# Introduction

Giovanni Sommaruga

## Part I

This book's topic are formal theories of information. It may be useful to start explaining this topic a little further, and then to say what the structure of the book is like and what motivates it, and finally how this book compares with other works grappling with the same or a similar topic.

What is meant by the term 'formal' in 'formal theories of information'? All of the formal theories presented or discussed in the sequel either have a strongly mathematical or logical flavor or are downright mathematical. And how should the term 'information' be understood in the expression 'formal theories of information'? A first clarification of this term is provided by L. Floridi's introductory philosophical considerations in this book; a second attempt at a clarification is made by G. Sommaruga's concluding remarks.

What is the structure of this book? This book's structure could be represented by some sort of a circular model: The innermost circle will be called the syntactical one: it constitutes the basic skeleton or the set of essential components of any formal theory of information. The second, larger circle is called the semantical one: it adds the crucial feature of meaning to the information-theoretical consideration of mere signs (or well-structured data) in the smallest, innermost circle. The third, even larger, outermost circle might be called the pragmatic one: it adds the crucial feature of real-life usage of meaningful signs by humans to the information-theoretical consideration of mere meaningful signs in the intermediate circle.

This structure is motivated by a doubly unificatory purpose: on the one hand by the question of 'unification' of different approaches inside a given circle; on the other hand by the question of 'unification' underlying the various circles: is it possible to think of one unique concept of information which is gradually built up, developed over several stages represented by the different circles?[1]

K. Kornwachs and K. Jacoby's *Information. New Questions to a Multidisciplinary Concept*(1996) seems to pursue a similar unificatory purpose. The two editors reason in the introduction to their book as follows: There appear to be

[1] An alternative structure of this volume could have been the result of interchanging its second ('the syntactical approach') and its third part ('the semantical approach') for the following reasons: as argued for in sect. 3.3 of my contribution to this volume, the center concept of information is the semantical one which can be phrased in terms of questions and answers. A very sensible way of presenting the following articles would have been to start with the contributions to this center concept and to carry on with two extensions of it: the technical extensions of this center concept (i.e. the syntactical approach) and a pragmatical extension of it (i.e. beyond the semantical approach). I owe this interesting suggestion to Jürg Kohlas.

G. Sommaruga (Ed.): Formal Theories of Information, LNCS 5363, pp. 1–12, 2009.

only the following three kinds of concept of information: a) Shannon's (syntactical) concept modified in many different but not essential ways. It is so limited as to be almost uninteresting (according to a comment by E.U. v. Weizsäcker). b) a very vague concept in everyday language. It is so broad that it is just about meaningless. c) an economical concept of information as a commodity which, however, until now defied any attempt to define it. Thus, Kornwachs and Jacoby reach the following conclusion: The search for a unified concept of information is a hopeless endeavor; information is a multidisciplinary concept, i.e. every scientific discipline has its own concept of information. (1996:1ff) They carry on with the following observation: Scarce applications of Shannon's, i.e. statistical information theory could be made in cognitive science (psychology), biology, system theory, philosophy of science, linguistics and the social sciences. Therefore, all these sciences have started developing their own concept of information. That is why there is no unified concept of information available. Kornwachs and Jacoby continue by making a claim which seems to contradict their earlier conclusion. Claim: A unified concept of information can be reached by a multidisciplinary approach only. What does that claim mean? Does it mean that a unified concept of information has to account for the different concepts of information used in the different scientific disciplines? Is this what is meant by a multidisciplinary approach being a necessary condition for a unified concept? And what does the expression 'account for' imply in this context? Their explanations following their claim are by no means illuminating. Kornwachs and Jacoby's book amounts eventually to presenting various aspects of the concept of information and discussing various uses of the term 'information' in physics, biology, system theory, philosophy of science, philosophy and linguistics, all of this in agreement with their original conclusion, namely that information is (and cannot be but) a multidisciplinary concept.

