This is a post-print version formatted by the author of a chapter published in the *Routledge Handbook of Philosophy of Information* (2016), pp.~164-179.

The Philosophy of Mathematical Information

MARCELLO D'AGOSTINO

1. Introduction: is there any content to this chapter?

It is somewhat ironical that, for some philosophically sophisticated readers, the very topic of this chapter, the concept of mathematical information, should be empty. Indeed, according to a time-honored philosophical tradition, there is no such thing as "mathematical information". This tradition, which started with the Vienna Circle and later became known as "logical neopositivism" or "logical empiricism" or also "logical positivism", maintained that the trademark of logical and mathematical statements, which separates them from all the other scientific statements, is that they are "tautological" and convey no information. This view was one of the basic tenets of the Vienna Circle, as clearly stated in the manifesto of logical neopositivism:

> The conception of mathematics as tautological in character, which is based on the investigations of Russell and Wittgenstein, is also held by the Vienna Circle. It is to be noted that this conception is opposed not only to apriorism and intuitionism, but also to the older empiricism (for instance of J.S. Mill), which tried to derive mathematics and logic in an experimental-inductive manner as it were (Hahn et al., 1973, p. 311).

According to Rudolf Carnap, one of the most eminent members of the Vienna Circle, there is a fundamental distinction between the empirical sentences and the "mathematico-logical" ones. The former are "real sentences", they are "synthetic" and have content, while the latter are "analytic", have no content and are "merely formal auxiliaries" (Carnap, 1937, p. xiv). Unlike genuine informative sentences, mathematico-logical sentences cannot possibly be false, unless they are formally contradictory, simply because they "say nothing" about the world (Wittgenstein, 1961, 4.461). They are true or false in virtue of the rules of language that fix the meaning of the words, not in virtue of fact (Carnap, 1937, p. 116).

This position was intended to reconcile the basic tenet of empiricism — experience is the *only* source of knowledge — with the fundamental claim that logic and mathematics are infallible. In this way it appeared to provide a satisfactory explanation of logic and mathematics that avoided Kant's appeal to the creative power of "pure reason" as well as the excessive empiricism of John Stuart Mill, which was seen as undermining the necessity and certainty of mathematical truths. Despite its philosophical appeal, however, it is out of question that the idea that mathematics is utterly uninformative clashes with the layman's intuition and, as J.S. Mill would have put it, "a person must have made some advances in philosophy to believe it".[1]

Suppose that next week someone comes up with a proof that settles one of the six still unsolved millennium problems, such as the "$P = NP$?" question or the Riemann hypothesis. It would be very hard to say that a solution to one of them would not be informative: anyone will remain uncertain about the solution until a proof is obtained. So, recalling "the old important idea that information equals elimination of uncertainty" (Hintikka, 1973, pp. 228–229) we cannot help admitting that a positive or negative answer would provide us with genuinely new information, although it would not be a case of empirical discovery. Thus, the fundamental question we address in this chapter is: How can we vindicate the layman's intuition? In what sense, if any, is mathematics informative?

2. A minimal answer to a big question

We can already hear the more analytically oriented reader complain that, before starting our investigation, we should first clarify what we mean by "information". Unfortunately, taking this question seriously would amount to a show stopper. We cannot but agree with Floridi that "this is the hardest and most central problem in the philosophy of information", and that "information is still an elusive concept" (Floridi, 2011, p. 30). Although such

---

[1] Sometimes "men fly to so paradoxical a belief to avoid, as they think, some even greater difficulty, which the vulgar do not see" (Mill, 1882, 315).

a clarification is of paramount importance — and we hope that this Handbook will substantially contribute to it — we also maintain that a good deal of interesting problems can be fruitfully discussed independently of a fully satisfactory theory on the nature of information. Therefore, we shall adopt, as a working hypothesis, a *minimal* notion of information that is precise enough to support philosophical investigation and, at the same time, reasonably close to the layman's intuitive notion.

We take as our starting point the *operational* view that, whatever its nature may be, information manifests itself in an agent's disposition to *answer questions*. The fact that I hold information about Luciano Floridi's mobile number manifests itself in my disposition to answer the question "what is Luciano's mobile number?" Conversely, lack of information manifests itself in an agent's disposition to *abstain* from answering (or, equivalently, answering "I don't know"). It may well be that the answer is not immediately available to me, but that I have a *procedure* for obtaining it with relatively little effort and consumption of resources. For example, it may be that I can't remember Luciano's number, but I can quickly retrieve it from the contact list stored in my own mobile, in which case it would certainly sound bizarre to answer "I don't know". On the other hand, if I have no easy access to Luciano's number, I would answer "I don't know" even if there is an effective procedure that would allow me, *in principle*, to find it out, for example by dialling all the possible 10-digit numbers (assuming that 10 is the usual number of digits of a mobile number). What counts then is whether or not we possess not only an effective procedure that allows us to answer a question "effortlessly", but a *feasible* one that provides an answer within *bounded resources*. Reasonable bounds may then depend on (i) the resources available to us, (ii) the resources required to understand the question and, (iii) maybe also the "value" of the information in question. The existence of such a procedure ensures that the answer is always *practically* available for further use (including further question answering), and so is not informative, while its absence implies that the answer conveys genuinely new information.

Although admittedly vague, this notion appears to be firmly rooted in the ordinary usage of the word "information" and we can safely adopt it as a first approximation to an operational definition. Of course, there is no claim that this is the only reasonable operational notion of information, and not even that it does not require further analysis (it does), but only that it is a reasonable minimal notion that considerably overlaps with ordinary usage, complies with its operational connotations and is good enough for our philosophical task in this chapter.

