CAN ARTIFICIAL INTELLIGENCE POSSESS FIRST PRINCIPLES? AQUINAS' *HABITUS PRINCIPIORUM* AND THE NATURE OF HUMAN RATIONALITY

By

Mary Christine Ugobi-Onyemere, IHM (PhD)
Department of Philosophy and Religious Studies
Dominican University, Ibadan
gigamugobi@gmail.com; ugobi-o@dui.edu.ng

Abstract

Recent advances in artificial intelligence (AI) have intensified philosophical debates about intelligence, reasoning, and machine cognition. While discussions often compare human and artificial intelligence in terms of computational efficiency, learning, and problem-solving, they seldom address a more fundamental question: can artificial systems possess the first principles that make rational knowledge possible? This paper revisits Thomas Aquinas' doctrine of *habitus principiorum* as a framework for addressing this question. Aquinas argues that demonstrative reasoning presupposes an intrinsic intellectual habit through which the human intellect immediately apprehends self-evident first principles, including the principle of non-contradiction and other foundational truths. These principles are not acquired through inference but constitute the precondition of subsequent knowledge and rational judgement. The paper argues that *habitus principiorum* is neither an innate stock of propositions nor merely a psychological disposition, but an ontological and epistemological perfection of the intellect that grounds objective rationality. From this perspective, contemporary AI systems, despite their capacities for pattern recognition, optimisation, and probabilistic inference, operate without the intrinsic intellectual disposition required for genuine understanding

of universal first principles; their outputs are therefore algorithmically generated rather than intellectually apprehended. By bringing Aquinas into dialogue with contemporary philosophy of mind, epistemology, and AI, the paper argues that the distinction between computation and intellect remains philosophically significant. Aquinas' account challenges reductionist conceptions of intelligence by locating rationality in the metaphysical structure of the human person rather than computational processes alone, thereby offering a neglected resource for reassessing AI's limits and clarifying the distinctive character of human rationality.

Keywords: *Habitus principiorum*; First Principles; Artificial Intelligence; Human Rationality; Thomistic Epistemology

Introduction

The rapid development of artificial intelligence (AI) has renewed a longstanding philosophical question: what does it mean to be intelligent? Contemporary AI systems can perform tasks once regarded as distinctive of human intelligence, including complex reasoning, language generation, pattern recognition, learning, planning, and problem-solving. Indeed, the history of AI has been closely associated with attempts to formalise reasoning and construct systems capable of rational behaviour.[1] Yet the increasing sophistication of artificial systems also exposes a deeper philosophical problem. The ability to produce rationally structured outputs does not by itself settle whether an artificial system possesses the kind of rationality involved in understanding, judging, and knowing. The question, therefore, is not simply whether machines

[1] Margaret A. Boden, *Artificial Intelligence: A Very Short Introduction* (Oxford: Oxford University Press, 2018), 1–12; Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. (Hoboken, NJ: Pearson, 2021), 19–23.

can reason, but "**what makes reasoning genuinely intellectual in the first place?"**

This question becomes particularly significant when considered in relation to first principles. Every demonstrative argument presupposes principles that are not themselves established by prior demonstration. In Aristotelian-Thomistic epistemology, these principles are *per se nota*: self-evident propositions whose truth becomes manifest through an understanding of their terms. Aquinas identifies the principle of non-contradiction as the first indemonstrable principle and maintains that the intellect is not deceived concerning first principles.[2] Such principles are consequently not merely propositions stored in a knowledge system; they constitute foundational conditions for rational demonstration and scientific knowledge.[3]

Aquinas' account raises an important question for contemporary philosophy of AI. If an artificial system can manipulate logical principles, identify contradictions, formulate valid inferences, and generate apparently rational explanations, does it thereby possess first principles? Or does possession of first principles require a mode of intellectual apprehension that cannot be reduced to the computational manipulation of representations? These questions become especially pressing because contemporary AI incorporates both formal and probabilistic approaches to reasoning, while machine-learning systems can exhibit sophisticated forms of

[2] Thomas Aquinas, *Summa Theologiae*, *(ST)* I, q. 79, a. 9; I–II, q. 94, a. 2, trans. Fathers of the English Dominican Province (New York: Benziger Brothers, 1947).

[3] See, for example, Selmer Bringsjord & Naveen Sundar Govindarajulu, "Artificial Intelligence," *Stanford Encyclopedia of Philosophy*, 2018, especially the discussions of reasoning-based and behaviour-based conceptions of AI and the philosophical problem of machine intelligence. https://plato.stanford.edu/entries/artificial-intelligence/. Accessed, 24 June 2024.

behaviour without necessarily providing an account of how their operations relate to understanding or intellectual knowledge.

Aquinas' doctrine of *habitus principiorum* offers a distinctive framework for examining this problem. For Aquinas, the first principles of speculative reasoning are apprehended through *intellectus principiorum*, an intellectual habit by which the human intellect grasps indemonstrable principles immediately. This habit should not be understood as a storehouse of innate propositions. Rather, *habitus* denotes a stable perfection or disposition of an intellectual power ordered towards its proper acts.[4] Aquinas accordingly treats the habit of first principles as intrinsic to the perfection of human intellectual activity. In the practical order, he associates the corresponding foundational disposition with *synderesis*, through which the first principles of practical reason are known.[5]

The importance of this distinction lies in the difference between "**having access to a principle and intellectually apprehending it as a principle**." An artificial system may be programmed with logical rules, trained on vast bodies of textual material, or statistically generate outputs consistent with principles of reasoning. None of these capacities, however, by themselves establishes that the system possesses the intellectual habit through which first principles are understood as the foundations of rational judgement. The distinction invites a more fundamental examination of the relation between computation, representation, understanding, and intellect.