Another weighty attempt at providing a unified theory of information is provided by W. Hofkirchner's *The Quest for a Unified Theory of Information. Proceedings of the Second International Conference on the Foundations of Information Science*(1999). In his introduction to this volume Hofkirchner raises several questions. The first of these questions is: Which are 'the philosophical and/or formal scientific suppositions [that] seem best suited to serve as a basis for a unified theory of information (UTI)'? (1999:xxi) Hofkirchner answers this question as follows: a UTI ought to be conceived of as a general theory of information-generating systems. (1999:xxii) This answer appears unsatisfactory for at least two reasons: First, in order to identify and construct theories about information-generating systems one has to know what information is, i.e. one has to know the concept of information. Hence Hofkirchner's answer is somewhat viciously circular. Second, these information-generating systems are (according to Hofkirchner) to be considered as particular kinds of systems, as physical, chemical, biotical etc. systems, depending on the material context. This means that UTI has to be conceived as a 'material' theory of information. And this conception implies that the underlying concept of information will at best be an analogous one and at worst equivocal. This consequence is hardly in the spirit of a UTI. Hofkirchner seems to point at a way out of these difficulties: A concept of information should be flexible enough to perform two functions: 'It must relate to the most various manifestations of information, thus enabling all scientific disciplines to use a common concept; at the same time, it must be precise enough to fit the unique requirement of each individual branch of science.' (1999:xxii) No theory can fulfil these two requirements other than a formal (mathematical) theory of information. The term 'formal' should not be understood in a purely formalistic sense, but at most in a sense that S. Shapiro calls deductivist. (cf. (Shapiro 2000:ch.6.2)) It is a logical mathematical theory of information, expressing or incorporating a formal concept of information and applicable to a wide range of scientific disciplines. It thus comes as no surprise when Hofkirchner writes: '[ ] the conference was unable to answer unambiguously the question of whether a UTI is possible at all, and, if so, if a theory of evolutionary systems represents suitable foundations for this; in which way different properties of information-generating self-organizing systems can be subsumed;. . . ' (1999:xxiii).[2,3]

---

[2] D.F. Flückiger distinguishes in his ph.d. thesis *Beiträge zur Entwicklung eines vereinheitlichten Informations-Begriffs*(1995) two types of information theory: the so-called structural-attributive ones whose prototype is D. MacKay's descriptive information theory, and the so-called functional-cybernetic ones whose prototype is Shannon's statistical information theory. (1995:2,69; cf. also his (1999)) He makes an attempt at combining two essential perspectives on information, namely the perspective of information transmission (Shannon) and the perspective of information accumulation (Nauta). Flückiger's goal is to find a (consistent) concept of information underlying both these perspectives (1995:63). On the way to finding such a concept, he makes extensive use of modern brain biology. Flückiger's approach has a similar objective as the Barwise-Seligman theory of information and information flow, but unlike the latter one, it suffers from the same flaw as Hofkirchner's approach, namely from not being a really formal theory.

[3] The objectives of P. Keller's thesis *Information Flow. Logics for the (r)age of information*(2002) are somewhat similar to those of this book: (i) 'to give a conceptual analysis of the notions of information, data and knowledge and their interrelations' – where in this book the concept of knowledge plays no role whatsoever –, and (ii) 'to apply this analysis to the theory of information flow' (2002:I) – where in this book, the analysis is partially applied, partially extracted from the theory of information flow and other formal theories of information –. Keller carries out task (i) by comparing different theories of information with each other, such as Dretske's philosophical theory of information, situation-theoretic information theory and epistemic modal logic of information. He mentions three possible reasons for the apparent fact that the different theories of information considered by him are incommensurable. (2002:VII/VIII) One may be tempted to add a fourth reason, namely that Keller's choice of theories to be compared with each other wasn't particularly fortunate, or say, too heterogeneous. His conclusion at the end of his thesis is disappointed and delusive: 'the concept [of information, G.S.] is elusive and there is not much to be hoped from a 'philosophy' of information' (2002:240), and by no means shared by the editor of this book. It is one among other objectives of this book that the reader may come, after reading this book, to the opposite conclusion.