3. Are the axioms of a mathematical theory really uninformative?

If we focus on axiomatized mathematical theories, like Euclidean Geometry or Peano Arithmetic, the claim that such theories are (not) "tautological" can be analyzed into two distinct components: (i) the axioms are (not) tautological; (ii) the inference process that leads from the axioms to the theorems is (not) tautological. (Typically, in mathematics the inference process is assumed to be some form of logical deduction.) In the context of this chapter by saying that a *sentence* is "tautological" we mean that it carries no information, while by saying that an *inference* is "tautological" we mean that its conclusion carries no *new* information with respect to its premises.[2]

According to J.S. Mill's radical empiricism, logical and mathematical axioms are not tautological for the simple reason that their nature is just as empirical as that of any other scientific truth:

> It remains to inquire, what is the ground of our belief in axioms — what is the evidence on which they rest? I answer, they are experimental truths; generalizations from observation. (Mill, 1882, p. 286).

As mentioned above, Carnap and the other members of the Vienna Circle found this view quite unpalatable. First, it appeared to derogate from the certainty of logic and mathematics. Second, as Carnap himself explains in his intellectual autobiography, the Viennese were strongly influenced by Frege and Russell's opposite view that "all mathematical concepts can be defined on the basis of the concepts of logic and that the theorems of mathematics can be deduced from the principles of logic" (Carnap, 1963, p. 46).[3] In

---

[2] Thus, we do not use "tautological" in its technical sense of "being true in all possible worlds". The two senses coincide if Bar-Hillel and Carnap's notion of semantic information (see Chapter six of this Handbook) is adopted.

[3] Indeed, this is not quite accurate a picture, since as Michael Dummett puts it: "Frege attempted to show that some mathematical propositions, those of number theory and analysis which he jointly

Frege's view, the axioms of a mathematical theory are "self-evident", in the sense that they are "beyond a reasonable doubt by someone who fully understands the relevant propositions" (Burge, 1998, p. 312). Here "self-evident" does not mean "obvious" in a psychological sense, but something that justifies itself, i.e., requires no other grounds to be recognized as true. Occasionally, Frege conceded that correctly recognizing the self-evidence of some propositions may be somewhat demanding for a less-than-ideal agent.[4] From this point of view, a self-evident proposition may not be obvious, and can be perceived as informative by non-ideal agents. Conversely, some propositions may seem obvious to non-ideal agents without being self-evident or even true. These remarks were dramatically confirmed by the discovery of Russell's paradox (1901), which showed that Frege's Axiom V of his *Grundgesetze der Arhitmetik* (1893) was so far from being self-evident as to be inconsistent.

In any case, Frege's view suggested to the neopositivists what Imre Lakatos would have called "a progressive problemshift", from Kant's old question "how is pure mathematics possible" to the more accessible question "how is logical certainty possible"? They thought that the answer to this new question was to be found in Wittgenstein's view "that all logical truths are tautological, that is they hold necessarily in every possible case, therefore they do not exclude any case, and do not say anything about the facts of the world" (Carnap, 1963, p. 46). So, a typical logical truth such as "it rains or it does not rain" conveys no information: "I know nothing about the weather, when I know that it rains or does not rain" (Wittgenstein, *Tractatus*, 4.461). In this spirit, at the half of the 20th century Bar-Hillel and Carnap's theory of "semantic information"[5] – which was meant to provide a semantic counterpart to Shannon and Weaver's purely syntactic theory of information – closed the circle, by providing what is, to date, the most comprehensive theoretical framework for Wittgenstein's thesis. Roughly speaking, this theory simply identifies the information conveyed by a sentence with the set of all "possible worlds" that are excluded by it. Bar-Hillel and Carnap showed also how this theory can be extended to first-order languages, i.e., to

---

classified as 'arithmetic', had the same character as, and in fact were, logical propositions; he never believed this to be true of the whole of mathematics" (Dummett, 1991, p. 10).

[4] On this point see (Burge, 1998, §IV).

[5] See Chapter six of this Handbook. See also (D'Agostino, 2013b) for a discussion.

languages containing the usual quantifiers "for all" and "for some" as well as the standard Boolean operators. A straightforward consequence of Bar-Hillel and Carnap's theory is that all logical truths of the theory of quantification are equally uninformative (they exclude no possible world).[6]

4. Is deduction really uninformative?

Whatever position one may hold about the axioms of a mathematical theory, the problem of mathematical information remains at center-stage in the following form: is the information carried by a theorem really *contained* in the information (if any) carried by the axioms? Even if the axioms are regarded as informative, the conditional "If $A_1, \ldots , A_n$, then T", where the axioms $A_1, \ldots , A_n$ are incorporated as hypotheses in the antecedent, is itself a logical truth. Now, if logical truths are tautological, deductive reasoning "cannot lead to new knowledge but only to an explication or transformation of the knowledge contained in the premises" (Carnap, 1963, p. 46), which again clashes with the fact that many mathematical theorems do provide us with information that does not seem to be *available* just by grasping the axioms. Should we just dismiss this intuitive difficulty as philosophically naive, and content ourselves with claiming that the information carried by T was indeed already available, albeit implicitly, in the axioms? Or should we take it seriously and look into the possibility that deductive reasoning may, under certain conditions, generate information that is genuinely new?

The first position was strongly supported by most members of the Vienna Circle. Moritz Schlick, for example, in his *Allgemeine Erkenntnislehre* remarked that "exact inference is often credited with making a greater contribution than is within its power" (Schlick, 1974, p. 108). By contrast, he insisted that "the conclusion of a syllogism never contains knowledge that is not already assumed as valid in the major premiss or, perhaps, in both premisses of the inference" and that the syllogism "is not an instrument by

---

6 Another inevitable and undesirable consequence is that contradictions, like "it rains and it does not rain", carry the maximum amount of information, since they exclude all possible states. This is usually referred to as the "Bar-Hillell and Carnap Paradox". See Chapter six of this Handbook.

which new knowledge can be procured" (Schlick, 1974, p. 108). Indeed, a complex deduction may well have a great *psychological value*. "There may of course be cases where the conclusions of syllogistic procedures, say the results of some calculation, do astonish us and present us with unexpected findings." But this "shows merely that psychologically the final outcome was not conceived of along with the major premiss" and "does not mean that the end result is not contained logically in the major premiss" (Schlick, 1974, p. 111).

