of the reference machine).[11]

The third formal concept of information as wellformed data, namely the universal probability of an object, is a concept derived from the one of Kolmogorov complexity. The fact that there are three formal concepts of information as wellformed data, all essentially equivalent with each other, provides a certain evidence for the view that there might be a thesis **(T')** about the concept of information as wellformed data analogous to thesis **(T)**.[12]

The second conclusion is due to the following observations by Cover & Thomas: They draw attention to a striking similarity between $H(X)$ and $\log \frac{1}{P_X(x)}$ in information theory[13], and $K_{\mathcal{U}}(x)$ and $\log_2 \frac{1}{P_{\mathcal{U}}(x)}$ in algorithmic complexity theory. The Shannon code length assignment $l(x_i) = \lceil \log_2 \frac{1}{P_X(x_i)} \rceil$ achieves an average description length $H(X)$; while in Kolmogorov complexity theory $\log_2 \frac{1}{P_{\mathcal{U}}(x)}$ is almost equal to the algorithmic complexity $K_{\mathcal{U}}(x)$ of x. From this they conclude that $\log_2 \frac{1}{P(x)}$ looks like the fundamental or natural form of the descriptive complexity of a string x in algorithmic as well as probabilistic settings (Cover and Thomas 1991:170). That is, $\log_2 \frac{1}{P(x)}$ might be the fundamental form of the coextensional formal concepts of information as well-formed data.

A third and last conclusion is also due to Cover & Thomas: Chronologically, the concept of Shannon entropy is prior to the one of Kolmogorov complexity.[14] However, as was mentioned earlier on, the concept of Kolmogorov complexity is an absolute concept (unlike the one of Shannon entropy), independent of any, or equivalently, universally good for all probability distributions P. Shannon's relative concept of descriptive complexity of an object can therefore be considered as a relativization of the concept of Kolmogorov complexity to particular probability distributions. In this sense, the formal concept of Kolmogorov complexity is conceptually prior to the one of Shannon entropy (Cover and Thomas 1991:144).

2.4 A Third Digression on Facts and 'Wishful Thinking'

Since thesis **(T)** is about meaningful well-formed data mwd, or what is often called semantical information, the proof of some sort of an equivalence of a

[11] Li & Vitanyi just notice that it would have been troublesome had this not been so, as both are intended to express the content of information (Li and Vitanyi 1997:190). And Cover & Thomas call this approximate equality an amazing fact (Cover and Thomas 1991:144) Indeed, if one restricts the theory of Kolmogorov complexity to prefix-free programs (cf. the concept of prefix Kolmogorov complexity), the entire theory is formally equivalent to or coextensional with Shannon's information theory.

[12] For a hint in this direction, cf. (Li and Vitanyi 1997:525).

[13] Remember $H(X) = E[\log_2 \frac{1}{P_X(X)}]$.

[14] (Shannon 1948), (Kolmogorov 1965, 1968).

(strongly) semantical IT with the algebraic IT is of particular interest. It might provide evidence that our wishful thinking is not without at least a grain of truth.

One variant of a semantical IT is van Rooij's logic of questions and answers or at least the starting point of his logic.[15] The demonstration of this sort of equivalence will be carried out by demonstrating that it as well as the first-order predicate logic as information algebra (an algebraic IT) both are examples of a more general structure, exhibited by Jeremy Seligman, namely the Gentzen-theory-based information algebra, or GTBIA for short. The so-called Gentzen theories and the links determined by Gentzen theories are a fundamental component in Seligman's signalling/indicating model of information flow.[16] Thus, there might be something like an embedding of van Rooij's logic of questions and answers and Kohlas' information algebra in the Barwise-Seligman theory of information and its signalling/indicating model in particular.[17]

As a first step, the structure of a Gentzen-theory-based information algebra GTBIA will be introduced.[18]

Let $\langle \Sigma, \vdash \rangle$ be a Gentzen theory (i.e., closed under Identity, Weakening and Cut). And suppose that it is extensional: if $\sigma \vdash \tau$ and $\tau \vdash \sigma$ then $\sigma = \tau$. For $\sigma, \tau \in \Sigma$, say $\sigma \& \tau \in \Sigma$ is the conjunction of σ and τ iff for all $\Gamma, \Delta \subseteq \Sigma$,

(&L) $\sigma \& \tau, \Gamma \vdash \Delta$ iff $\sigma, \tau, \Gamma \vdash \Delta$
(&R) $\Gamma \vdash \Delta, \sigma$ and $\Gamma \vdash \Delta, \tau$ iff $\Gamma \vdash \Delta, \sigma \& \tau$

(Note that conjunctions, if they exist, are unique, by extensionality.)

