 DOI: http://dx.doi.org/10.5216/phi.v29i1.78657

Inferential limits of Machine's Intelligence: Can Kant teach us anything about the content of A.I. judgments?[1]

Lucas Vollet[2,3]
luvollet@gmail.com

Abstract: We argue that a framework for comprehending the basic differences between the mental structures of humans and machines (as they currently exist) is established by Transcendental Analytics' argument in the *Critique of Pure Reason*. It will be demonstrated that Kant's theory of the synthetic unity of apperception, as established by Transcendental Analytics' argument in the *Critique of Pure Reason*, along with Dummett's theory of meaning for meaning-theoretical predictions of inferential connections, can assist in establishing this framework. When combined, these form a framework for organizing a coherent differentiation between what we refer to as the conscious grasp of the unity that is present during judgment and the machine-performed manipulation of signs. In the end, we will present an appendix on the underdevelopment of the Kantian framework for distinguishing artificial intelligence from human intelligence.

Keywords: Inferentialism, artificial intelligence, synthetic unity of apperception, Dummett, Kant.

---

[1] Recebido: 02-03-2024/ Aceito: 07-06-2024/ Publicado on-line: 30-06-2024.
[2] É doutor em Filosofia pela Universidade Federal de Santa Catarina (UFSC), Florianópolis, Santa Catarina, Brasil.
[3] ORCID: https://orcid.org/ 0000-0002-6300-491X.

Programmatic Introduction

The goal of this paper is to offer a framework for structuring a meaningful difference between machine-performed manipulation of signs and conscious representation in relation to performing inferences. It will be demonstrated that Kant's theory of the synthetic unity of apperception, as established by Transcendental Analytics' argument in the *Critique of Pure Reason*[4], along with Dummett's theory of meaning for meaning-theoretical predictions of inferential connections, can assist in establishing this framework. To accomplish this, the strategy employed is to emphasize the inferential aspect of representation. The article commences by illustrating how contemporary inferentialism addresses the philosophical requirement for a notion of representation conditioned by its inferential function. We will get to see how this reflection takes shape as the harmony theory, which is derived from Gentzen's *sequent calculus* (1969) and draws on the symmetry criteria between the rules of introduction and elimination of symbols. Subsequently, the article will engage in a discussion regarding the various means through which this symmetry can be established, while demonstrating that 1. Dummett remains attached to a theory of meaning-theoretical harmony applied to language and linguistic learning, in a program similar to Davidson's (2001) program, but with greater depth on the intensional dimension. 2. that Kant goes beyond a mere theory of linguistic harmony conditions, establishing a theory on the harmonic enrichment of our conceptual repre-

[4] Abbreviation for *Critique of Pure Reason*: *KrV* Kritik der reinen Vernunft (1781, 1787). Cited by A/B pagination.

sentations through their unity with intuition.

In the most direct argumentative part of the article, we will ask whether computers' inferential capacity is as rich as ours. It will be seen that there is an indifference between the inferential competencies of a machine and our own, if we consider both from the perspective of their practical success. Because Dummett's harmony theory remains tied to the test of inferential success or failure, it is unable to detect an observational difference between a machine's successful inference and ours.

The subsequent step of the article is to argue that this indifference is gradually eliminated as we consider the possibilities of specific enrichment of inferential content, made by a different competence, such as that described by Kant's theory of *a priori* syntheses. We will then have arrived at a provisional philosophical framework for understanding the disparity between human consciousness and artificial intelligence machines. This differentiation is made on entirely Kantian bases: and can only be established on the bases of his theory of syntheses, his transcendental logical theory and his conception of judgment as a synthetic unity between concepts and intuitions. However, it should be noted that this differentiation is not flawless, as there is no certainty that the underlying components supporting it will not be invalidated in the future by new insights into machines and our consciousness. Following the article's conclusion, we will address the problems left unanswered by Kant in an appendix that will include a brief study of the anthropocentric foundations of his paradigm for differentiating between artificial intelligence and human intelligence.