This paper therefore asks: "**Can artificial intelligence possess first principles in the Thomistic sense?"** It argues that Aquinas' *habitus*

[4]Aquinas, *ST*, I–II, qq. 49–51. Aquinas treats habit as a quality and as an intrinsic principle of human acts; see *ST*, I–II, q. 49, a. 1.

[5]Aquinas, *ST*, I–II, q. 94, a. 2. Aquinas explicitly compares the precepts of natural law in practical reason with the first principles of demonstration in speculative reason.

principiorum provides a philosophically significant criterion for distinguishing sophisticated computational performance from intellectual apprehension. The argument does not deny the genuine achievements of contemporary AI; rather, it questions whether functional competence in reasoning is sufficient for the possession of first principles. By bringing Aquinas into dialogue with contemporary philosophy of mind, epistemology, and philosophy of AI, the paper contends that rationality cannot be adequately characterised solely in terms of information processing or successful problem-solving. In the Thomistic account, first principles belong to an intellect capable of apprehending truth, and *habitus principiorum* names a constitutive perfection of that intellect.

The paper proceeds by first clarifying Aquinas' conception of *habitus* and the nature of first principles, before examining *intellectus principiorum* and its relation to the human intellect. It then considers the corresponding role of *synderesis* in practical reason and brings these distinctions into dialogue with contemporary conceptions of artificial intelligence. The final section assesses whether current AI systems can properly be said to possess first principles or whether their remarkable capacities remain distinguishable from the intellectual apprehension that, for Aquinas, grounds rational knowledge itself.

Aquinas on *Habitus Principiorum*

Aquinas' account of first principles begins from a fundamental feature of human cognition: some truths are known without demonstration and nevertheless serve as the foundations of demonstrative knowledge. Such principles cannot themselves be established through a prior chain of reasoning without generating an infinite regress. Aquinas therefore locates their knowledge in a natural intellectual habit. He distinguishes this habit according to the

two principal orders of rational activity: speculative and practical. In speculative cognition, it is called *intellectus principiorum*; in practical cognition, its corresponding natural habit is *synderesis*.[6]

Aquinas' most explicit account occurs in *De Veritate*, where he argues that human beings, insofar as their nature reaches towards the angelic mode of cognition, possess knowledge of truth without discursive inquiry in both speculative and practical matters. Such knowledge must precede subsequent cognition because principles are more certain and stable than the conclusions derived from them. It must also be natural and habitual: natural because it is not acquired through prior demonstration, and habitual because it must be readily available for use when required.[7] Aquinas consequently describes the speculative habit as that by which the soul knows the principles of the sciences and identifies it as *intellectus principiorum*. In the practical order, he identifies a corresponding natural habit of the first principles of action, pertaining to synderesis.[8]

The significance of *habitus* here requires precision. Aquinas does not posit a collection of innate propositions deposited in the human mind prior to experience; nor does he reduce the habit to a temporary psychological state. A habit is a stable quality by which a power is disposed towards its proper operation. In the case of first principles, therefore, the relevant habit perfects the intellect for the immediate apprehension of truths that function as principles of subsequent reasoning. The human intellect does not first acquire a demonstrated knowledge of these principles and then employ them; rather, their

[6] Thomas Aquinas, *De Veritate*, q. 16, a. 1. Aquinas explicitly distinguishes the natural habit ·of speculative first principles, which he calls *intellectus principiorum*, from the natural habit of practical first principles, which pertains to *synderesis*.

[7] Aquinas, *De Veritate*, q. 16, a. 1.

[8] Aquinas, *De Veritate*, q. 16, a. 1.

immediate apprehension makes demonstration itself possible.

This point also clarifies the relation between *habitus principiorum* and certainty. First principles cannot derive their certainty from conclusions, because conclusions depend upon principles for their demonstrative justification. Aquinas accordingly maintains that the knowledge of principles is more certain than demonstrative knowledge insofar as the former does not depend upon demonstration for its certitude.[9] The habit of first principles thus occupies a foundational position within the architecture of human knowing: it does not constitute all knowledge, but it provides the intellectual starting points from which demonstrative knowledge proceeds.

The natural character of this habit is equally significant. Aquinas explicitly states that the knowledge of first principles must be naturally present because it functions as a kind of seed of subsequent knowledge.[10] This metaphor should not be interpreted as though complete scientific knowledge were already contained in the intellect in an explicit form. Rather, the habit provides the principles and intellectual readiness from which further cognition can develop. Its natural possession therefore concerns the capacity to apprehend foundational truths, not the possession of the conclusions subsequently derived from them.

Aquinas further links *intellectus principiorum* to the immediacy characteristic of first-principle cognition. In his account, self-evident principles are known as soon as their terms are understood; their truth is therefore not reached by a process of demonstration.[11]This immediacy distinguishes *intellectus* from *ratio*. Reason proceeds

[9] Thomas Aquinas, *Commentary on the Posterior Analytics*, bk. 1, lect. 7; bk. 2, lect. 20.