A question of the theory is a subset $q \subset \Sigma$ such that $\vdash q$ and $\sigma, \tau \vdash$ for each $\sigma \neq \tau \in q$. Say that q_1 is a strong refinement of q_2, written $q_1 \sqsubseteq q_2$, iff for each $\sigma \in q_1$ there is a $\tau \in q_2$ such that $\sigma \vdash \tau$. (This is a partial order.) The restriction of q to σ, written $q|_\sigma$ is the set $\{\tau \in q \mid \sigma, \tau \not\vdash\}$.

Now say that $[q]\sigma \in \Sigma$ is the answer that σ gives to q iff for all $\Gamma, \Delta \subseteq \Sigma$,

($[q]L$) $[q]\sigma, \Gamma \vdash \Delta$ iff $\tau, \Gamma \vdash \Delta$ for each $\tau \in q|_\sigma$
($[q]R$) $\Gamma \vdash \Delta, [q]\sigma$ iff $\Gamma \vdash \Delta, q|_\sigma$

(Again, answers, if they exist, are unique, by extensionality.)

Now suppose that Q is a set of questions closed under joins (least upper bounds in the $\sqsubseteq$ order) and A is a subset of Σ closed under conjunctions and answers to questions in Q, i.e., if $\sigma, \tau \in A$ and $q \in Q$ then both $\sigma \& \tau$ and $[q]\sigma$ are in A.

[15] His logic actually starts off as a semantical IT and soon turns into some sort of a pragmatical IT large parts of which at least can however still be accomodated within a semantical framework. i.e. in terms of a semantical IT.

[16] Cf. Seligman's contribution to this volume, sect. 3.

[17] This latter speculative remark won't be elaborated in the sequel.

[18] I owe the following comments to Jeremy Seligman. Since this point (namely the relationship of the Barwise-Seligman theory to Kohlas' information algebra and to van Rooij's logic of questions and answers) is of considerable importance for my argument, I'm very grateful to him for these comments and his very generous support.

Proposition. By defining $\sigma \otimes \tau$ to be the conjunction $\sigma \& \tau$ and $\sigma^{\Rightarrow q}$ to be the answer $[q]\sigma$, we get an information algebra $\langle A, \otimes, Q, \sqsubseteq, \Rightarrow \rangle$.

This proposition is demonstrated by a fairly straightforward verification that $\langle A, \otimes, Q, \sqsubseteq, \Rightarrow \rangle$ satisfies the axioms of an information algebra.[19]

Call any information algebra of this form a 'Gentzen-theory-based information algebra', or GTBIA for short. This algebra interprets the conjunction of the Gentzen theory as combination of information, as expected, and uses questions as a filtering mechanism: $\sigma^{\Rightarrow q} = \tau^{\Rightarrow q}$ just in case any difference between σ and τ is irrelevance to the answering of q.

Two examples of GTBIAs:

1. For a language L of first-order predicate logic with a set V of individual variables, and a model M with domain D, let $\Sigma = \wp(D^V)$, the powerset of the set of assignment functions, and let $\Gamma \vdash \Delta$ iff $\bigcap \Gamma \subseteq \bigcup \Delta$, which is an extensional Gentzen theory. For each formula φ of L, let $[\![\varphi]\!]$ be the set of assignments that satisfy φ in M. For each set $X \subseteq V$ of individual variables and $\alpha, \beta \in D^V$, define the equivalence relation $\alpha \sim_X \beta$ iff for each $x \in X$, $\alpha(x) = \beta(x)$, and let q_X be the set of $\sim_X$-equivalence classes, which is a question in $\langle \Sigma, \vdash \rangle$.

   Now, let A be the set of $[\![\varphi]\!]$ for each formula φ in L and let Q be the set of q_X for each finite set of variables X. Then Q is closed under joins because $q_X \sqcup q_Y = q_{X \cap Y}$. Also A is closed under conjunction because $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \& [\![\psi]\!]$, and under answers because $[q_X][\![\varphi]\!] = [\![\exists \overline{X} \varphi]\!]$ where $\overline{X}$ is the set of free variables of φ not in X. Then $\langle \{[\![\varphi]\!] \mid \varphi \in L\}, \wedge, \{q_X \mid \text{finite } X \subset V\}, \sqsubseteq, \exists \rangle$ is a GTBIA and shows first-order predicate logic to be a GTBIA.[20]
2. The language LQ of van Rooij's logic of questions and answers extends L by recursively adding a modal operator $?\varphi$ for each formula φ.[21] Evaluate $\psi \in LQ$ in a modal model with possible worlds W and constant domain D, and with $?\varphi$ interpreted by the following accessibility relation

   $u \sim_\varphi v$ iff for each assignment function $\alpha \in D^V$, α satisfies φ in world u iff α satisfies φ in world v[22]

   This is clearly an equivalence relation, so $?\varphi$ is an S5 diamond. Now let $\Sigma = \wp(W)$ and let $\Gamma \vdash \Delta$ iff $\bigcap \Gamma \subseteq \bigcup \Delta$, which is an extensional Gentzen

[19] Cf. Kohlas and Schneuwly's article in this volume, sect. 2.