Despite being at odds with the layman intuitive judgment, this view that deductive reasoning is "objectively" uninformative caught on and became part of the logical folklore. In fact, it was not just a philosophical oddity of the neopositivists, but an idea firmly rooted in modern philosophy that can be found in such diverse authors as Bacon, Descartes, Kant and Mill. In his *Grundlagen*, Frege referred to it as to "the legend of the sterility of pure logic" (Frege, 1960, p. 24) and eighty years later Jaakko Hintikka still called it a true "scandal of deduction":

> C.D. Broad has called the unsolved problems concerning induction a scandal of philosophy. It seems to me that in addition to this scandal of induction there is an equally disquieting scandal of deduction. Its urgency can be brought home to each of us by any clever freshman who asks, upon being told that deductive reasoning is "tautological" or "analytical" and that logical truths have no "empirical content" and cannot be used to make "factual assertions": in what other sense, then, does deductive reasoning give us new information? Is it not perfectly obvious there is some such sense, for what point would there otherwise be to logic and mathematics? (Hintikka, 1973, p. 222).

This scandal has recently received renewed attention leading to a number of original contributions (e.g., Primiero 2008, Ch. 2, Sequoiah- Grayson 2008, Sillari 2008, D'Agostino & Floridi 2009, Duzí 2010, Jago 2013, Allo & Mares 2012). In the next three sections we shall illustrate the attempts made by distinguished philosophers such as Mill, Frege and Hintikka to account for an objective and non purely psychological sense in which deductive reasoning, the only allowed kind of inference in mathematics, can indeed be informative.

5. Mill on the inductive nature of deductive reasoning

According to Mill a real inference is always "a progress from the known to the unknown: a means of coming to the knowledge of something we did not know before" (Mill, 1882, p. 227). Is then deductive reasoning ("ratiocination") a real inference? The answer, for Mill, is that what is usually called "deduction" is nothing but induction in disguise. The following inference:

All men are mortal
The Duke of Wellington is a man
―――――――――――――――――
The Duke of Wellington is mortal

is just a device to "decode" the inductive information, based on past experience, that had been previously encoded in the general statement "All men are mortal". The allegedly "deductive" step does nothing but extending this empirical knowledge to a new case (Mill, 1882, p. 224).

> When, therefore, we conclude from the death of John and Thomas, and every other person we ever heard of in whose case the experiment had been fairly tried, that the Duke of Wellington is mortal like the rest; we may, indeed, pass through the generalization, All men are mortal, as an intermediate stage; but it is not in the latter half of the process, the descent from all men to the Duke of Wellington, that the inference resides. The inference is finished when we have asserted that all men are mortal. What remains to be performed afterward is merely deciphering our own notes (Mill, 1882, p. 232).

So, every mathematical inference is, in essence, inductive and "the general proposition on which it is said to depend, may, in certain cases, be altogether omitted, without impairing its probative force" (Mill, 1882, p. 237). A mathematical proof can always be described as a sequence of steps from particular truths to particular truths, each of which is grounded on experience. This is apparent in the ordinary process of proving a geometrical theorem by means of a diagram, where we reason on a *specific case*, by performing a thought-experiment, as it were, whereby we perceive that the same reasoning process could be indefinitely replicated in similar conditions (see Mill, 1882, p. 238). However, although each single step is ultimately based on experience, in mathematics this experience is so "superabundant" (Mill, 1882, p. 310) that the steps appear to be "obvious". Thus, the appeal to induction, by itself, fails to explain the apparent informativity of mathematical reasoning. Mill is aware of this difficulty:

> It might seem to follow, if all reasoning be induction, that the difficulties of philosophical investigation must lie in the inductions exclusively, and that when these were easy, and susceptible of no doubt or hesitation, there could be no science, or, at least, no difficulties in science. The existence, for example, of an extensive Science of Mathematics, requiring the highest scientific genius in those who contributed to its creation, and calling for a most continued and vigorous exertion of intellect in order to appropriate it when created, may seem hard to be accounted for on the foregoing theory (Mill, 1882, p. 267–268).

This brings us back to square one. What is, then, the source of mathematical information? According to Mill, an answer can be found by considering that, in most interesting cases, a reasoning process does not consist in a single chain of obvious inductive steps but in the combination of several such chains that converge to a conclusion. Hence, the mystery of mathematical information is dissipated simply by observing that "even when the inductions themselves are obvious, there may be much difficulty in finding whether the particular case which is the subject of inquiry comes within them; and ample room for scientific ingenuity in so combining various inductions, as, by means of one within which the case evidently falls, to bring it within others in which it can not be directly seen to be included" (Mill, 1882, pp. 267–268). To use the working operational definition of Section 2, we might say that a mathematical theorem is informative when, because of the unforeseeable complexity of the reasoning process, we have no feasible procedure to obtain it within reasonably bounded cognitive/computational resources.

6. Frege on the ampliative power of deduction

As is well known, Frege totally rejected Mill's "preconception that all knowledge is empirical" and severely criticized Mill's inductivist view of logic and mathematics (see Frege, 1960, §§7–10). However, like Mill, he maintained that deductive reasoning is informative. Interestingly enough, his main argument for this claim was based on the very same premise — namely the thesis that arithmetic can be reduced to logic — that, later on, led the neopositivists to claim that mathematical statements have no information content. While the neopositivists used this premise together with Wittgenstein's idea that logic is tautological to conclude that arithmetic is also tautological, Frege used it together with the intuitive judgement that arithmetic is not tautological to conclude, by *modus tollens*, that logic is not tautological either. Given the reduction of arithmetic to logic, he believed that:

> the prodigious development of arithmetical studies, with their multitudinous applications, will suffice to put an end to the widespread contempt for analytic judgments and to the legend of the sterility of pure logic (Frege, 1960, p. 24).