## Part II: The Individual Contributions

In his contribution **Philosophical Conceptions of Semantic Information**, L. Floridi sets out to explore and clarify the wide and messy conceptual field surrounding the concept of information. Even though he starts by declaring that there is a considerable number of concepts of information depending on the level of abstraction and the requirements of one's perspective, he essentially zooms in on three fundamental concepts: information as (well-structured) data, information as meaningful well-structured data (meaningful content), and information as truthful meaningful well-structured data. He then provides a philosophical discussion of the nature of (well-structured) data. After a brief philosophical presentation of statistical information theory (called MTC by Floridi), he examines the concept of information as semantic content and especially the one he calls factual semantic content (factual information) and he presents a sketch of the debate on whether factual information ought to be truthful or not in order to correctly be called information. At the end, he considers the relationship between MTC and a semantic theory of information, thereby continuing the previous sketch on the level of theories: for the weakly semantic theories of information, information as semantic content is alethically neutral, whereas for the strongly semantic theories, information as semantic content has to be truthful.

The canonical measure of probabilistic uncertainty is *Shannon's entropy* (1948), whose properties and applications constitute *Information Theory.* In Information Theory, the entropy of a message limits its minimum coding length, in the same way that, more generally, the complexity of the message determines its compressibility in the Kolmogorov-Chaitin-Solomonov *Algorithmic Information Theory.* In his contribution **Information Theory, relative Entropy and Statistics**, F. Bavaud summarizes and revisits the classical Shannonian framework from a statistical inferential perspective: besides coding and compressibility interpretations, the relative entropy $K(f||g)$ (or Kullback-Leibler divergence) possesses a direct probabilistic meaning, and measures the badness-of-fit between an empirical distribution $f$ and a model distribution $g$ - a theme first explored by authors such as Kullback, Sanov, Jaynes, Billingsley, Csiszár, and Cover among others. Through about twenty examples, Bavaud illustrates a few formal properties of the functional $K(f||g)$, rich enough to capture the various aspects of the confrontation between models (= what we believe) and data (=what we observe), that is the art of classical statistical inference, including Popper's falsificationism as a special case. In particular, the asymmetry of $K(f||g)$ nicely matches the epistemological asymmetry between data and models, as illustrated by Fisher's single hypothesis testing, the Neyman-Pearson testing between two hypotheses, and Bayesian model selection. Also, the exact additive decomposition of the relative entropy holds in two dual contexts, namely for convex families of empirical distributions, or for exponential families of model distributions. Moreover, the principles of Maximum Likelihood and Maximum Entropy clearly emerge as dual to each other, which clarifies the (often misunderstood) epistemological meaning of the former, namely as a method of reconstructing under incomplete observation the most likely data under some prior model - which is highlighted in the so-called EM algorithm consisting of an alternating use of both principles. In the last section, Bavaud demonstrates how the relative entropy formalism extends beyond independence models, and can be used to test independence or to test the order of a Markov chain. His conclusion, in the spirit of convex and exponential models, illustrates the heating and cooling of texts by a few textual simulations, and the mixing (in an additive or multiplicative way) of English and French texts.