Dummett's Inferentialism and Harmony of a meaning-theory

Modern inferentialism developed more recently in reaction to the representationalism of analytic philosophers in the first part of the 20th century, especially Bertrand Russell, who helped the setting of neo-empiricist boundaries in analytic philosophy: "our world is not wholly a matter of inference" (Russell, 1959, p. 23). Logical empiricism brought classical empiricism back to life, as Robert Brandom would say "by appealing to the new quantificational predicate logic Russell had developed out of Frege" (Brandom, 2007, p. 653). Sellars (1997) critique of the "myth of given," which is defined as the idea that certain beliefs are "given" to us through direct sensory experience and represent privileged lines of contact with the outside world, is one of these contributions that forms the core of inferentialism. Richard Rorty commended this critique, stating that it was crucial in transitioning analytic philosophy from its Humean phase to its Kantian phase (Rorty, 1997, p. 3).

We can add the premise that inferentialism adds a level of reasoning about the basis of our inferences that is lacking in pure representationalism, even when it is compatible with functionalist types of representationalism (Fodor, 1975). In this dimension of reasoning, the portrayal of the unity of the foundations authorizing a judgment is at issue. However, what sort of reasoning is this? According to Dummett, this is the reasoning behind our conception of a meaning theory. For the author: "we use the word meaning in such a way that any difference in meaning between two expressions involves a different in effect" (Dummett, 1991, p. 21).

Unlike Sellars, who focused on an epistemological agenda, Dummett developed his inferentialism as a reaction to language theories based on extensionalist and truth-centered conceptions. He created an inferentialist theory that represents the meaning-theoretical contribution to a more robust meaning theorization project than truth-functional extensionalism. For him, instead of reducing the concept of meaning to that of truth-functionality, "the relevant application of the concepts of truth and falsity must be governed by whatever it is that these notions are supposed to have with the meaning of the sentences" (Dummet, 1991, p. 22).

Dummet's objective, then, fit the then-current philosophical paradigm: the development of philosophical meaning theories concerning the establishment of meaning-theories to support language acquisition ability. Donald Davidson led this philosophical movement in the latter part of the 20th century and carried on with the truth-theoretical quest to develop a theory of language interpretation and learning: "There is a sense, then, in which a theory of truth accounts for the role each sentence plays in the language in so far as that role depends on the sentence's being a potential bearer of truth or falsity; and the account is given in terms of structure" (Davidson, 2001, p. 61).

However, Dummett created his theory in order to make up for the shortcomings of a purely truth-functional or extensional theory of meaning, because he could establish the conditions for learning 'p' based on a more robust theory that could identify the conceptual opposition between 'p' and 'not-p,' even for non-classical conditions: "the interesting cases are those in which the logic is non-classical, and

hence the explanation of at least some of its logical constants are not even relatively straightforward" (1991, p. 26). While Davidson stayed within the parameters of extensional detection of meaning, using the Tarskian scheme as a tool to identify learnable sentences, Dummett build a meaning-theory that can forecast sentence usages that are consistent with their inferential outcomes. This meaning-theory can predict language usages in situations where their inferential function cannot be separated from them. A meaning-theoretical theory of this kind relies on the speaker's knowledge of "harmony" that connects between truth-grounds and its implications.

The following circumstances apply when a harmonic meaning theory predicts how a language will be used. If the language is meaning-harmonic, then the difference between 'p' and 'non-p' in that language must be representable by a predicate 'R' that cannot be agreed upon and disputed upon under the same conditions. In order to identify a consistent predicate of the language that establishes the incompatibility between any statement and its negation: "we require a harmony which obtains only if a statement that has been indirectly established always could (in some sense of 'could') have been established directly" (Dummett, 1975, p. 227). Dummett's theory is consistent with an inferentialism compatible with intuitionism and anti-realism, since the premise of Dummett's theory of harmony is that 'p' and 'not-p' are not in a relationship of obvious and direct opposition, but rather in an indirect and inferential relationship, which is obtained under harmonious conditions such as those enunciated by Gerard Gentzen: "In eliminating a symbol, we may use the

formula with whose terminal symbol we are dealing only 'in the sense afforded it by the introduction of that symbol'" (Gentzen, 1969, p. 80).

Languages with harmonic theories of meaning prohibit compatible sentences from having contradictory extensions and incompatible sentences from having the same extension. There would be no different outcomes in these languages from accepting or agreeing with 'p.' This particular language would be highly advantageous for a foreign linguist to acquire, as it would eliminate any significant differences in the behavior of two individuals (who are being interviewed by the linguist) who share the belief in 'p', unless one or both of them lack comprehension of said language.

Because Dummett's theory of language could predict substantial inferential contents involved in extra-extensional connections, it outperforms Davidson's theory in terms of predictive value (and as we will see, it presents challenges to the supposition that machines can make inferences as rich as ours).