[10] Aquinas, *De Veritate*, q. 16, a. 1. The passage explicitly describes this natural cognition as quasi *seminarium quoddam totius cognitionis sequentis*.

[11] Aquinas, *De Veritate*, q. 1, a. 12; *ST*, I, q. 17, a. 3, ad 2.

discursively from one cognition to another, whereas intellect apprehends what is immediately intelligible. The habit of first principles consequently belongs to the foundational activity of intellect upon which discursive reasoning depends.

This structure becomes especially important for understanding the principle of non-contradiction. Aquinas treats it as the first and most certain of the indemonstrable principles, since an attempt to demonstrate it would necessarily presuppose the very principle at issue.[12] First principles therefore function not simply as propositions within a logical system but as conditions of rational demonstration itself. Their epistemic priority derives precisely from their indemonstrability and immediate intelligibility.

The speculative and practical dimensions of this account are structurally parallel without being identical. *Intellectus principiorum* concerns the first principles underlying speculative knowledge, whereas *synderesis* concerns the universal first principles of practical action. Aquinas' comparison is therefore not accidental: both represent natural habits through which the human rational power is oriented towards foundational truth without prior discursive inquiry[13]. This parallel will become important when assessing artificial intelligence. A system may manipulate propositions expressing logical or moral principles, but such manipulation does not by itself establish possession of the intellectual habits Aquinas describes.

The Thomistic account therefore gives *habitus principiorum* a precise role in the constitution of human rationality. It explains how intellectual cognition can begin without beginning from an already demonstrated principle, while avoiding the need to posit an innate store of fully formed knowledge. First principles are naturally

[12] Aquinas, *ST*, I, q. 2, a. 1, ad 3; Thomas Aquinas, *Commentary on the* Metaphysics, bk. 4, lect. 6.
[13] Aquinas, *De Veritate*, q. 16, a. 1.

available to the intellect as foundational objects of immediate understanding, and the corresponding habit disposes the intellect to apprehend and employ them. The question for contemporary philosophy of artificial intelligence is consequently not merely whether a machine can reproduce conclusions based on first principles, but whether it can possess the kind of intellectual disposition by which such principles are apprehended as principles. That question requires a closer examination of *intellectus principiorum* itself.

The Foundational Totality of *Intellectus Principiorum*

The foundational significance of *intellectus principiorum* becomes clearer in relation to scientific knowledge. It is not itself a demonstrative science but the intellectual habit by which first principles are grasped. Aquinas therefore holds that *intellectus principiorum* is more noble than the science of conclusions: *intellectus principiorum est nobilius principium quam scientia* conclusionum.[14] Scientific reasoning proceeds from principles to conclusions; *intellectus principiorum* provides the immediate apprehension on which that reasoning depends.

Bogliolo's account highlights both the 'mode' and 'content' of this apprehension. Its mode is immediate rather than demonstrative, although the concepts involved ultimately arise through sensible experience. Its content begins with *ens* and the most universal intelligible principles, thereby opening the intellect to reality as intelligible.[15] This does not imply exhaustive knowledge but an openness to being that makes progressively differentiated knowledge possible.

[14] Aquinas, *ST*, I–II, q. 51, a. 2, ad 3.

[15] Luigi Bogliolo, *Essere e conoscere* (Città del Vaticano: Libreria Editrice Vaticana, 1983), 36–37, 49–56; see also C. M. J. Vansteenkiste, "Il metodo di san Tommaso," in Le ragioni del tomismo: dopo il centenario dell'enciclica 'Aeterni Patris', ed. A. Livi, C. Fabro, F. Ocariz, and C. M. J. Vansteenkiste (Milan: Ares, 1979), 163–166.

Aquinas' account of demonstration confirms this foundational role. Demonstration cannot establish its own ultimate principles without circularity or infinite regress; immediate principles are therefore presupposed by scientific knowledge. The cognition of these principles is, moreover, more certain than demonstrative cognition because they ground the certainty of conclusions derived from them[16]. The sciences consequently depend upon intellectus principiorum without containing it as one of their conclusions.

The "totality" of *intellectus principiorum* should therefore be understood formally and foundationally, not quantitatively. It does not mean possession of all possible knowledge but the intellect's universal orientation towards being and intelligibility. Vansteenkiste rightly emphasises the methodological importance of beginning with the 'given' and developing knowledge through the ordering of concepts, judgements, and reasoning.[17] The human intellect is thus open to an indefinitely extensible field of knowledge without being omniscient.

This distinction is crucial for the question of artificial intelligence. If first principles are merely propositions that can be stored or manipulated, an artificial system may appear to possess them. If, however, *intellectus principiorum* is an intellectual habit through which principles are immediately apprehended as principles, the question becomes more demanding: can an artificial system possess the intellectual actuality by which a first principle is understood as necessarily true? The distinction between representing or employing a principle and intellectually apprehending it will be central to the argument that follows.

[16] Thomas Aquinas, *In I Posteriorum Analyticorum*, lect. 7; In II Posteriorum *Analyticorum*, lect. 20.

[17] Vansteenkiste, "Il metodo di san Tommaso," 166.