Notice that from this passage it is clear that Frege does not use the word "analytic" as synonymous with "tautological" in the sense of "uninformative" (on this point see also footnote 12 in Section 10 below), but in the foundational sense of being grounded on purely logical truths. Indeed, it is clear enough from §23 of the *Grundlagen* that Frege regarded arithmetical truth as informative albeit being "analytic" in this sense. Then, the same question keeps coming up: "Can the great tree of the science of number as we know it, towering, spreading and still continually growing, have its roots in bare identities? How do the empty forms of logic come to disgorge so rich a content?" (Frege, 1960, p. 22). Frege's answer is that the informativeness of deductive reasoning stems from a single source, namely the labor required to extract remote conclusions from the premises. In deductive reasoning the conclusions are contained in the premises "as plants are contained in their seeds, not as beams are contained in a house" (Frege, 1960, p. 101). The growth of a plant from seed to tree is indeed a resource-consuming process. As explained by Dummett (Dummett, 1991, Chapter 4), deduction is for Frege a creative process that cannot be reduced to a mechanical one. Its creative aspect lies in the discernment — by a process of dissection — of suitable patterns that are hidden in judgments and may be used to connect them in a deductive chain. When quantifiers are involved, this creativity is necessary even to carry out very basic steps, such as inferring the conclusion:

(1) There is a body such that, either Jupiter is larger than it and it is larger than Mars, or Mars is larger than it and it is larger than Jupiter.

from the premise

(2) Either Jupiter is larger than Neptune and Neptune is larger than Mars, or Mars is larger than Neptune and Neptune is larger than Jupiter.

This amounts to what we call an introduction of the existential quantifier, which is a primitive rule in Gentzen's "natural deduction".[7] The creative act here is that of discerning, in the premise (2) the complex predicate "either Jupiter is larger than *x* and *x* is larger than Mars, or Mars is larger than *x* and *x* is larger than Jupiter" as a pattern that allows one to see that the conclusion follows from the premise. As Dummett points out, this discernment is, in fact, preliminary to the formation of the quantified proposition that, in Frege's own language, represents (1), namely:

(3) For some x, either Jupiter is larger than x and x is larger than Mars, or Mars is larger than x and x is larger than Jupiter.

As Dummett puts it:

> This predicate is not a component of the proposition it was extracted by dissection, in that we do not have to recognize its presence in order to grasp the content of the proposition; but it is a component of the quantified proposition (Dummett, 1991, p. 42).

Although, in some sense, the predicate "is there" to be discerned in (2), its discernment involves an act of pattern recognition, which is sometimes rendered complex by the fact that a single sentence may contain several such patterns, depending on how it is "dissected" into "functions" and "arguments". For example, consider the sentence "Cato killed Cato". Then:

> If we here think of "Cato" as replaceable at its first occurrence, "to kill Cato" is the function; if we think of "Cato" as replaceable at its second occurrence, "to be killed by Cato" is the function; if, finally, we think of "Cato" as

[7] For an excellent exposition of natural deduction see (Tennant, 1990).

> replaceable at both occurrences, “to kill oneself” is the function (Frege, 1967, p. 22).

Thus, for Frege constructing a deduction, unlike checking it for correctness, is not a mechanical process, but has “a creative component involving the apprehension of patterns within the thought expressed and relating them to one another, that are not required for a grasp of those thoughts themselves” (Dummett, 1991, p. 42). This is again in tune with our working definition of information given in Section 2. The informativity of deduction stems from the fact that we don’t have a practical procedure to construct a proof in a quantificational language, and that this activity requires considerable, and in general unpredictable, effort in order to recognize those patterns, among the many that can be discerned in the sentences that enter the proof, which are shared by other sentences and can therefore be fruitfully used to connect them. This seems to solve the paradox of deduction, namely the paradox of how deduction can, at the same time, be valid and fruitful:

> Since [deduction] has this creative component, a knowledge of the premises of an inferential step does not entail a knowledge of the conclusion, even when we attend to them simultaneously; and so deductive reasoning can yield new knowledge. Since the relevant patterns need to be discerned, such reasoning is fruitful; but since they are there to be discerned, its validity is not called in question (Dummett, 1991, p. 42).

In conclusion of this section, we observe that (i) Frege’s explanation of the source of deductive information applies only to reasoning in quantificational languages and does not apply at all to propositional logic; (ii) it, admittedly, applies to the process of “dissection” of sentences and concept-formation, which is preliminary to the specification of sentences in a formalized quantificational language; when the expressive resources are specified from the outset, it fails to capture the labor that is anyway needed to obtain the conclusion from the premises. This is now particularly evident in light of the then unknown undecidability of logical consequence in quantificational languages.