C.S. Calude's contribution **Information: The Algorithmic Paradigm** has almost the form of a dialogue: questions are raised, answers are given which in turn may raise new questions etc. Moreover, a central theme of Calude's with variations is incompleteness. After introducing bits, i.e. binary digits, and bits-strings, Calude raises the question: How efficiently can all the non-negative numbers be coded? In order to answer this question, he introduces a special type of Turing machine, namely the self-delimiting universal Turing machine $U$, and he also explains the following coding problem: If one considers all Turing machines of length at most $n$, i.e. $2^{n+1}-1$ Turing machines, some Turing machines halt on a certain input $x$, others don't. If all the Turing machines of length $n$ are ordered lexicographically and for each Turing machine, one asks whether it stops or not, one gets a bit-string of length $2^{n+1}-1$ encoding the whole information. Can the same amount of information be encoded with fewer bits? The answer is yes, and expressed by the Omega number $\Omega_U$ whose binary expansion is $0.\omega_1\omega_2\dots\omega_m\dots$. The halting information for all Turing machines $p$ s.t. $n \geq |p|$ can then be compressed into a string of length $n : \omega_1\omega_2\dots\omega_n$. It is now possible to answer the original question: The most efficient coding of all the non-negative numbers is provided by the domain of a self-delimiting universal Turing machine. Calude continues showing that many problems in mathematics can be rephrased in terms of the halting/non-halting status of appropriately constructed self-delimiting Turing machines. The next question to be discussed is whether computers can produce new information? The amount of information $H_U(x)$ contained in a bit-string $x$ is the smallest length of a Turing machine by means of which a self-delimiting universal Turing machine $U$ produces $x$. To produce new information then means to start with an input $x$ and produce an output $y$ s.t. $H_U(x) < H_U(y)$. The question just asked becomes: Is there any computable process which can produce infinitely many outputs each having more information than its corresponding input? Calude demonstrates that the answer is essentially negative: a computer cannot create much new information. The ensuing question is: But how much can one expect to be created? If a Gödelian theory is roughly speaking a formal theory for which Gödel's incompleteness theorems hold, such a theory can be used to prove theorems having a bit more information than the theory itself. The next point raised concerns a link between algorithmic and statistical information theory: Calude presents an algorithmic version of Shannon's noiseless coding theorem. Next, he treats the relationship between algorithmic randomness and incompleteness: An infinite sequence $x_1x_2\dots x_n\dots$ is algorithmically random if there exists a positive constant $c$ s.t. $H_U(x_1x_2\dots x_n) \geq n-c$. It has then been proved that (the sequence of bits of) $\Omega_U$ is algorithmically random. Questions: i) Are there other natural algorithmically random sequences? And ii) Are there any computably enumerable and algorithmically random numbers other than $\Omega_U$? The answer to the first question is positive: If $\zeta_U$ is the so-called zeta number of a self-delimiting universal Turing machine $U$, $\zeta_U$ can be shown to be algorithmically random. The answer to the second question, however, is negative: for one can prove that a real number $\alpha \in (0,1)$ is computably enumerable and algorithmically random iff there exists a self-delimiting universal Turing machine $U$ s.t. $\alpha = \Omega_U$. The link to incompleteness is establised by the following result: A Gödelian theory cannot determine more than finitely many digits of $\Omega_U$. Calude also comes back to incompleteness in his last point: If one expresses the property of $\Omega_U$ being algorithmically random as the uncertainty relation $\Delta_s \cdot \Delta(\omega_1 \dots \omega_s) \geq 1$, one can derive from it Gödel's incompleteness theorem, that is, uncertainty implies incompleteness.

J. Kohlas starts his article **Information Algebra** by explaining intuitively the basic components and ideas of his algebraic theory of information. In the second section, he gives an axiomatic presentation of the algebra of information which he motivates by showing that the relational algebra associated with relational databases is its prototype. In the following subsections, he reinterprets the projection operation of an information algebra in two ways: by interpreting it as variable elimination, he points to the connection between information algebra and logic; by interpreting it as a transport operation, he prepares the ground for the definition of an interesting equivalent version of the information algebra, i.e. the so-called domain-free one. The third section is dedicated to a variety of examples, non-logical and logical, of information algebras: fuzzy set theory (or parts thereof) can be conceived of as an information algebra, and as for the logical examples, propositional logic, first order logic, and so-called contexts can equally be conceived of as information algebras. The contexts are designed as a more general logical framework for obtaining information algebras. The fourth and last section links information algebra to statistical information theory. The first subsection explains how information algebra gives rise to a natural partial order of information content. In the second subsection, attention is drawn to the fact that in a relational (information) algebra, the information content of a relation depends on the question of one's interest. Since this fact is related to the Boolean character of a relational algebra, the so-called Boolean information algebras are introduced.The next subsection shows how this partial order of information content can be used to define particular information algebras based on basic, finest information pieces, called atoms; these algebras are subsequently called atomic information algebras. The fourth and last subsection deals with the measurement of information content in the case of atomic information algebras, using Hartley's measure. This quantitative information measure measures the reduction of uncertainty by an information element of an atomic information algebra and it also respects the qualitative, partial order of information content. Despite these connections between information algebra and statistical information theory, Kohlas keeps emphasizing that atomic information algebra is not a statistical information theory as entropy doesn't develop its full power in it.