[20] In (J. Langel and J. Kohlas 2007), the language L of first-order logic is taken to be the language of a many-sorted first-order predicate logic. Technically, this is of no significance whatsoever.

[21] This is a sloppy or even a not entirely correct claim: LQ is closely related to the language of van Rooij's logic, but it isn't its language (its language doesn't involve modalities; LQ rather is a natural extension of van Rooij's language). It looks like everything van Rooij says about questions and answers can be said in LQ. But LQ allows iteration of the question-forming modality which is something, van Rooij doesn't consider.

[22] In other words, $u \models ?\varphi\, \psi\ [\alpha]$ iff there is a $v \in W$ such that $u \sim_\varphi v$ and $v \models \psi\ [\alpha]$. If van Rooij presented the Kripke semantics of his logic as explicitly as it is presented here, he (presumably) would have an equivalent definition.

theory. For each sentence φ of L, let $[\![\varphi]\!]$ be the set of worlds at which φ is satisfiable. For each φ, let q_φ be the set of $\sim_\varphi$-equivalence classes, which is a question in $\langle \Sigma, \vdash \rangle$. Then, as before, let A be the set of $[\![\varphi]\!]$ for each formula φ in LQ and let Q be the set of q_φ for each formula φ in LQ. Then Q is closed under joins because $q_\varphi \sqcup q_\psi = q_{\varphi \cap \psi}$. Also A is closed under conjunction because $[\![\varphi \wedge \psi]\!] = [\![\varphi]\!] \& [\![\psi]\!]$, as before, and under answers because $[q_\varphi][\![\psi]\!]$ is the union of those $\sim_\varphi$ equivalence classes that intersect with $[\![\psi]\!]$, which is $[\![?\varphi\psi]\!]$ and so in A. Then the construction above yields a GTBIA, that is, $\langle \{[\![\varphi]\!] \mid \varphi \in LQ\}, \wedge, \{q_\varphi \mid \varphi \in LQ\}, \sqsubseteq, \cup \rangle$ is a GTBIA and shows van Rooij's logic of questions and answers (or something close to it) to be a GTBIA.

3 Some Conclusions from These Conceptual Analyses and Considerations

3.1 The Upshot

The upshot of these conceptual analyses is the following. An examination of the possibility to make out a reductionist approach in the theoretical study of information yields a negative result: There is no such reductionist way. Thus, the answer to the initial question **(Q)** [How many informal theoretical concepts of information are there?] is: several. The answer to question **(Q*)** [How many formal theoretical concepts of information are there?] is: at least several.

Now, the nonreductionist approach appears to be the right one, as the antireductionist approach is just a lazy answer to the question: how many; since it abstains from any consideration of connections, similarities, networks between different concepts of information. And the next question is, following Floridi's scheme of options, whether it is a centralized or multi-centered approach.[23]

3.2 Further Evidence for These Conclusions

The first prima facie conclusion seems to be what might be called a multi-centralized approach. Evidence for this prima facie conclusion is provided by the fact – already alluded to earlier on – that there is hardly one formal theory of information based on a single informal theoretical concept of information. However, a second look at this conclusion and a closer inspection and analysis of the resp. concepts should or will show that this multi-centralized approach can actually be so to speak 'reduced' to a centralized one.

3 examples:

i) The situation-theoretical IT: In (Israel and Perry 1990) the term 'information' is used ambiguously meaning

[23] The completely decentralized one drops out for the same reason as that speaking against the antireductionist approach. As a matter of fact, the difference between the antireductionist approach and the completely decentralized nonreductionist one is not entirely clear.

a) information content carried by a fact and relative to a constraint.
b) (just) information content.

(Devlin 1991) distinguishes 3 meanings of the term ‘information’, namely

a) true infonic proposition
b) infonic proposition (i.e. information content)
c) parameterfree infon

ii) the algebraic IT: (Kohlas 2003) distinguishes 2 informal theoretical concepts of information, namely
a) a relative one: an information item relative to a domain (question)
b) an absolute one: a domainfree information item (which is an equivalence class of information items relative to a domain)

This example is insofar less telling as the two concepts are shown to be formally equivalent; and moreover, Kohlas makes it clear that the relative concept is the primary one, whereas the absolute concept is derivative and secondary.