7. Hintikka on the informativity of quantification logic

In 1935–1936 we learned, as a result of independent investigations of Alan Turing and Alonzo Church, that the logic of quantification is undecidable, that is, there is no mechanical procedure that allows us to answer "yes" or "no" to all questions of the form "is ϕ a logical consequence of $\psi_1,\ldots,\psi_n$?" when $\phi, \psi_1,\ldots, \psi_n$ are sentences of a fixed quantificational language containing at least one non-monadic predicate (e.g., a binary relation). The completeness theorem proved by Kurt Gödel (1929) ensured that there is, however, a semidecision procedure. Since there are proof systems (e.g., those based on the axioms and inference rules given by Whitehead and Russell in their *Principia Mathematica*) that allow us to prove all valid inferences, we do have a trivial procedure — namely that consisting of enumerating all possible proofs — that always allows us to answer "yes" after a finite number of steps (as soon as we hit a proof of the conclusion from the premises) whenever the answer is "yes". But there is no mechanical procedure that allows us to answer "no" whenever the answer is "no". If there were such a procedure we would be guaranteed to obtain always a yes-or-no answer after a finite number of steps, simply by running the two procedures simultaneously. But this is exactly what is ruled out by the undecidability theorem. Thus, if the answer is "no", it might be the case that we never come to know it, although in some cases we may be able to find a counterexample. But, even if the answer is "yes" there is no guarantee that we can obtain it within *bounded resources*, for there cannot be any general upper bound on the time required to obtain the answer as a function of the input question.[8] So, if we run our semidecision procedure on a supercomputer and the answer has not come after a long time, this does not allow us to conclude that the answer is "no", since there is still a chance that a positive answer will come up tomorrow. We must therefore admit that there are cases in which we do not *actually* possess the information carried by the conclusion of a valid inference, even if we possess the information carried by the premises. So, while "objectively" — according to the theory of semantic information — we already possess the information carried by the conclusion, "subjectively" we may not know that we possess it.

This has a strong paradoxical flavour. What kind of notion of

[8] If there were such an upper bound, the logic of quantification would be decidable, since the answer would be "no" whenever no answer has been obtained within the upper time limit.

information is one that does not allow us, in general, to know whether we possess information about something? This kind of information is not fully manifestable and has no clear *operational* value. Hintikka's positive proposal consists in distinguishing between two objective and non-psychological notions of information content: *surface information*, which may be increased by deductive reasoning, and *depth information* (equivalent to Bar-Hillel and Carnap's semantic information), which may not. While the latter justifies the traditional claim that logical reasoning is tautological, the former vindicates the intuition underlying the opposite claim. In Hintikka's view, quantificational deductive reasoning with polyadic predicates may increase surface information, although it never increases depth information. In the rest of this section we shall give the reader a flavour of Hintikka's distinction without even attempting to do proper justice to its technical content.[9]

According to Hintikka, in a quantificational language every quantifier, as well as every individual constant, "invites" us to consider a certain individual in relation with other individuals. For example:

- $(\exists x)$ *loves*$(x, Mary)$ ("someone loves Mary") invites us to consider an unspecified individual, call it a, such that *loves*$(a, Mary)$;
- $(\forall x)$ *loves*$(x, Mary)$ ("everyone loves Mary") invites us to consider *an arbitrary* individual in his relation to Mary, say an individual b that is somehow representative of all individuals in the domain and tells us that *loves*$(b, Mary)$

Despite the universally quantified sentence appearing to invite us to consider all the individuals that inhabit the domain in their relation to Mary, clearly we do not (and sometimes cannot, when the domain is unspecified or infinite) review in our intuition all these individuals and consider their relation to Mary, but just focus on a single, albeit generic, representative. (This is exactly what we do when we use figures to demonstrate a geometrical theorem.) Usually the depth of nesting of quantifiers in a sentence and the number of constants in their scope contribute to the difficulty of grasping the

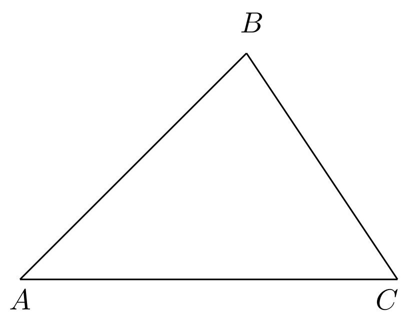
[9] For Hintikka's own account, the reader is referred to (Hintikka, 1965, 1973). For a (very) critical exposition, see (Sequoiah-Grayson, 2008).

meaning of a sentence. For example, $(\forall x)(\exists y)$ (*loves*(x, y) and *loves*(y, *Mary*)) ("everyone loves someone who loves Mary") cannot be properly understood without imagining three individuals mutually related by the dyadic predicate "love"; an arbitrary one instantiating the variable x, a second one, which is also unspecified but depends on the first, instantiating y and the one denoted by "Mary".

To give an informal explanation of Hintikka's ideas, let us consider the geometrical example discussed by Kant in his *Critique of Pure Reason*, namely Euclid's Proposition 32 of the first book of the *Elements*:

(4) In any triangle the sum of the internal angles is equal to two right angles (180°).

This statement invites us to consider an arbitrary triangle *ABC*, its internal angles and their sum:

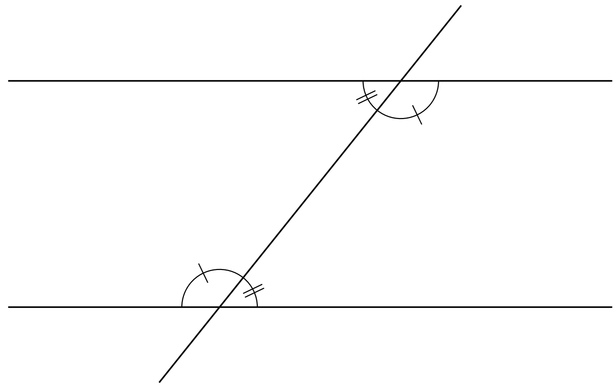


$$\widehat{BAC} + \widehat{ACB} + \widehat{CBA}$$

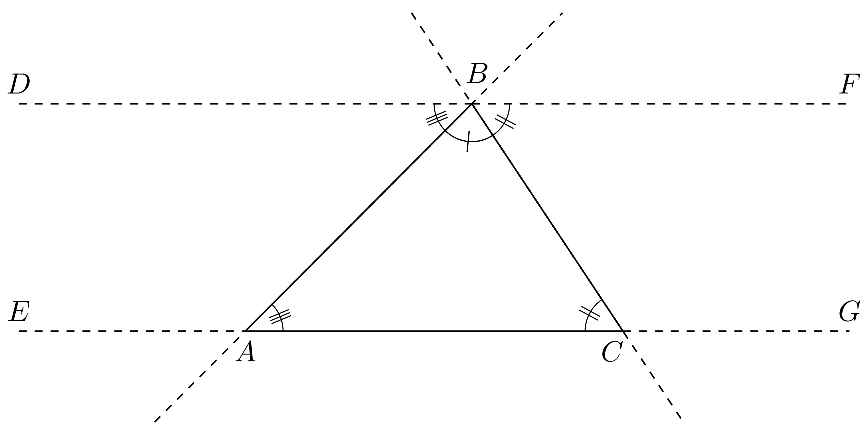
In proposition 29 Euclid had already proven (among other things) that

(5) if two straight lines are parallel, then a straight line that meets them makes the alternate angles equal.