In information algebra, information is represented by an algebraic structure in which pieces of information refer to precise questions and can be combined and focussed on other questions. Uncertain information arises when a piece of information is known to be true under certain assumptions which themselves are not necessarily known to be true. Varying these assumptions leads to different information by means of assumption-based reasoning. If the likelihood of different assumptions is varied and described by a probability measure, it is possible to measure the degree of support of a piece of information in terms of the probability of those assumptions supporting this piece of information. This is the basic tenet of J. Kohlas and Ch. Eichenbergers approach **Uncertain Information**. Section two starts off with a presentation of functional models describing the process whereby a data (an answer) is generated from a parameter (question) and some random element (an assumption). The basic idea of assumption-based reasoning is to suppose that a random element generated some data and then to determine the consequences of this supposition on the parameter (and to determine these consequences in terms of the probabilities of the resp. random element(s)). The last technical term introduced in this section is the one of a hint: A hint is essentially a mapping from a probability space into a certain set, and, more precisely, a mapping of an assumption to the smallest set of possible answers to a given question, containing for sure the right answer. Intuitively, a hint represents a piece of information concerning the right answer to some question, if this answer depends on certain assumptions. Section three introduces a generalisation of the hints, namely random variables with values in an information algebra: A (simple) random variable is a mapping from a (finite) probability space whose elements represent uncertain assumptions into an information algebra. Since it is shown that (simple) random variables form themselves an information algebra, they are on the hand information, and, due to their relation to a probability space, they are on the other hand uncertain information. Section four associates random variables with probability distributions: These probability distributions arise from the probabilities of the assumptions supporting the answers to some question. A degree of support of an answer to some question, as well as a degree of possibility (or plausibility) of some answer are defined by means of the random variables. The support and the possibility function actually represent distribution functions of the random variables and can, in the case of simple random variables, be defined in terms of basic probability assignments. In the basically last section five, the fact is exploited that uncertain, i.e. assumption-based, information is also information, i.e. constitutes an information algebra of random variables. This fact allows for the definition of an order between the elements of this information algebra. This order is induced by the algebra and reflects a comparison of random variables w.r.t. information content taking into account that this information content is also related to assumptions. If the information algebra of random variables is Boolean, it can be generalised in such a way as to admit also of varying probability spaces of assumptions. In this latter case, a measure of information content can be introduced in a way analogous to the one presented in the article Information Algebra provided the respective Boolean information algebra is atomic. This measure of information content can be regarded as the reduction of uncertainty by the random variable w.r.t. full ignorance (where uncertainty is measured by Shannon's concept of entropy adapted to the information algebra of random variables). Thereby, a link to Shannon's theory of information is established.