The indifference between consciousness and machines in formulating effective inferences

Dummett's theory was so robust that it allowed testing the consistency between the consequences of 'p' and the rational disposition to assert that 'p', that is, a theory that went far beyond the mere recognition of the extension of p. In Dummett's theory, whoever knows how to assert that p must also know the inferential role of p. A theory such as this has the benefit of tracking the flow of clarification or explanation of inferential commitments connected to sentences, as

Robert Brandom rightly pointed out: "insofar as the theory of semantic or inferential harmony makes sense at all, it must take the form of an investigation of the ongoing elucidative process, of the 'Socratic method' of discovering and repairing discordant concepts" (2001, p. 75).

Indeed, for Dummett, provided we have clarity about our meaning-theory we can establish a test parameter that effectively distinguishes between effective and defeatist strategies of assertion. If the test is decidable: "there must be an observable difference between the behavior or capacities of someone who is said to have that knowledge and someone who is said to lack it" (Dummett, 1975, p. 7).

However, different conceptual frameworks can align with the same inferential information, which can make it difficult to distinguish between individuals who are believed to have that inferential knowledge and those who are believed to not have it. Detecting this difference, in the case of machines, can become a tiring task. Turing tests will become increasingly difficult to become effective tests, as the imitative capabilities of machines become more complex and indistinguishable from our intelligent capacities.

Since the criteria and parameters used by machines, like those used by humans, are determined by directives to succeed, predictions about their results might take on a variety of meaning-theoretical forms without any special characteristic allowing them to be separated from one another: "a number of significantly different theories of truth will fit the evidence equally well" (Davidson, 2001, p. 62). An extensionalist like Davidson holds the following beliefs because, if the machine is successful in completing its work, all theories

predicting that accomplishment will be extensionally identical: "It would be no easier to interpret what l'homme machine means by what it 'says' than to interpret the words of a man, nor would the problem be essentially different" (Davidson, 1973, p. 258).

Success seems to be an anthropomorphic concept. But even if it may be strange to envision a machine that is structured by anthropomorphic aggressive criteria of victory, that kind of cost-cutting logic is precisely what is instilled in it through algorithmic guidelines for successfully completing tasks. It is still possible to argue that the machine doesn't dramatize its aspirations for success, doesn't get angry when things don't work out, and doesn't tell itself stories about the kind of successful machine it wants to be. But if it's the only thing that separates machines from humans, then the portion of this split that belongs to us is not attractive.

Transition to Kant's theory: the challenging part of a meaning-theory

Dummett's harmony theory does not pose any philosophical challenge for machine learning of language's inferential connections, since it just outlines the intra-linguistic requirements for producing the harmonic reconciliation of sentence outcomes and its verification procedure. Although his theory adds an extra-extensional layer of complexity to meaning theories, it is consistent with predicting the inferential behavior of computers with artificial intelligence. Dummett's theory dedicated his allegiance to the professional fixation of his time – that is, to provide a theory (stronger than Davidson's extensionalist semantics) – in order to establish

an effective theoretical prediction of meaning-theoretical capacity to speak a language.

But the challenging part of a meaning-theory is not intra-linguistic. Of course, it is the part where we have to define the inferential function of hypothetical, ungrounded sentences, or sentences whose grounds need to be indirectly anchored. We will call these underspecified sentences. As we shall see in the next section, synthetic *a priori* sentences[5] are typical instances of underspecified sentences because they have a vague intuitive connection, or can only be determined by higher-order synthesis. Before we go to that section of the essay, we call these sentences – which make up the upper layer of theoretical statements in any language – underspecified sentences because, in contrast to directly grounded sentences, they stray further from an empirical boundary.

However, there are several ways to obtain a theory for grounding those sentences. Since the early debates between rationalism and empiricism, there have been many different philosophical views about how to reconcile the inferences that might be made from an underspecified statement with the grounding-conditions for asserting the statement.