Here we are invited to consider two parallel lines, a third line that meets them and the alternate angles formed thereby:



We can safely claim that we understand the two propositions (4) and (5) taken separately and our understanding is helped by the exhibition of the two diagrams. However, simply understanding them is by no means sufficient to realize that (4) follows from (5). In order to “see” this we have to introduce a more complex diagram that results from an ingenious combination of the previous two:



After this construction has been carried out, we can indeed see that (4) is true. But to achieve this, we need to conceive an aggregate of objects in their mutual relations, in short a configuration of objects, that is significantly more complex than the configurations that were needed to understand either the premise (5) or the conclusion (4).

Roughly speaking, Hintikka defines the surface information conveyed by a sentence ϕ as the set of types of “possible worlds” excluded by ϕ that can be described within bounded expressive resources, without increasing the complexity (number of distinct mutually related individuals) of the configurations that we have to master in order to understand any of the premises or the conclusion. Not surprisingly, given its essentially bounded nature, surface information is effectively computable and provides "an objective sense of information in which deduction can increase one's information — a sense that definitively confutes logical positivists on this point" (Hintikka, 1973, p. 230).

Without going into further details, we observe only that Hintikka's proposal classifies as non-tautological only some inferences of the polyadic predicate calculus, and so leaves the "scandal of deduction" unsettled in the domains of propositional logic and of the monadic predicate calculus.

8. The lesson from computational complexity

The theory of computational complexity can be considered a refinement of the traditional theory of computability taking into account the resources (time and space, that is, number of steps and amount of memory) used by algorithms. Its principal innovation consists in having replaced the concept of "effective procedure" with that of "feasible procedure".[10] An effective procedure or "algorithm" by and large consists in a "mechanical method", i.e. one executable in principle by a machine, to solve a given class of problems (answering a certain class of questions). An effective procedure is *feasible* when it can also be carried out in practice, and not only in principle. The expression "in practice" involves a certain degree of vagueness that computational complexity researchers have removed by agreeing to consider as feasible, executable in practice, algorithms that can answer the questions they are designed to answer within polynomial time. This means that there exists a fixed polynomial p such that, for any input of size n (i.e., that is encoded by a string of n occurrences of symbols), the algorithm yields an answer in a number of steps $\leq p(n)$.

The class P is the class of all *tractable* decision problems, i.e. those that can be solved by means of a polynomial time algorithm. On the other hand, if a problem is *intractable* — i.e., does not belong to P — then, even if

[10] For an excellent exposition, still valid after thirty-five years, see (Garey & Johnson, 1979).

decidable in principle, it is regarded as “undecidable in practice”. The reason is readily explained: if the running time of an algorithm is not bounded above by a polynomial in the size of the input — for example, if it is expressed by an exponential function such as 2^n — its execution may in some cases require a number of steps that grows fast beyond any practical limit as the input’s size grows. In extreme cases, the time required could be longer than the life of the universe.

The bad news is that most interesting decidable problems that one encounters in many branches of mathematics do not belong to *P*. Among these there stands out the problem of establishing whether a certain proposition is a theorem of elementary Euclidean geometry, namely the geometry that can be expressed in a standard quantificational language. Although this problem was shown to be “algorithmically solvable” by Tarski in 1951, twenty-three years later Fischer and Rabin proved that it is, in fact, intractable.[11] According to our minimal notion of information of Section 2, this implies that, in some cases, the information carried by a theorem in elementary geometry is not *actually* contained in the information carried by the axioms — although it is indeed *potentially* contained in it — because there is no feasible procedure to answer the question of whether an arbitrary sentence is true assuming that a yes-answer has been given for all the axioms.

Now, although there is no proof yet that propositional logic is intractable, a very general pessimistic result was obtained by Stephen Cook (1971), who showed that a large class of very difficult computational problems (the so-called *NP-complete problems)*[12] can be translated in polynomial time into the problem of deciding whether a given Boolean sentence is a tautology. Thus, a polynomial time algorithm for the tautology problem would automatically generate a polynomial time algorithm for each of the *NP*-complete problems. Since, despite all efforts, no polynomial time algorithm has ever been found for any of them, the dominant conjecture is that the tautology problem is intractable. If this conjecture is correct, we will never be able to guarantee that an answer to all questions concerning potential consequences of a set of

---

11 See (Rabin, 1977) for this and other intractability results.

12 See (Garey & Johnson, 1979) and (Stockmeyer, 1987) for an introduction to the theory of *NP*-completeness.

assumptions can be actually obtained effortlessly, by means of a *feasible* mechanical procedure, even when these consequences follow by propositional logic only. This strongly suggests that both Frege's and Hintikka's solution of the "scandal of deduction" are not fully satisfactory, for the core of the problem seems to lie at the heart of propositional logic for which a mechanical decision procedure has been known since the 1920's.