The general theme of R. van Rooij's article **Comparing Questions and Answers: A bit of Logic, a bit of Language, and some bits of Information** is the informative value of questions and answers and its measurement. Van Rooij's contribution is set up in a dialectical way: He begins with a first definition of this informational value and then points out its limitations. He goes on giving a second definition which takes into account the crtiticism of the first, but then he points out the limitations of this second attempt. And so he carries on presenting a third definition etc. Van Rooij starts (in the second section) by explaining the meaning of questions as well as the entailment relation between questions within the framework of Groenendijk and Stokhof's partition semantics. In the third section, he first discusses Groenendijk and Stokhof's semantic comparison of (relevant) answers and questions and then observes that the state of information of a questioner influences the relevance of questions and answers. This observation leads to a pragmatic comparison of questions relevant w.r.t. an information state K as well as a comparison of relevant answers to a question w.r.t. K. Van Rooij ends this section by pointing out that the qualitative notion of relevance in the pragmatic comparions is too rough, and that the partial ordering relations between questions and answers should be extended to total orderings by measuring the informativity and relevance of answers and questions in a quantitative way. The next fourth section sets out to explain how this could be achieved. Van Rooij follows the lead of Carnap and Bar-Hillel by defining the informational value of an answer (a proposition) $A$, $\inf(A)$, as the negative logarithm (base 2) of its probability. As the inf-function is monotone increasing w.r.t. the entailment relation between propositions, the total ordering relation induced by the inf-function is exactly an extension of the partial ordering induced by the entailment relation. He then defines the informational value (or entropy) of a question in a formally analogous way to Shannon's definition of the entropy of a coding system as the average informational value of its answers. This definition allows likewise to extend the partial ordering on questions mentioned earlier to a total ordering. Let $B$ be the question (partition) an answer to which provides total information about the world; $B$ has a certain entropy. Van Rooij now defines the informational value of an answer $q$ to question $B$ as the reduction of entropy of $B$ upon learning $q$, and the informational value of question $Q$ w.r.t. question $B$ as the average reduction of entropy of $B$ upon learning an answer to $Q$. As soon as question $B$ is replaced by a mutually exclusive and exhaustive set of hypotheses $H$ (a partition), the use of Shannon's conditional entropy becomes unavoidable. The informational value of question $Q$ w.r.t. question $H$ serves to define the informational usefulness of $Q$ w.r.t. $H$, which in turn is important when an agent is faced with the decision problem which of the hypotheses to choose. Something analogous can be done for answers. But then van Rooij points out the limitations of this approach: the measure of usefulness of questions and answers w.r.t. a decision problem is reasonable only for those kinds of decision problem where the decisions depend alone on the probabilities involved. As soon as desirabilities or utilities influence the actions to be chosen, the approach followed up to now is not appropriate any longer. To take into consideration not only probabilities, but also desirabilities is the approach presented in the fifth and last section. Suppose an agent has to deal with a decision problem. W.r.t. assertions (answers), van Rooij distinguishes between the highest expected utility according to the original decision problem and the utility value of making an informed decision conditional on learning a certain assertion, and defines the utility value of the assertion by the difference between these two values. The expected utility value of a question can then be defined in terms of the utility values of the possible answers. He carries on quoting a theorem according to which measuring the expected utility value of a question w.r.t. a decision problem corresponds to the qualitative 'measurement' of the resp. question. It might now be expected that something similar holds w.r.t. assertions. This, however, is not the case: the utility value of an assertion or answer resp. does not only not behave monotone increasing w.r.t. the entailment relation between propositions, it also doesn't behave monotone increasing w.r.t. the informational value of an assertion w.r.t. a set of hypotheses. The last subsection's starting point is the observation that not only is there in general no connection between the utility value of an assertion and the informational value of an assertion w.r.t. a set of hypotheses, there is in general no connection between the expected utility value of a question and the expected informational value of a question (w.r.t. the most fine-grained partition) either.