We shall examine how the question of meaning-theoretical harmony evolved in Kant's day by focusing on the capacity for mental synthesis rather than language-learning in the upcoming sections. We anticipate some of this theory's general characteristics before delving into its specifics. For Kant the question of meaning awareness needs to be posed in

[5] In Kant's words, a synthetic *a priori* judgment shares the synthetic characteristic of being based on an amplificative and, therefore, non-conceptual/analytical connection and, at the same time, the *a priori* characteristic of not being able to count on the help of sensible experience (*KrV* A 9/ B 13).

terms of the possible representation of the synthetic unity of the content of the judgment, which gives *objective reality* to pure concepts. Once we are able to generate this synthesis, we will have conceptual clarity on the success conditions for the hypothesis that p is true, and we will know precisely what the best tactical defensive stance is to assert 'p'. Even if 'p' is only a hypothesis or an underspecified sentence (even a synthetic *a priori* one), the knowledge of the ideal conditions for proving 'p' is what is referred to as the "objective reality".

We shall now demonstrate how this kind of inferentialism is consistent with Kant's synthetic theory for representing the "objective reality" of synthetic *a priori* judgments, according to his view of the synthetic unity of apperception.

Kant's theory of harmonizing intuitions and concepts in a transcendental Logic

In the Analytic of Principles, Kant makes the case that Transcendental Logic "offer a general but sufficient characterization of the conditions under which objects in harmony with those concepts can be given, for otherwise they would be without content" (*KrV* A 136 / B 175). The key to understanding this process lies in Kant's separation between general logic and transcendental logic: "although general logic can give no precepts to the power of judgment, things are quite different with transcendental logic" (*KrV* A 135/B 174). The author had already underscored how purely logical differentiation lack content: "the difference between an indistinct and a distinct representation is merely logical, and does not concern the content" (*KrV* A 44/B 61).

The emphasis on this distinction has been reiterated

numerous times as a fundamental factor in comprehending Kantian philosophy. To explore this thesis in a novel way, we will employ a different interpretation approach that specifically focuses on the inferentialist aspect, as advocated by Dummett's theory of harmony.

It is easy to develop linguistic learning principles when the conditions for meaning-theoretical harmony are met. However, this is much more challenging when we try to establish "objective reality" for sentences that have a high theoretical and predictive (i.e., non-analytic) content and do not directly involve a verification source (synthetic *a priori* judgments). A meaning-theoretical framework must be established in order to justify this mismatch for synthetic *a priori* judgments and this is much more challenging (Kant thought that this would require a *Deduction of Pure Categories*)[6]. This requires a different kind of harmony, intuitively richer than just logical harmony that comes from analytically connecting the sentence with its premises.

For Kant, the consciousness of logical differences has a content as well as a form. But for that to happen, some sort of logical characterization of the synthetic unity of this difference is needed. In section nineteen of the Transcendental Deduction (*KrV*, B 141) the author tries to illustrate the development of this thesis by suggesting an example of what is thought when we represent the proposition "The body is

[6] In this article, we take the stance that Transcendental Deduction is a component of a larger theory concerning human cognition and its productive capacities to generate the content of synthetically problematic sentences (which may be at odds with its intuitions). On the other hand, the transcendental aspect of his argument – which came to be known as the model for transcendental arguments – represents a separate and distinct aspect of his contribution.

heavy" as a combination of weight and body. The author says:

> I do not mean that these representations necessarily belong to one another in the empirical intuition, but rather that they belong to one another in virtue of the necessary unity of the apperception in the synthesis of intuition (KrV B141/143).

From now on we will represent "the body is heavy"[7] as just 'p'. The simple generic-logical form of 'p' does not confer any cognitive or intuitive distinction between the hypothesis that 'p' is true and the hypothesis that 'not-p' is false. The difference between 'p' and 'not-p' can be formally represented as a contradiction by general logic, but general logic is unable to capture the cognitive content involved in various patterns of computation of 'p,' such as the various mental rules used in the cognition of the beliefs that 'p' is true and 'not-p' is false. To repeat succinctly: While *general logic* can represent the difference between 'p' and 'not-p' formally as an extension and its anti-extension, it fails to capture the cognitive content involved in different patterns of computation for 'p'. This means that the cognitive paths to prove that 'not-p' is false may be richer than that to prove that 'p' is true, with the consequence that the "content" of this difference goes beyond the criterion of non-contradiction. The synthetic unity considered in this context implies a richer thought – although no thought can be as rich as a complete intuition – than the one perceived solely through the generic and syntactic form of the connection.

[7] We are using the two inverted commas to represent propositions.