9. Actual vs virtual information

This tension between the (presumptive) intractability of propositional logic and its alleged uninformativity is the motivation of a new approach put forward in (D'Agostino & Floridi, 2009) and further elaborated in subsequent papers (D'Agostino et al., 2013; D'Agostino, 2014). In this approach there is a sense in which propositional logic is indeed informative. Here we shall give just a flavour of the underlying ideas. Consider the following simple schematic inferences:

(Modus Ponens)

$$\frac{\begin{array}{l}\phi \rightarrow \psi \\ \phi\end{array}}{\psi}$$

(Disjunctive Syllogism)

$$\frac{\begin{array}{l}\phi \vee \psi \\ \neg\phi\end{array}}{\psi}$$

Here for each instance of the schematic letters, whenever we possess the information carried by the premises, we *immediately* infer the conclusion in virtue of the very meaning of the logical operators with no need for introducing and discharging assumptions (as in Gentzen's natural deduction). (D'Agostino & Floridi, 2009) presents a system of natural deduction consisting only of simple operational rules like these. Moreover, it can be shown that, whenever a conclusion follows from a set of premises by a chain of such simple inferences, there is a feasible procedure to construct such a chain. So, in accordance with our working notion of information of Section 2, we can say that we actually possess the information carried by the conclusion whenever we actually possess the information carried by the premises: the *marginal cost* of inferring the conclusion is negligible with respect to the cost of grasping the meaning of the sentences. All the information that can be extracted from the premises by means of these rules is *actual information*.

Now, not all arguments in propositional logic can be fully justified in this way. Consider the following inference:

(1) If n is an integer, n is either even or odd
(2) If n is an even integer, then there is an integer m such that $n = 2m$
(3) If n is an integer, then $2n$ is an even integer
(4) n is even if and only if $n + 1$ is odd
(5) n is an integer

$n^2 + 3n + 5$ is odd.

Here is a simple "proof by cases". By (5) n is an integer and so, by (1) n is either even or odd.

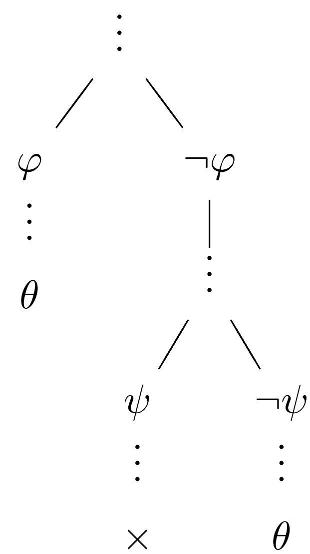
Case 1: n is even. Then, by (2) and Modus Ponens, n = $2m$ for some integer m. So: $n^2 +3n+5 = (2m)^2 +3(2m)+5 = 2(2m^2 + 3m + 2) + 1$. Since by (3) and Modus Ponens $2(2m^2 + 3m + 2)$ is even, then by (4) and Modus Ponens $n^2 + 3n + 5$ is odd.

Case 2: n is not even. So, by Disjunctive Syllogism n is odd. Then, by (2), (4) and Modus Ponens $n=2p+1$ for some integer p. So, $n^2 + 3n + 5= (2p + 1)^2 + 3(2p + 1) + 5 = 2(2p^2 + 5p + 4) +1$. Since, by (3) and Modus Ponens $2(2p^2 + 5p + 4)$ is even, then $n^2 + 3n + 5$ is odd. Since either n is even or it is not, and in either case $n^2 + 3n + 5$ is odd, it follows that the conclusion is true independent of our being informed of whether n is even or not.

□

Here, in each case, we simulate information that we do not actually possess. This is what we call *virtual information*. We pretend we have an answer to the question of whether or not n is even and check what the consequences of each alternative are by means of simple inferences such as Modus Ponens or

Disjunctive Syllogism. In each case we simulate an information state that is essentially richer than the actual one, containing information that is not even implicitly contained in it. Inference patterns of this kind, that essentially involve the manipulation of virtual information require, for each realistic agent, an *objective* effort that increases with the *depth* at which virtual information is introduced. In general, the propositional structure of an argument making use of virtual information is the following:



where the conclusion θ occurs in all branches that are not explicitly inconsistent (marked with ×). Each branching increases the complexity of a logical inference from both the cognitive and the computational point of view and, therefore, also its informativity. In the papers cited above it is shown that if we fix an upper bound *k* on the depth at which virtual information can be introduced, the relations of *k*-depth consequence are all tractable, albeit requiring increasing computational resources, and converge to classical propositional logic.

10. Conclusions: a Kantian view?

One of the trademarks of modern, or "logical", empiricism was the rejection of the possibility of "synthetic a priori" judgments in Kant's sense. While the precise sense attached by Kant to the opposition analytic/synthetic is not our concern here, there is very little doubt that in his view synthetic judgments — and by extension synthetic inferences — are capable of extending our knowledge, and are therefore informative, while analytic judgments and inferences are not.[13] Kant also thought that all analytic judgments or inferences are *a priori*, independent of experience, and that while synthetic judgments or inferences are usually *a posteriori*, in that they do depend on experience, in some important cases — such as those of mathematical statements — they may also be a priori.

The neopositivists' opposition to this Kantian view was clearly stated in the *manifesto* of the Vienna Circle as "the basic thesis of modern empiricism":

> The scientific world-conception knows no unconditionally valid knowledge derived from pure reason, no "synthetic judgments a priori" of the kind that lie at the basis of Kantian epistemology and even more of all pre- and post-Kantian ontology and metaphysics. [...] It is precisely in the rejection of the possibility of synthetic knowledge a priori that the basic thesis of modern empiricism lies. The scientific world-conception knows only empirical statements about things of all kinds, and analytic statements of logic and mathematics. (Hahn et al., 1973, p. 308).

In his (1973) Hintikka presented his view on the informativity of quantification logic, that we have briefly described in Section 7, as a vindication of Kant's old idea. Some laws or inferences of quantification logic are synthetic in that they cannot be obtained by simply analyzing the configurations of objects that are "given" with the very sentences by means of which they are expressed.