The Principle of Non-Contradiction as an Expression of *Intellectus Principiorum*

The principle of non-contradiction (PNC) provides the clearest test of *intellectus principiorum*. For Aquinas, its intelligibility is grounded in *ens*, which is the first object apprehended by the intellect. As Giacon observes, the PNC expresses the non-contradictory character of reality as grasped by intelligence.[18]

Aquinas explains that the intellect has two fundamental operations: apprehension and judgement. In apprehension, *ens* is first conceived; in judgement, the PNC is first because it depends upon the intelligibility of being and non-being.[19] The order is therefore significant: **being is first in conception; the PNC is first in judgement**. The PNC is consequently not an arbitrary rule imposed upon thought but expresses the intelligible structure presupposed by rational judgement.

Aquinas formulates the principle as the impossibility of being and not-being simultaneously and, in the *Summa theologiae*, as the impossibility of affirming and denying simultaneously.[20] Its non-demonstrability does not make it irrational. Rather, it is known naturally when its terms are understood; it is a principle from which reasoning proceeds, not a conclusion reached through reasoning.[21]

The PNC is consequently both **ontological and epistemological**. Ontologically, it expresses the incompatibility of being and non-being under the same respect and at the same time. Epistemologically, it provides the basic intelligibility required for affirmation, negation, and demonstration. Hence Aquinas treats it as the firmest of principles while denying that it can itself be demonstrated.[22]

---

[18] C. Giacon, *Verità, esistenza, causa* (Bologna: Zanichelli, 1973), 51–84, esp. 55.

[19] Thomas Aquinas, *In IV Metaphysicorum*, lect. 6.

[20] Aquinas, *In IV Metaphysicorum*, lect. 6; ST, I–II, q. 94, a. 2.

[21] Aquinas, *De Veritate*, q. 1, a. 12; *ST*, I–II, q. 57, a. 2.

[22] Aquinas, *In IV Metaphysicorum*, lect. 6; *In I Posteriorum Analyticorum*, lect. 7.

This is where *intellectus principiorum* becomes decisive. The intellect does not infer the PNC from more fundamental propositions; it apprehends it immediately once the relevant terms are understood. The PNC therefore exemplifies the distinctive operation of first-principle knowledge: it is not produced by reasoning but makes reasoning possible.

The point has direct implications for artificial intelligence. An AI system may represent the PNC, detect contradictions, reject inconsistent propositions, or employ formal rules equivalent to it. Such capacities establish functional competence, but they do not by themselves establish intellectual apprehension of *ens* and *non-ens* or possession of the PNC as a first principle. The Thomistic question is therefore not merely whether an AI system can 'use' a first principle, but whether it can 'intellectually possess' one.

Ens, Intellectus, and the Structure of First Judgement

The relation between *ens* and the PNC further clarifies the structure of first judgement. Aquinas does not understand the intellect as beginning with a fully formed proposition and deriving the notion of being from it. Rather, *ens* is first apprehended in conception, while the PNC is first in the order of judgement because it arises from the intelligibility of being and non-being.[23]

The PNC is therefore neither merely a formal law detached from reality nor an empirical generalisation derived from repeated observation. Its necessity is rooted in the intelligible structure of being. It can consequently function both as a metaphysical principle concerning reality and as a logical principle governing judgement.

This distinction is decisive for assessing artificial intelligence. A computational system can be programmed or trained to maintain

[23] Aquinas, *In IV Metaphysicorum*, lect. 6.

formal consistency without thereby demonstrating possession of the metaphysical concepts on which the Thomistic account of the PNC rests. **Formal conformity to a principle is not yet evidence of intellectual apprehension of that principle.** The PNC thus provides a paradigm for distinguishing computational performance from the intellectual activity that grounds rational knowledge.

Synderesis and the Practical *Habitus Principiorum*

The practical counterpart of *intellectus principiorum* is *synderesis*. Aquinas identifies it not as a separate faculty but as a natural habit of the first principles of practical reason.[24] *Synderesis* therefore belongs to the foundational level of practical rationality, whereas conscience and prudence concern the subsequent application and direction of practical judgement. Conscience is an act through which knowledge is applied to a particular case, while prudence is *recta ratio agibilium*, right reason concerning things to be done.[25]

Synderesis thus provides the first practical principles from which practical reasoning proceeds. It should not, however, be identified with the whole of natural law. Aquinas understands natural law as the rational creature's participation in the eternal law, while *synderesis* concerns the habitual possession of the first principles through which practical reason recognises the fundamental orientation of human action.[26] The principle that good is to be done and evil avoided is therefore a foundational practical principle rather than the entirety of natural law; from such principles, practical reason proceeds towards more determinate moral judgements.[27]

---

[24] Aquinas, *ST*, I, q. 79, aa. 12–13; I–II, q. 57, a. 4.

[25] Ralph McInerny, "Prudence and Conscience," *The Thomist* 38, no. 2 (1974): 291–305, esp. 292. See Aquinas, *In II Sententiarum*, d. 39, q. 3, a. 1; ST, I–II, q. 94, a. 2; ST, II–II, q. 47, a. 6, ad 1; q. 57, a. 4. Aquinas describes prudence as

[26] Mary Christine Ugobi-Onyemere, *The Knowledge of the First Principles in Saint Thomas Aquinas* (Austria: Peter Lang AG, International Academic Publishers, 2015), 238–242.