Kant's theory of Truth and Logical Consequence

In KrV, Kant defines truth as an agreement of the understanding and object (*KrV* A 58 / B83). Nonetheless, this description is limited. A more accurate explanation may be found in *Logik*[8]: "Truth is, above all, the primary perfection, since it is the ground of the unity between our knowledge and the object" (AK 39). In contrast to the initial description's emphasis on agreement and correspondence, the second description highlights the significance of the ground of unity as a pivotal element for Kant: the capacity to assess or identify the judging grounds of a representation (thereby introducing an inferentialist component):

> All judgments are accordingly functions of unity among our representations, since instead of an immediate representation a higher one, which comprehends this and other representations under itself, is used for the cognition of the object, and many possible cognitions are thereby drawn together into one (*KrV* A 69/B 94).

Similar to how Boole's extensional logic was rejected by Frege, Kant had already devalued the definition of logicians for the formal relation of judgments: "it is not here determined wherein this relation consists" (B 141). In *KrV* 267/B 323, the author mentions: "In every judgment one can call the given concepts the logical matter (for judgment), their relation (by means of the copula) the form of the judgment".

Here, a relationship between concepts through copula expresses an assertoric categorization. Kant's quandary with the "logicians" is that, however, when there is a relationship

[8] From now on cited by pagination AK.

between two judgments (like disjunctive judgments or Filo's conditional), this extension needs to be resolved in relation to something else, such as the tendency or inclination to affirm 'q' in the condition that 'p' is affirmed. But inclination or tendency are psychological concepts, determined subjectively either by imagination or by the matter of contingent cognition. In both those cases, the grounds for inference would remain problematic, which is why the logician's characterization do not fit "hypothetical and disjunctive judgments"[9] (*KrV* B 141).

What the author devalues is the logician's haste to determine the subject of judgments through a "form" without first explaining what this form imparts to the cognition that comprehends it. By categorizing the form of the relationship between two judgments in the conditional 'if p then q', we cannot claim to know what the relationship between 'p' and 'q' consists of, as we must first ascertain the link between 'p' and 'q' as a non-psychological unit: "that is, the aim of the

[9] This lengthy footnote is required to make clear that Kant was unable to discuss the material implication – that is, the truth-functional representation of the conditional – because he was obviously unable to access the further historical advancement of logic. It is important to note that, in this aspect, if our theory about Kant's being an inferentialist is correct, then his reply would be that the material implication is merely a hypothetical extension, something that could be employed in a judgment, but it isn't a judgment just yet. There are a few noteworthy works among the extensive body of work on the material implication conditional that we can quote to make this same point. According to Gilbert Ryle, the conditional's truth-functional representation is simply an inferential ticket: "Knowing 'if p, then q' is, then, rather like being in possession of a railway ticket" (Ryle, 2009, p. 250). Etchemendy claims in a critique of representationalist semantics "the fact that the target sentence would have been false in a row of the table was taken to indicate that the sentence would been false in a row (...). But the third row itself (...) is just a handy surrogate, for the aims of our theory" (Etchementy, 1999, p. 20). Etchemendy concludes a few pages later: "Although these models would given us complete partition of possible worlds, the partition would not have been fine-grained enough" (1999, p. 24). This digression demonstrates how similar the potential responses to the truthfunctional theory of hypothetical extensions can be. All of them assert that while the hypothetical extension establishes a particular relationship between two judgments, it only provides a problematic extension.

copula (...) to distinguish the objective unity of given cognitions from the subjective" (*KrV* B 142). The affirmative favorability of q under the condition p does not signify a singular extension, but rather a plethora (a manifold) of potential extensions in relation to other possible extensions. According to Kant, there is not a transition from a problematic to a non-problematic dimension in this case. According to the author's assessment of mathematics' instrumental reach, the best we can do is create an abstract depiction of the relationship between two "hypotheticals": "Mathematics (...) is an excellent organon containing the ground of the expansion of our cognition in regard to a certain use of reason" (AK13, p. 15). But, for Kant "Logic is not (...) an algebra that helps to discover hidden truths" (AK 20). The author requires here the theoretical part involved in knowing this unity of a manifold. If we are not to be blind to the content of the inferences, the unity of this manifold needs to be demonstrated by a higher order concept (an apperceptive synthesis).