In the same vein, one could say that, in our discussion of propositional logic in Section 9, analytical inferences are those that are recognized as sound via basic logical steps that process *actual* information, such as Modus ponens and Disjunctive Syllogism, while synthetic ones involve the intuitive

[13] The reader is warned that the expressions "analytic" and "synthetic" have been used in several different meanings, albeit related by a "family resemblance", in the philosophical literature. See (Hintikka, 1973) for a detailed discussion of a variety of different meanings. Here we only stress that, according to Frege's use of the term "analytic" in (Frege, 1967), analytic judgments may well be informative. See also Section 6 above and (Dummett, 1991).

simulation of *virtual* information that is in no way "contained" in the premises. These two views can be seen as "orthogonal" in that they account for two different dimensions of mathematical information: in one of them, it stems from the complexity of the configurations needed to carry out an inference; in the other, it stems from the depth of the nested patterns of virtual information that are needed to obtain the conclusion. In either view, informativity is crucially linked to the cognitive and computational resources consumed in the inferential process: those needed to master a more complex configuration of individuals in their mutual relations, or to keep track of several sub-processes running in the tree-like structure of an argument by cases. Finally, both views allow for the definition of *degrees of informativity* of a logical inference depending on the maximum complexity of the configurations of objects required to carry it out (for the quantificational case) or on the maximum depth at which the use of virtual information is required (for the propositional case).

References

Allo, P., & Mares, E. 2012. Informational semantics as a third alternative? *Erkenntnis*, 77(2), 167–185.

Burge, T. 1998. Frege on knowing the foundations. *Mind*, 107(426), 305– 347.

Carnap, R. 1934. *Logische Syntax der Sprache*. Wien: Springer.Carnap, R. 1937. *The Logical Syntax of Language*. London: Routledge & Kegan Paul. English translation of (Carnap, 1934) by A. Smeaton.

Carnap, R. 1963. Intellectual autobiography. Pages 3–84 of: Schilpp, P.A: (ed), *The Philosophy of Rudolf Carnap*. La Salle: Open Court.

Cook, S. A. 1971. The complexity of theorem-proving procedures. Pages 151–158 of: STOC '71: *Proceedings of the third annual ACM symposium on Theory of computing*. New York, NY, USA: ACM Press.

D'Agostino, M. 2013. Semantic Information and the trivialization of logic. Floridi on the scandal of deduction. *Information*, 4, 33–59.

D'Agostino, M. 2014. Analytic inference and the informational meaning of the logical operators. *Logique et Analyse*, 227, 407–437.

D'Agostino, M., & Floridi, L. 2009. The enduring scandal of deduction. Is propositionally logic really uninformative? *Synthese*, 167, 271–315.

D'Agostino, M., Finger, M., & Gabbay, D.M. 2013. Semantics and proof theory of depth-bounded Boolean logics. *Theoretical Computer Science*, 480, 43–68.

Dummett, M. 1991. Frege: Philosophy of Mathematics. Duckworth.

Duzí, M. 2010. The Paradox of Inference and the Non-Triviality of Analytic Information. *Journal of Philosophical Logic*, 39, 473–510.

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

Frege, G. 1960. *The Foundations of Arithmetic: a Logico-Mathematical Enquiry into the Concept of Number* (1884). 2nd edn. New York: Harper & Brothers.

Frege, G. 1967. *Begriffscchrift, a formula language, modeled upon that of arithmetic, for pure thought* (1879). Pages 1–82 of: van Heijenoort, J. (ed), *From Frege to Gödel. A Source Book in Mathematical Logic*. 1879– 1931. Harvard University Press.

Garey, M. R., & Johnson, D. S. 1979. *Computers and Intractability. A Guide to the theory of NP-Completeness*. San Francisco: W.H. Freeman & Co.

Hahn, H., Neurath, O., & Carnap, R. 1973. The scientific conception of the world [1929]. In: Neurath, M., & Cohen, R.S. (eds), *Empiricism and Sociology*. Dordrecht: Reidel.

Hintikka, J. 1965. Are logical truths analytic? *The Philosophical Review*, 74(2), 178–203.

Hintikka, J. 1973. *Logic, language games and information. Kantian themes in the philosophy of logic*. Oxford: Clarendon Press.

Jago, M. 2013. The content of deduction. *The Journal of Philosophical Logic*, 42, 317-334.

Mill, J.S. 1882. *A System of Logic Ratiocinative and Inductive*. 8th edn. New York: Harper & Brothers.

Primiero, G. 2008. *Information and Knowledge. A Constructive Type- Theoretical Approach*. Springer.

Rabin, M.O. 1977. Decidable Theories. Pages 595–630 of: Barwise, J. (ed), *Handbook of Mathematical Logic*. Amsterdam: North-Holland.

Schlick, M. 1974. *General Theory of Knowledge*. New York: Springer-Verlag. Translation of the 2nd German edition of *Allgemeine Erkenntnislehre*, Berlin 1925.

Sequoiah-Grayson, S. 2008. The Scandal of Deduction. Hintikka on the Information Yield of Deductive Inferences. *The Journal of Philosophical Logic*, 37(1), 67–94.

Sillari, G. 2008. Quantified Logic of Awareness and Impossible Possible Worlds. *Review of Symbolic Logic*, 1(4), 1–16.

Stockmeyer, Larry. 1987. Classifying the Computational Complexity of Problems. *The Journal of Symbolic Logic*, 52, 1–43.

Tennant, N. 1990. *Natural Logic*. Edinburgh: Edinburgh University Press.

Wittgenstein, L. 1961. *Tractatus Logico-Philosophicus* (1921). New York: Routledge and Kegan Paul. Translated by D.F. Pears and B.F. McGuinnes.