[27] *Aquinas, ST*, I, q. 79, a. 12. See Juan Fernando Sellés, *Los hábitos intelectuales según Tomás de Aquino* (Pamplona: EUNSA, 2008), 417–425.

Aquinas also associates *synderesis* with the *scintilla rationis*, the "spark of reason", describing the natural capacity of the rational soul to apprehend first practical principles without discursive investigation.[28] This does not imply a store of fully formed moral propositions existing independently of cognition and experience. Rather, the habit is, in Aquinas' words, *quodammodo innatus menti nostrae ex ipso lumine intellectus agentis*—"in a certain way innate in our mind from the very light of the agent intellect".[29] The point is that the capacity to apprehend practical first principles belongs naturally to human rationality. Like *intellectus principiorum*, *synderesis* is not extinguished, even though its application may be impeded by error, imprudence or defective judgement.[30]

This also clarifies the relationship among *synderesis*, conscience and prudence. *Synderesis* provides the first practical principles; conscience applies knowledge to the particular act and may testify, bind, accuse, excuse or rebuke; prudence deliberates about what

[28] Aquinas, ST, I, q. 79, a. 12. Aquinas states that synderesis "incites to good" and "murmurs at evil" insofar as practical reason proceeds from first principles to discover and judge what it has discovered.

[29] Thomas Aquinas, *In II Sententiarum*, d. 7, q. 1, a. 2, ad 3. See Fernando Molina, *La sindéresis* (Pamplona: Servicio de Publicaciones de la Universidad de Navarra, 1999), 16.

[30] Aquinas, ST, I–II, q. 94, a. 1; I, q. 79, a. 12. See also Ralph McInerny, "Prudence and Conscience," The Thomist 38, no. 2 (1974): 291–305, esp. 292. Aquinas, *De Veritate*, q. 17, a. 1, ad 1. Aquinas distinguishes the infallibility of synderesis concerning universal principles from error in their application to particulars: synderesis "does not fail in the universal, but in the application to singulars."

ought to be done in concrete circumstances.[31] These are therefore complementary rather than interchangeable operations of practical rationality. The distinction also prevents *synderesis* from being reduced to a psychological feeling of moral unease: Aquinas' *murmurare de malo* expresses the orientation of the natural habit towards the good and away from what contradicts its first principles.[32]

The parallel between *intellectus principiorum* and *synderesis* completes the Thomistic account of *habitus principiorum*. Both are natural habits through which first principles are apprehended without demonstration, yet they operate in different orders: *intellectus principiorum* concerns the speculative apprehension of first principles, while *synderesis* concerns their practical counterpart[33]. Rationality, therefore, is grounded not simply in the capacity to perform inferences but in an intellectual actuality through which foundational truths are apprehended.

This distinction becomes decisive in considering artificial intelligence. An artificial system may employ a principle, generate conclusions consistent with the principle of non-contradiction, detect

---

[31] Aquinas, *ST*, I, q. 79, a. 13; *De Veritate*, q. 17, a. 1. Aquinas explicitly defines conscience as an act consisting in the application of knowledge to what one does. See De Veritate, q. 17, a. 1. See also Tobias Hoffmann, "Conscience and synderesis." In Brian Davies (ed.), *The Oxford Handbook of Aquinas*, Oxford Handbooks (2012; online edn, Oxford Academic, 1 May 2012), 255–264. https://doi.org/10.1093/oxfordhb/ 9780195326093.013.0020, accessed 24 July 2024.

[32] Aquinas, *In II Sententiarum*, d. 7, q. 1, a. 2, ad 3: "In synderesi autem sunt universalia principia juris naturalis; unde oportet quod remurmuret omni ei quod contra jus naturale fit. Sed tamen istud murmur est actus naturae."

[33] Aquinas, *In II Sententiarum*, d. 24, q. 2, a. 3, ad 5. Aquinas states that a natural habit is never lost, comparing synderesis to the habit of speculative first principles. See Juan Fernando Sellés, *Los hábitos intelectuales según Tomás* de *Aquino*, 420–425; Aquinas, *In II Sententiarum*, d. 39, q. 3, a. 1.

contradictions, prove theorems or produce recommendations within an ethical framework. Such performances demonstrate functional competence, but they do not by themselves establish possession of *intellectus principiorum* or *synderesis*. The relevant question is therefore not merely whether an artificial system can manipulate first principles, but whether it can possess them in the Thomistic sense—as principles apprehended by an intellect and constitutive of the rational activity they ground.

First Principles, Rationality, and the Possibility of Artificial Intelligence

The foregoing analysis shows that first principles are not simply propositions with which reasoning begins. For Aquinas, they are apprehended by the intellect without demonstration and constitute the foundation upon which demonstrative knowledge depends.[34] The possession of first principles therefore belongs to the conditions of rational knowledge itself. Rationality involves not merely manipulating propositions or producing correct conclusions, but an intellectual act through which foundational truths are apprehended as necessary.

Aquinas accordingly distinguishes *intellectus*, which grasps first principles, from *scientia*, which proceeds from them through demonstration. He explicitly maintains that *intellectus principiorum* is a nobler principle than *scientia conclusionum*.[35] First principles are thus not conclusions reached through increasingly sophisticated reasoning; they are the principles upon which such reasoning depends. This distinction provides the central criterion for assessing artificial intelligence: the question is not simply whether an artificial system can reason according to first principles, but whether it can possess them as principles.