J. Seligman starts in his article **Channels: from Logic to Probability** from the assumption that information arises in conditions of uncertainty: uncertainty is reduced by gaining information. Essential for any mathematical model of information is the representation of a state of uncertainty and the change of state induced by the acquisiton of a piece of information. Probability theory provides one such model: Shannon showed how this theory can be used to give a precise measure of uncertainty and to model the movement of information in a system of communication channels. Dretske tried to extend Shannon's model to an information-based semantics and epistemology, which was developed by Barwise, Perry and others as 'situation semantics' and 'situation theory'. Formal logic provides another such model: Barwise and Seligman worked out an account of information flow using a more abstract model of channels and based on formal logic. This account is called the Barwise-Seligman theory. Seligmans aim is to adapt the Barwise-Seligman theory in order to give a similarly abstract account of Dretske's conception of information. The Barwise-Seligman theory of information and information flow makes use of various structures called 'classifications', maps between classifications called 'infomorphisms' and combinations of infomorphisms called 'channels'. In sect. 1, Seligman presents all these basic structures as well as different types of channel related to different fields. If two classifications combined by two infomorphisms satisfying certain conditions are probability spaces, the resulting binary channel is called a 'Shannon channel'; if the two classifications similarly combined by two infomorphisms are formal language classifications, the resulting channel is called a 'Tarski channel'; and if the tokens of classifications are actual concrete events rather than possible configurations, the channel constructed from them is called a 'concrete event channel'. After a brief review in sect. 2 of Shannon's definition of information flow in Shannon channels and Dretske's information-based account of knowledge and belief, Seligman points out a structural similarity between information flow in Shannon channels and information flow in Tarski channels, but he also demonstrates i) that this similarity cannot be formulated within the Barwise-Seligman theory in terms of strong information flow: the infinite Shannon channels elude this attempt (the Strength Problem); and ii) that the model underlying this similarity cannot simply be adapted to information flow in concrete event channels (required for Dretske's epistemological project)(the Modality and the Context Problem). To solve the Modality and the Context Problem, Seligman needs on the one hand a suitable 'linking relation' between sets of types in the core of a concrete event channel to model the regularities on which information flow depends, and on the other hand a suitable set of 'normal tokens' to characterise the contextual connections between particular events: this linking relation and this set of normal tokens are used to define the concept of link on that classification, and ultimately to define information flow relative to a link. At the level of types and tokens this means: There is information flow relative to a link if both components are given: a receiver event type 'indicates' a source event type, and a particular receiver event 'signals' a particular source event within a (core of a) channel C. Sect. 3 serves to determine the value of the link, introduced in the previous sect. The ultimate philosophical goal is to actually find a definition of information flow relative to a link determined by any theory whatsoever, while avoiding the 3 just mentioned problems. Now, a set of pairs of subsets of the set of types of a formal language classification A satisfying certain conditions is called a theory or Tarski theory of A. If the classification is of a probability space P, the resp. theory is called a Dretske theory of P. Seligman then axiomatically characterises the so-called 'Gentzen theories' and shows that all Tarski and all Dretske theories are Gentzen theories, and he succeeds in characterising the Tarski theories. He subsequently raises the question whether the relationship between Gentzen, Tarski and Dretske theories is duplicated w.r.t. the links they determine and he answers it in a negative way: the reason being that a link can be determined by more than one theory. In sect.s 4 and 5, Seligman sets out to find a characterisation of the Dretske theories, both axiomatically and situation semantically. In sect. 4, he characterises the Dretske theories as the theories of extensional Barwise structures satisfying the principle of No Countable Mystery. In sect. 5, he discovers a few properties characterising the class of Dretske theories of a probability space P. In the last sect. 6, Seligman calls his analysis of information flow developed throughout the sections 2-5 the signalling/indicating model of information flow, and he compares it to a model presented in his joint book with J. Barwise (1997), which he calls the logic-movement model of information flow. He notes that in this book, information flow is not modelled as a relation between individual types (or tokens) in the source and receiver, but as a movement of local logics around a network of classifications, whereby a local (Tarski) logic is roughly speaking a Tarski theory on a classification A restricted to a subset (of normal tokens) of the set of tokens of A. Local logics on a classification represent information about the regularities within it. Seligman finally observes that movement of local logics yields a more coherent model of information flow in concrete event channels than the signalling/indicating model, which can be seen as a special case.