There is, however, a more telling ambiguity, namely whether the term ‘information’ refers to

a) a ‘body of information’ (i.e. a set of information items)
or
b) an individual information item

And if it is taken to refer to an individual information item, there appears to be another ambiguity: since an information item is conceived of as an answer to a question, the term ‘information item’ can either be taken to refer to

b1) something like an ordered pair of question and answer (remember: information is a relative concept)
or
b2) to one member of this pair, namely to the answer

iii) The statistical IT: In (Shannon and Weaver 1949) one can find at least two meanings of ‘information’ which obviously cannot be said to be the same:
a) information (from the point of view of the communication engineer) is conceived of as a choice of one message from a set of possible messages
b) information is the decrease (or change) of (in) uncertainty induced by the production of a particular event

3.3 Final Conclusion

Lets come back to question **(Q)** or rather ask a question following from it:

($\mathbf{Q}^f$) Is the multi-centered or the centralized answer to question **(Q)** the more appropriate one?

For the multi-centered answer one might argue as follows: There appears to be something like a center at the syntactical level (cf. thesis **(T')**). And there might be something like a center on the semantical level (cf. thesis **(T)**). One could even be tempted to speculate about the existence of a further center on a 'higher', pragmatical level, a level including epistemic, social, political or other considerations.[24]

It is possible to argue for the centralized answer, and more seems to speak in favor of this latter answer: The center on the syntactical level (e.g. Shannon's statistical IT) is an integrated part of probably every formal IT on the semantical level (i.e. it doesn't just exist side by side and unrelated with a center conjectured on the semantical level).[25] The same line of thought as before could be carried on in the following way: it is possible that, due to the development of information science and information philosophy, a new center of formal ITs will be formed on some sort of pragmatical level and that the center of formal ITs on the semantical level will be an integrated part of this new center. Furthermore, one could add the observation – already contained in a philosophical statement by Seligman – that a center concept on the semantical level is the reduction of uncertainty by gaining information. If uncertainty is conceptually related or tied to (asking a) question, another way of phrasing this center concept is in terms of questions and answers. Again, the model of question and answer plays a crucial role in Kohlas', van Rooij's and Seligman's contribution to this volume.

All this of course is 'information science fiction'. But it is a philosophical-speculative attempt to not only understand where the research on information in the form of formal ITs comes from, but also where it may be heading to.

References

[Cover/Thomas (1991)] Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)

[Devlin (1991)] Devlin, K.: Logic and Information. Cambridge University Press, Cambridge (1991)

[Floridi (2003a)] Floridi, L.: The Blackwell Guide to the Philosophy of Computing and Information. Blackwell, Oxford (2003a)

[Floridi (2003b)] Floridi, L.: Information' in Floridi (2003a), pp. 40–61 (2003b)

[Floridi (2005a)] Floridi, L.: Information, Semantic Conceptions Of. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy (2005a), `http://plato.stanford.edu/entries/information-semantic/`

[24] A hint at such pragmatic theorizing about information can already be found in (Floridi 2005a:41 and also 57); cf. his reference to philosophers such as Baudrillard, Foucault, Lyotard, McLuhan, Rorty and Derrida. Needless to say, the latters' kind of theorizing is not (yet) of the form of a theory intended here.

[25] A hint at this type of relationship is given in (Floridi 2005a:53); cf. MTC as a rigorous constraint.

Moreover, the statistical IT is extensively treated in Kohlas', van Rooij's and Seligman's contribution to this volume.

[Floridi (2005b)] Floridi, L.: Is Information Meaningful Data? Philosophy and Phenomenological Research 70(2), 351–370 (2005b)

[Glock (1996)] Glock, H.-J.: A Wittgenstein Dictionary. Blackwell, Oxford (1996)

[Hanson (1990)] Hanson, P.: Information, Language and Cognition. University of British Columbia Press, Vancouver (1990)

[Israel/Perry (1990)] Israel, D., Perry, J.: What is information? in Hanson, pp. 1–19 (1990)

[Kohlas (2003)] Kohlas, J.: Information Algebras. Generic Structures for Inference. Springer, London (2003)

[Kolmogorov (1965)] Kolmogorov, A.N.: Three approaches to the quantitative definition of information. Problems Inform. Transmission 1(1), 1–7 (1965)

[Kolmogorov (1968)] Kolmogorov, A.N.: Logical basis for information theory and probability theory. IEEE Trans. Inform. Theory IT-14(5), 662–664 (1968)

[Langel/Kohlas (2007)] Langel, J., Kohlas, J.: Algebraic Structure of Semantic Information and Questions. Predicate Logic: an Information Algebra (Technical Report, Fribourg University) (2007)

[Li/Vitányi (1997)] Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and its Applications. Springer, New York (1997)

[Shannon (1948)] Shannon, C.: The Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623-656 (1948)

[Shannon/Weaver (1949)] Shannon, C., Weaver, W.: The Mathematical Theory of Communication. University of Illinois Press, Urbana (1949)