[34] Aquinas, ST, I, q. 79, a. 11; *De Veritate*, q. 1, a. 12.

[35] Aquinas, ST, I–II, q. 51, a. 2, ad 3. Aquinas states: *"Intellectus principiorum est nobilius principium quam scientia conclusionum."*

This question differs from the classical question, "Can machines think?" Turing deliberately reformulated that question in terms of observable performance, shifting attention from the definition of thought to a machine's behaviour in the imitation game[36]. Contemporary AI has considerably extended such functional capacities, including logical reasoning, planning, prediction, learning and language generation.[37] These achievements demonstrate that artificial systems can perform operations that resemble important aspects of human rational activity. Philosophical accounts of AI have likewise examined whether artificial systems can represent content and participate in forms of communication without assuming that their cognitive architecture is identical to that of human beings.[38]

Functional success, however, does not by itself establish possession of first principles. An artificial system may preserve logical consistency, identify contradictions, apply rules, prove theorems or generate conclusions from premises. It may even represent the principle of non-contradiction explicitly within a formal system. Such operations demonstrate significant forms of rational performance, but they do not establish that the system apprehends the principle as a first principle.

From a Thomistic perspective, this distinction is fundamental. *Habitus principiorum* is not merely the presence of propositions within a cognitive architecture; it is the perfection of intellect through which first principles are naturally and immediately apprehended[39].

---

[36] A. M. Turing, "Computing Machinery and Intelligence," *Mind* 59, no. 236 (1950): 433–460.

[37] Luciano Floridi et al., "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations," *Minds and Machines* 28, no. 4 (2018): 689–707. 9482-5.

[38] Herman Cappelen and Josh Dever, *Making AI Intelligible: Philosophical Foundations* (Oxford: Oxford University Press, 2021), esp. 1–29.

[39] Aquinas, *Summa Theologiae* I, q. 79, a. 11; *De Veritate*, q. 1, a. 12.

The issue, therefore, is not whether an artificial system can be supplied with first principles or reliably operate according to them, but whether it possesses an intrinsic intellectual capacity through which their principial necessity is grasped.

This distinction also prevents two opposing errors. On the one hand, it would be unwarranted to deny that AI can produce rationally valid conclusions or perform increasingly sophisticated forms of reasoning. On the other hand, such performance does not by itself demonstrate possession of *intellectus principiorum*. The relevant distinction is therefore between **rational performance** and **the intellectual ground of rationality**. An artificial system may instantiate procedures whose outputs conform to rational norms; the Thomistic question is whether conformity to those norms entails possession of the principles that make rational judgement possible.

The question can consequently be stated more precisely: **Can an artificial system possess first principles in the Thomistic sense—not merely as represented rules or operational constraints, but as principles intellectually apprehended as necessary truths?** The answer to this question will depend on whether artificial systems can possess anything corresponding to the intrinsic intellectual habit that Aquinas calls *habitus principiorum*. This provides the basis for the subsequent examination of whether AI can satisfy the Thomistic conditions of genuine intellectual apprehension.

Computation, Representation, and Intellectual Apprehension

The distinction becomes clearer when computation is separated from intellectual apprehension. Computation operates through formally specified transformations of states, symbols, or representations. Contemporary AI has greatly expanded the complexity and flexibility of such operations, including systems that construct and manipulate representations in ways that can support sophisticated

reasoning.[40] But representation of a principle is not identical with apprehension of its intelligibility.

This point recalls the longstanding philosophical distinction between syntax and semantics. Searle, for example, argues that the formal manipulation of symbols does not, merely by virtue of its formal correctness, establish understanding or intentionality.[41] Whatever position one takes on Searle's broader argument, the distinction is useful here: a system may process a representation of the principle of non-contradiction without thereby understanding what the principle asserts or why it is necessarily true.

The Thomistic distinction is more precise still. Intellectual apprehension is not merely the manipulation of a representation but an act of intellect directed towards an intelligible object. First principles are therefore not merely items of information possessed by a system; they are principles because the intellect apprehends their necessity and employs them as the foundations of further judgement and reasoning.[42]

Consequently, the decisive difference between human intellect and artificial computation, on the present account, lies not in computational speed, complexity, or even the ability to reproduce particular forms of reasoning. It lies in the status of the principle within the activity of the knower. An artificial system may represent, instantiate, apply, and even discover formulations corresponding to first principles. The stronger claim—that it possesses *habitus*

[40] For the centrality of representation and computation in contemporary accounts of artificial and cognitive systems, see, for example, Mollo D. Coelho, & A. Vernazzani. "The Formats of Cognitive Representation: A Computational Account." *Philosophy of Science* 91, no. 3 (2024): 682-701. doi:10.1017/psa.2023.123.

[41] John R. Searle, "Minds, Brains, and Programs," *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–457.

[42] Aquinas, *In IV Metaphysicorum*, lect. 6; *In I Posteriorum Analyticorum*, lect. 7.

principiorum—would require showing that it intellectually apprehends first principles as principles. That is precisely the point at which computation and Thomistic intellectual apprehension cease to be synonymous.