While these passages may not seem noteworthy to Kant, they are among the few that reveal the author's primary motivation behind several of his other theses. Because it's evident from these excerpts that the author, like Dummett, values inference over representation[10], even when it comes to

[10] Patricia Kitcher in *Kant's Transcendental Psychology* argues that Kant's answer to Hume's empiricism stems from his representationalist functionalist interpretation of the concepts of judgment and apperception. The theory would provide richer and more complicated representational operations – what Kant refers to as "syntheses" – in order to get around some of the drawbacks of Humean associationism: "A synthesis is an act, or to be more neutral, a process that produces a representation, by assign or combining diverse elements contained in different cognitive states in a further state" (Kitcher, 1990, p. 74). However, the historical progression of this matter indicates that the reconciliation process, which often occurs between conflicting philosophical propositions, also takes place in this particular case. The fact that some aspects of this reconciliation are already

choosing how to depict the synthetic unity that connects the truth-grounds to their outcomes in a theoretical understanding of the grounds for judgment. In describing the intensional magnitude of a concept that refers "to the content (*Gehalt*) (...), insofar as it is considered the foundation of many great consequences," Kant also demonstrates the inferentialist component of his theory (AK 40).

Machines under-specifications of content: Can robots think as richly as humans?

Kant would concede that computers and even people incapable of making judgments may successfully produce algorithmic instructions to learn to agree and disagree with 'p' under the same favorable and unfavorable conditions as those experienced by people with the faculty of judgment: "A dull or limited head, which is lacking nothing but the appropriate degree of understanding and its proper concepts, may well be trained through instruction, even to the point of becoming learned" (*KrV* A 134/B 173).

But we must acknowledge that the synthetic unity of the content goes beyond associative patterns, namely, those "whose synthesis is subject solely to empirical laws of association" (*KrV* B152). This variation in richness is due to an intuition that is formed spontaneously by the synthesis of the content. If Kant and Dummett are right, human inferential intelligence would, apparently, have richer patterns for

present in Kant's thesis indicates that it is one of the most comprehensive and intricate theories on the development of cognitive content. One method to foresee the necessity of combining representational and inferential elements in an account of cognition is to adopt the Kantian assumption that concepts are nothing without intuition, and vice versa (*KrV* A 51/B 75).

thinking and computing than mere associations between signals. This wealth of specification is reflected in the greater determination of the role of this thought as a premise or conclusion of arguments.

It is now clear to us that this content cannot be adequately defined by the straightforward syntax of 'p'. To do this, the meaning-theoretic prerequisites for a logical distinction between p and non-p must be stated clearly, and they must be objectively framed as incompatible rather than merely associatively framed as opposed. We can state that underspecifications in the content of 'p' can be eliminated by human thought, which allows for a gradual improvement in the scope of 'p' in a proof context – as 'p' is meant to work as premise and conclusion of inferences.

Dummett developed his theory of harmony, as we have seen in previous sections, to offer meaning-theoretical predictions of fine-grained inferential outcomes from inputs, i.e., balancing truth-grounds with their implications. We then observed that Kant's theory of *a priori* syntheses was created under the pressure of a similar necessity: the pressure for bringing theoretical statements – which are not fully specified – in line with their verificational base. Dummett's theory is an echo of an earlier discussion, which existed in Kant's time as a theory of *a priori* syntheses, despite his adherence to the professional obsession of his time, which was to provide a theory (more powerful than Davidson's) to establish a successful theoretical detection of meaning-theoretical competence to speak a language. In the Kantian form, a theory of harmony provides a theory of mind structure for the conceptual representations required for 'p' to be

specified in order to progressively define a particular scope for 'p' in a proving-inferential context. Of course, that theory present ideal cases. We won't delve into the controversy around whether or not those theories adequately represent the true mental processes of humans or whether they require modification to fit the constraints and requirements of methodological paradigms of modern science (like Kant's categories and its alignment with Newtonian physics), etc. We don't need as much here. All that theory has to do is offer a theory of mental structure explaining the ideal conditions under which the definition of p would be sufficiently stable to justify its distance from not-p. This is enough due to the purpose of our article, which has as its sole purpose a comparison with artificial intelligence's inferential capabilities (in its current condition).

Can machine inferences work with such specifications? is the question. Can machines get rid of 'p''s logical underspecification? To put the question in a more Kantian format: can machines give an objective reality to synthetic sentences that are highly theoretical and therefore intuitively underspecified? It seems that there's no fundamental reason why machines can't accomplish this. Nothing would stop a machine from operating with the same computational richness as ourselves, even if Kant is correct and this operation requires the representation of a higher order unity, which is what defines human awareness of the unity of representations. However, we can know – and we can know it now – that, if Kant is correct, there are still some tasks that are not completed by machines. One such task is the transcendental (non-general) basis of the notion of logical incompatibility,

which characterizes our awareness of the unity of representations. Computer-machines possess the capacity to augment their computational abilities through programming; nevertheless, they lack the capability to spontaneously generate this enhancement through synthetic enrichment.