In his contribution **Modeling Real Reasoning** K. Devlin sets out to develop a mathematical model for real-life logical reasoning analogous to classical formal logic as a mathematical model for formal reasoning in pure mathematics. He starts off by presenting a couple of reasons why such a model cannot consist of an application of classical formal logic or simple modifications thereof. Next, Devlin treats the topic of information which is related to real-life logical reasoning (and also to other forms of reasoning) in the following way: reasoning is a specific and very important form of purposeful information gathering and information processing. In virtue of the following general observations concerning information, namely that information can arise by virtue of regularities in the world, and that anything can be used to represent information, two tasks have to be tackled with: first, provide a precise, representation-free definition of information, and second, study the nature of the regularities whereby things in the world represent information. These two tasks have been the main focus of attention and the main subject of situation theory (or situation-theoretical information theory). Next, Devlin provides a concise and elegant survey of parts of situation-theory. In the following section, he uses situation theory to model real-life logical reasoning. The basic evidential reasoning element in his model is reminiscent of a proposition in the situation-theoretical sense, and the evidential reasoning process of a proof in the formal logical sense. An evidential reasoning process is constituted by a certain number of evidential reasoning elements some of which are the result of basic reasoning steps. Devlin presents and explains several of these basic reasoning steps. By making explicit in the model the features of the context situation that provide direct support for the items of information considered in the reasoning, and by accounting for various aspects of the reasoning process, Devlin's model clearly goes beyond situation theory and makes it possible to obtain a finer-grained analysis of a specific reasoning process than could be obtained by situation theory. Next, Devlin applies his situation-theoretic model of real-life logical reasoning to three special cases, namely to mathematical reasoning, to reasoning from a common source, and to Bayesian reasoning. In the last section, Devlin motivates on the one hand his model against the background of situation theory, and on the other he briefly discusses ways other than understanding and analysing real reasoning processes that his model could be used for.

In his article **One or Many Concepts of Information?** G. Sommaruga carries on the conceptual work already carried out by L. Floridi, but at the same time trying to take stock of the articles on the various formal theories of information. In his first section, he introduces the distinction between ordinary language concepts, informal theoretical concepts and formal theoretical concepts and he applies this distinction to the concept of information and to the title question in particular. The second section consists of applying the conceptual apparatus developed in the first part to the formal theories of information. This application leads up to an information-theoretical analogue **(T)** to Church's Thesis **(CT)**. The remainder of section two is devoted to a philosophical reflection on **(T)** and to an attempt to provide evidence for **(T)**. The third and last section draws a few conclusions from the previous conceptual analyses and considerations: The most appropriate point of view w.r.t. the title question may very well be a centralized (but not a reductionist) one, and adopting such a point of view may also provide some directions for future work on formal theories of information. As I vaguely recall having read in one of Donald Davidson's articles: It's good to know that we won't run out of work.

## References


[Barwise/Seligman (1997)] Barwise, J., Seligman, J.: Information Flow. The Logic of Distributed Systems. Cambridge University Press, Cambridge (1997)

[Flückiger (1995)] Flückiger, F.: Beiträge zur Entwicklung eines vereinheitlichten Informationsbegriffs. Berne, ph.d. thesis, Berne University (1995), `http://splendor.unibe.ch/Federico.Flueckiger`

[Flückiger (1999)] Flückiger, F.: Towards a Unified Concept of Information: Presentation of a New Approach in Hofkirchner, pp. 101–111 (1999)

[Hofkirchner (1999)] Hofkirchner, W.: The Quest for a Unified Theory of Information. In: Proceedings of the Second International Conference on the Foundations of Information Science. Gordon and Breach Publ., Amsterdam (1999)

[Keller (2002)] Keller, P.: Information Flow. Logics for the (r)age of information. Berne University, math. diploma thesis (2002), `http://www.unige.ch/lettres/philo/enseignants/philipp/research/info.pdf`

[Kornwachs/Jacoby (1996)] Kornwachs, K., Jacoby, K.: Information:New Questions to a Multidisciplinary Concept. Akademie Verlag, Berlin (1996)

[Shapiro (2000)] Shapiro, S.: Thinking About Mathematics. The Philosophy of Mathematics. Oxford University Press, Oxford (2000)