Aquinas against a Merely Functional Account of Intelligence

Aquinas' account of intelligence resists any identification of intellect with the successful performance of cognitive functions. For him, *intellectus* and *ratio* are not two distinct powers: the intellect apprehends intelligible truth, while reason proceeds from one apprehended truth to another. Human reasoning therefore begins from what is already intellectually grasped and returns to first principles in judging what has been discovered. [43] Intelligence cannot consequently be defined adequately by the external performance of tasks alone.

This distinction exposes a limitation in a merely functional account of artificial intelligence. Functionalism identifies a mental state by the role it performs within a system rather than by the nature of its underlying constitution.[44] Such an approach can explain why an artificial system may be described as intelligent when it performs functions associated with human intelligence. It does not, however, settle the deeper question of what makes an intellect an intellect.

For Aquinas, the decisive issue is not simply what a system does but what kind of act is being performed. First principles are apprehended

[43] Aquinas, *ST*, I, q. 79, a. 8. Aquinas distinguishes intellectus, by which intelligible truth is apprehended, from ratio, by which one proceeds from one understood truth to another; he nevertheless holds that they are acts of the same intellectual power.

[44] Janet Levin, "Functionalism," *Stanford Encyclopedia of Philosophy*, Fall 2004 edition, defining functionalism in terms of the role or function performed by a mental state within a system, accessed, 3 November 2024. https://plato.stanford.edu/entries/functionalism/.

by intellect; reasoning depends upon them. A system may therefore reproduce the functional consequences of rationality without possessing the intellectual act that grounds them. The distinction is especially important in the case of *habitus principiorum*: the habit is not constituted by the successful execution of procedures but by the intellectual perfection through which first principles are apprehended.

Human Rationality and the Limits of Artificial Intelligence

The preceding analysis does not require the denial of the remarkable achievements of contemporary AI. Artificial systems can process immense quantities of information, detect patterns, generate language, optimise solutions and perform tasks that once appeared to require human intelligence. Their practical usefulness is therefore beyond serious dispute. What remains disputed is what these capacities establish about intelligence itself.

From a Thomistic perspective, they establish functional competence, not necessarily intellectual apprehension. Human rationality involves both *intellectus* and *ratio*: immediate apprehension provides the principles from which discursive reasoning proceeds.[45] This gives rationality a depth that cannot be measured solely by the correctness, speed or complexity of its outputs.

The limitation of AI, therefore, should not be stated merely as a deficiency in present computational power. A more fundamental question concerns the kind of entity that performs the operation. Human intelligence belongs to the human person and operates through an embodied intellect ordered towards truth.[46] AI, by

[45] Aquinas, *ST*, I, q. 79, a. 8; see also I, q. 79, a. 10, where intelligentia is identified with the act of understanding rather than with a distinct intellectual power.

[46] Aquinas, *ST*, I, q. 79, aa. 1, 8. Aquinas locates intellect as a power of the human soul and explains human rationality through the relation between understanding and discursive reasoning.

contrast, is an artefact designed to perform operations that can reproduce selected functions associated with human intelligence. The difference is consequently not exhausted by a comparison of present capabilities. It concerns the ontological and epistemological status of the activity itself.

This distinction also guards against two opposite errors. It would be premature to infer from present AI limitations that no artificial system could ever perform increasingly sophisticated forms of reasoning. But it would be equally unwarranted to infer from increasingly successful performance that such a system therefore possesses *intellectus principiorum*. The latter conclusion requires an argument about intellectual apprehension, not merely behavioural equivalence.

The Thomistic account thus provides a more exact criterion for the debate. The decisive question is whether an artificial system can genuinely apprehend first principles as necessary truths and thereby possess them as principles of its own intellectual activity. If it cannot, then however sophisticated its operations become, its conformity to rational structures remains distinguishable from the intellectual possession of those structures.

Conclusion

This paper has argued that Aquinas' doctrine of *habitus principiorum* provides a philosophically significant criterion for assessing claims about artificial intelligence. First principles are not merely rules incorporated into reasoning systems; they belong to the foundational activity of intellect. The *intellectus principiorum* apprehends what demonstration itself presupposes, while *ratio* proceeds from and returns to these principles.

The distinction between representation and apprehension is therefore decisive. An artificial system may represent the principle of non-contradiction, apply it correctly, identify contradictions, and produce conclusions consistent with it. Such achievements demonstrate

sophisticated computational competence. They do not, by themselves, demonstrate the intellectual apprehension of the principle as a first and necessary truth.

The argument does not establish that AI can never develop capacities presently beyond our understanding. Rather, it establishes a conceptual limit to what functional performance alone can demonstrate. Increasing computational sophistication may expand what machines can do without resolving what it means to understand. The Thomistic question consequently remains open at precisely the point where functional accounts tend to close the inquiry: **can an artificial system possess the intellectual habit by which first principles are apprehended as principles?**

If *habitus principiorum* is, as Aquinas maintains, a natural perfection of intellect rather than a collection of propositions or procedures, then the answer cannot be determined by performance alone. The deepest distinction between human intelligence and artificial computation lies not simply in what each can accomplish, but in the mode of knowing by which reality becomes intelligible. In this respect, Aquinas' account remains philosophically relevant to the contemporary debate: it shifts the question from whether machines can imitate intelligent behaviour to what must be true of a knower for genuine rationality to be possible.