Conclusion

The question of whether machine-computers' inferential capacity is as rich as ours became the focal point of the debate in this article. We delved into Dummett's definition of harmony, which serves as the theoretical and predictive basis for reconciling the utilization of sentences and their inferential outcomes. An individual's capacity for inference is represented by this meaning-theoretic base. We saw that this competence can be put to the test. But we also noticed that, in the case of underspecified sentences, multiple conceptual frames can be consistent with the same inferential data, making it impossible to establish an observable difference between the capacities or behavior of someone who is said to possess that inferential knowledge and someone who is said to lack it. We have seen that if "success" becomes the only criterion of distinction, we will have no recourse for discerning artificial intelligences and organic intelligence, other than the inadvisable fact that we represent our success or failure in dramatic ways – and the machine does not. Consequently, the more sophisticated the means of machine imitation become, the more challenging it will be to detect an observationally based distinction between humans and machines.

As we have not given up looking for elements to detect this difference, we have prepared the transition to an exposition of Kant's theory. When it came to the question of meaning-theoretic harmony inside a language, Dummett refrained from taking a radical stance and avoided discussing the processes of generating synthetic foundations for underspecifying sentences (synthetic *a priori)*. We observed how Kant formulates this question with a greater degree of radicalism: as the question of richer ways to achieve a harmony in order to establish a link between theoretical concepts and their verificational basis, i.e., a representation of the synthetic unity of intuitive and conceptual representations.

The question for cognition of *underspecified* sentences – like *a priori* synthetic judgments – is whether it is possible to construct a framework (the synthesis) that satisfies both the theoretical and empirical requirements for confirming the true of 'p'. Individuals who can establish this alignment between theoretical content (concepts) and verification methods (intuitions) possess the epistemic ability to gradually eliminate the underspecification of the *a priori* synthetic judgment, giving it a "objective reality", or bringing some "pure intuition" to the fore by the synthetic procedure. Finally, we questioned whether it is possible for machines to make inferences with this level of specifying richness, and we came to the conclusion that nothing would prevent them. However, we also observed that present computers cannot achieve the transcendental limit of this capacity because doing so would require the spontaneous generation of the synthetic content of these inferential connections. Put otherwise, computer machines have the ability to be programmed to enhance their

computational foundation, but they are unable to generate this enhancement synthetically and spontaneously. Thus, we have arrived at a coherent, if provisional, philosophical framework to account for the differences between machine-assisted signal processing and human inferential intelligence.

Appendix on the underdevelopment of the Kantian framework for distinguishing artificial intelligence from human intelligence

The final paragraph of our article brings it to a close. After this, the final section will read more like an appendix in terms of style and content, with more questions and angles for future research. The main point of this article's observations is to frame the difference between existing machine processes and human intellect. However, as with any broad and preliminary differentiations, numerous questions remain unanswered. Kant's elucidation of the distinction remains tentative, as it is yet to be definitively determined if there is potential for its advancement. While Kant does not neglect these foundational ideas in his writings, the progression he offers is underdeveloped. Kant's view of judgment was promoted in opposition to passive and unintelligent forms of cognitive representation. The author's theory avoided attributing intelligence or cognitive content to simple operations with signs (formulas), which is how his theory could present a straightforward lesson of opposition to the thesis that machines (in their current state) can think. However, this instructional unit lacks depth, given that the writer lacks the necessary tools to provide detailed explanations without relying on anthropological and anthropocentric reasoning to

define intelligence, inference, and judgment.

We will occupy this end of the article with some of these uncertainties, which need to be further explored in the future.

1. This thesis raises uncertainties regarding the convergence and conflation of intelligent and mechanical states in practical endeavors, as both can equally succeed in completing tasks.
2. It is also unclear whether intelligence emerges from non-intelligence or if it is inherently present, even during phases when it relies on mechanical processes for guidance.
3. The precise trigger or catalyst required for a machine to transition from merely following instructions to becoming self-aware of the content of those instructions remains unknown to us.
4. Additionally, a challenging issue arises concerning how phenomenological introspection – of the type required by Kant to determine the harmonic unity of judgment – can avoid descending into a circular anthropocentric representation of whatever an intelligence is supposed to be.