Bibliography

Aquinas, Thomas. *Commentary on Aristotle's Metaphysics*. Translated by John P. Rowan. 2 vols. Chicago: Henry Regnery, 1961.

———. *On Being and Essence* (*De ente et essentia*). Translated by Joseph Bobik. Notre Dame, IN: University of Notre Dame Press, 1965.

———. *On Being and Essence*. Translated, with an introduction and notes, by Armand A. Maurer. 2nd rev. ed. Toronto: Pontifical

Institute of Mediaeval Studies, 1968.

Aquinas, Thomas. *Summa Theologiae*. Translated by the Fathers of the English Dominican Province. New York: Benziger Brothers, 1947.

———. *Truth*. Translated from the definitive Leonine text by Robert W. Mulligan, James V. McGlynn, and Robert W. Schmidt. 3 vols. Chicago: Henry Regnery, 1952–1954.

S. Thomae de Aquino. *Corpus Thomisticum: Opera Omnia*. Fundación Tomás de Aquino. Iura omnia asservantur, 2019. https://www.corpusthomisticum.org/iopera.html. Accessed 14 February 2024.

Boden, Margaret A. *Artificial Intelligence: A Very Short Introduction*. Oxford: Oxford University Press, 2018.

Bogliolo, Luigi. *Essere e conoscere*. Studi Tomistici 21. Città del Vaticano: Libreria Editrice Vaticana, 1983.

Cappelen, Herman, and Josh Dever. *Making AI Intelligible: Philosophical Foundations*. Oxford: Oxford University Press, 2021.

Clavell, L. "Il primo principio della conoscenza intellettuale." In *L'uomo e il mondo nella luce dell'Aquinate VII: Atti dell'VIII Congresso Tomistico Internazionale*, 62–73. Città del Vaticano: Libreria Editrice Vaticana, 1982.

Coelho, Mollo D., and A. Vernazzani. "The Formats of Cognitive Representation: A Computational Account." *Philosophy of Science* 91, no. 3 (2024): 682–701 https://doi.org/10.1017/psa.2023.123.

Cottier, Georges. "Intellectus et Ratio." *Revue Thomiste* 88 (1988): 215–28.

Elders, Leo J. "The First Principles of Being in the Philosophy of St.

Thomas Aquinas." *Doctor Angelicus* 3 (2003): 59–95.

Fabro, Cornelio. "Il nuovo problema dell'essere e la fondazione della metafisica." *Rivista di filosofia neoscolastica* 66, nos. 2–4 (1974): 475–510.

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, and Effy Vayena. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4 (2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5.

Gardeil, H.-D. *Introduction to the Philosophy of St. Thomas Aquinas: IV. Metaphysics*. Translated by John A. Otto. Eugene: Wipf & Stock Publishers, 2012.

Giacon, Carlo. *Verità, esistenza, causa*. Bologna: Zanichelli, 1973.

Hoffmann, Tobias. "Conscience and Synderesis." In *The Oxford Handbook of Aquinas*, edited by Brian Davies, 255–64. Oxford Handbooks (2012; online edn, Oxford Academic, 1 May 2012). https://doi.org/10.1093/oxfordhb/9780195326093.013.0020 . Accessed 24 July 2024.

Levin, Janet. "Functionalism." *Stanford Encyclopedia of Philosophy*. Fall 2004 ed. Accessed 3 November 2024. https://plato.stanford.edu/entries/functionalism/.

Maritain, Jacques. *Man and the State*. Washington, DC: Catholic University of America Press, 1998.

McInerny, Ralph. "Prudence and Conscience." *The Thomist* 38, no. 2 (1974): 291–305.

Molina, Fernando. *La sindéresis*. Pamplona: Servicio de

Publicaciones de la Universidad de Navarra, 1999.

Peillaube, E., ed. *Initiation à la philosophie de Saint Thomas*. Paris: Marcel Rivière, 1933.

Pieper, Josef. *Prudence*. New York: Pantheon Books, 1959.

Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*. 4th ed. Hoboken, NJ: Pearson, 2021.

Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–57.

Sellés, Juan Fernando. *Los hábitos intelectuales según Tomás de Aquino*. Pamplona: EUNSA, 2008.

Thomason, Richmond. "Logic-Based Artificial Intelligence." *Stanford Encyclopedia of Philosophy*. Substantive revision February 27, 2024. https://plato.stanford.edu/entries/logic-ai/. Accessed 20 July 2024.

Turing, A. M. "Computing Machinery and Intelligence." *Mind* 59, no. 236 (1950): 433–60.

Ugobi-Onyemere, Mary Christine. *The Knowledge of the First Principles in Saint Thomas Aquinas*. Austria: Peter Lang AG, International Academic Publishers, 2015.

Vansteenkiste, C. M. J. "Il metodo di san Tommaso." In *Le ragioni del tomismo: Dopo il centenario dell'enciclica 'Aeterni Patris'*, edited by Antonio Livi, Cornelio Fabro, Fernando Ocáriz, and C. M. J. Vansteenkiste, 161–96. Milano: Ares, 1979.

Westberg, Daniel. *Right Practical Reason: Aristotle, Action, and Prudence in Aquinas*. Oxford: Clarendon Press, 1994.