These uncertainties raise a new question, about the possibility of machines replacing us in intelligent activities, even if they are not intelligent. Assuming machines can solve all our problems using our success criteria, what additional factors would be necessary to deem them intelligent? Furthermore, Kant skips over the discussion of how a human being lacking the faculty of judgment is represented in the process of gaining it, as people often use dramatization and political

narrative to express their failures as stories of maturation and progress. He also skips over how this differs from the failure of an automaton or machine to complete a task.

In the midst of so many unsolved questions, the entire corpus of Kant's writings provides a hint as to where attentive readers can go for answers. Kant addresses these issues drawing a connection between the capacity for judgment and the capacity to set normative standards. This relationship, consistent with freedom and practical reason, directs Kant toward a morally and practically regulative application of absorbing principles that completes the theoretical portion of understanding. However, our piece is unable to carry on the path of reflection after this.

Resumo: Argumentamos que um arcabouço para compreender as diferenças básicas entre as estruturas mentais dos humanos e das máquinas (tal como existem atualmente) é estabelecida pelo argumento da Analítica Transcendental na *Crítica da Razão Pura*. Será demonstrado que a teoria da unidade sintética da apercepção de Kant, conforme estabelecida pelo argumento da Analítica Transcendental, juntamente com a teoria do significado de Dummett para previsões "meaning-theoretical" de conexões inferenciais, pode ajudar no estabelecimento dessa estrutura. Quando combinados, esses dois elementos formam um arcabouço teórico para organizar uma diferenciação coerente entre o que chamamos de compreensão consciente da unidade que está presente durante o julgamento e a manipulação de sinais realizada por máquinas. Ao final, apresentaremos um apêndice sobre o subdesenvolvimento do arcabouço de teses kantianas para distinguir a inteligência artificial da inteligência humana.

Palavras-chave: Inferencialismo, inteligência artificial, unidade sintética da apercepção, Dummett, Kant.

References

BRANDOM, R. *Articulating Reasons: An Introduction to Inferentialism.* Harvard University Press, 2001.

BRANDOM, R. Inferentialism and Some of Its Challenges. *Philosophy and Phenomenological Research*, v. 74, n. 3, p. 651-676, 2007.

DAVIDSON, D. Psychology as Philosophy. In: BROWN, S. (ed.). *Philosophy of Psychology*. Londres: Macmillan, 1973. p. 41-52, 60-67, 1973. Reprinted in Davidson 1980a, p. 229-44.

DAVIDSON, D. *Inquiries into Truth and Interpretation*. Clarendon Press. Oxford, 2001.

DUMMETT, M. Truth. *Proceedings of the Aristotelian Society*, v. 59, n. 1, p. 141-62, 1958.

DUMMETT, M. Frege: *Philosophy of Language. London***:** Duckworth, 1973.

DUMMETT, M. *The Philosophical Basis of Intuitionistic Logic*. *Truth and Other Enigmas.* Cambridge: Harvard UP, 1975. p. 215-247.

DUMMETT, M. What is a theory of meaning? In: GUTTENPLAN, S. (ed.). *Mind and Language.* Oxford University Press, 1975.

DUMMETT, M. *The Interpretation of Frege's Philosophy.* Gerald Duckworth & Co Ltd, 1983.

DUMMETT, M. *The Logical Basis of Metaphysics.* Harvard

University Press, 1991.

ETCHEMENDY, J. *The Concept of Logical Consequence*. Cambridge Massachusetts: Harvard University Press, 1999.

FODOR, J. A. *The Language of Thought*. Harvard University Press, 1975.

GENTZEN, G. translation in: SZABO, M. E. (ed.). *The Collected Papers of Gerhard Gentzen*. Amsterdam: North Holland, 1969. p. 68-131.

KANT, I. *Logic*. New York: Dover Publications, 1988.

KANT, I. *Critique of Pure Reason*. Cambridge University Press, 1998.

KITCHER, P. *Kant's Transcendental Psychology*. Oup Usa, 1990.

RORTY, R. Introduction to Wilfrid Sellars, *Empiricism and the Philosophy of Mind*. Cambridge: Harvard University Press, 1997.

RUSSELL, B. *My Philosophical Development*. London: George Allen and Unwin; New York: Simon and Schuster, 1959.

RYLE, G. *Collected Essays*. 1929-1968. Volume 2. London and New York: Routledge, 2009.

SELLARS, W. *Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press. Vinci, T, 1997.