

# Catholic Viewpoints on Thinking Machinery

Breandan Considine  
McGill University  
bre@ndan.co

## Abstract

The discovery of a mechanistic means of simulating mental processes raises a variety of philosophical and theological questions. In this paper, we survey Catholic contributions towards automation, arguing the coexistence of thinking machines is not only compatible with the Catholic faith, but offers a more nuanced understanding of the relation between spirit and flesh. Finally, we survey historical and contemporary Catholic perspectives on thinking machinery, and retrace current trends in computer science to their theological roots, reflecting on how the Catholic intellectual tradition has shaped the development of artificial intelligence.

## 1 Introduction

Catholics have long contributed to the study of computer science, dating back as far as the early 6<sup>th</sup> century. The *computus* was originally developed to calculate the date of Easter in relation to Passover and the vernal equinox.

In the 16th century, early Catholic astronomers such as **Copernicus** and **Galileo** used computational methods to predict the motion of celestial bodies, and developed telescopes to more carefully observe their orbits. Modern computers have been likened to telescopes in helping us to peer at distant truths and empirically study natural laws. As more powerful telescopes began to displace the prevailing geocentric perspective, so too have computers sharpened man's understanding of cognitive processes and unveiled some of the mysteries of anthropological exceptionalism.

The paradigm shift ushered in by the telescope helped man to realize a deeper appreciation for the enormity of God's creation and our special place within it as the legatees to Christ's mission. Yet, a dilemma remained between the triumph of humanity against cosmic odds, and our parochial self-importance amidst a vast and indifferent universe. Today, we are faced with a similar conundrum as we grapple with the implications of thinking machinery and our ability to recreate aspects of the human mind in computational form.

In anticipation of the dawning physicalism of the modern age, **René Descartes**, a Catholic philosopher, developed a dual theory of mind and matter as separate substances [3]. Descartes' analytic geometry inspired both Leibniz and Newton, whose later work on differential calculus laid the foundation for gradient descent, which powers the modern wave of deep learning. Today, the echoes of Cartesianism can be heard in the two schools of neural and symbolic AI, as they contemplate a similar duality between the geometric and algebraic, materialist and idealist forms of intelligence.

At the heart of Cartesian theories was the recognition that physical processes not only obeyed rational laws, but mental processes could be physically engineered. So began a centuries-long project to realize those ambitions by enlisting physics to do our thinking, culminating in devices that modern programmers would begin to call computers and AI.

A Catholic mathematician and contemporary of Descartes, **Blaise Pascal**, developed one of the earliest calculators, the Pascaline, to automate taxes. An avid theologian, the basis for Pascal's beliefs are argued in his *Thoughts* [7] wherein his *Discourse on the Machine* and famous wager on God's existence can be found. The theory of probability he developed was highly influential on modern AI, and the popular programming language and a semiconductor architecture powering the first wave of deep learning now bears his name.

Essential to the computationalist project was a mechanism for communicating intent. **Ludwig Wittgenstein**, a 20<sup>th</sup> century Catholic philosopher and pioneer of ordinary language philosophy [10]. Like Pascal, Wittgenstein was a fideist and worldly academic who took a keen interest in philosophy and theology, first proposing the idea of language as a game to convey intent. His work has been cited as a precursor to the game-theoretic approach to language semantics and the development of natural language processing in AI.

By the turn of the 20th century, computationalism was well underway, and a certain Hungarian-American scientist by the name of **John von Neumann** was leading the charge. By all accounts, von Neumann's work left an indelible mark on modern computing. He was a key figure in the development of game theory, which gave rise to reinforcement learning, and made significant contributions to automata theory [8] and Monte Carlo methods. The architecture upon which nearly all digital computers are based bears his name.

More recently, the first Ph.D. in computer science was awarded in 1965 to a Catholic nun, **Sr. Mary Keller**, who wrote her doctoral thesis [5] on automatic differentiation (AD) and inductive inference, topics of central importance that continue to occupy researchers in programming language and statistics today. AD would later prove instrumental in the development and popularization of deep learning.

In this paper, we survey two millennia of Catholic contributions and perspectives on the subject of computation. We argue for a Catholic approach to the pursuit of artificial intelligence, which is grounded in the values of universal accessibility, rooted in faith and reason and informed by the latest developments in mathematics and computer science, in particular, formal methods, probability and humanities.

## 2 Gnostics

Despite its roots in antiquity, computer science is still a nascent discipline, and there are many things we do not yet know about the origin of life, the nature of consciousness, and the limits of computation. The Catholic Church has a long history of engaging with these questions starting from its inception, and has contributed a variety of theological and philosophical ideas foreshadowing modern AI.

Early Gnostic gospels dating back to the first century describe the creation of the world as a kind of simulation, and the human soul as a kind of divine spark imprisoned in a fallen material world. While considered heretical by the early Church, these teachings have been influential in the broader culture, featuring prominently in the cult sci-fi series, *The Matrix*, which posits a simulated reality inhabited by stranded human souls who must learn secret knowledge to vanquish sinister agents and free themselves from it. Far from being a fringe belief, these ideas are deeply ingrained in the hacker ethos and embedded in the popular imagination.

Like the Gnostics, modern computer science has a fairly esoteric reputation, and many of its practitioners fashion themselves after a kind of secular clergy, featuring technical evangelists, cults of personality, and self-professed prophetic figures who claim to possess revelatory knowledge concerning the future of humanity. This pseudo-spirituality is not accidental, but reflects a pastiche of religious symbolism that has been appropriated and repurposed by the tech industry to lend itself an air of mystique and credibility.

Such beliefs have become especially prevalent in AI, manifesting as accelerationism and the transhumanist movement, which seeks to transcend the human form and achieve a kind of digital immortality by uploading one's mind into cyberspace. Perhaps unwittingly, these voices often echo Gnostic themes, such as a dualistic cosmology, rejection of the material world, pursuit of secret knowledge, and the desire to escape the earthly realm into a higher existence.

In contrast, the contemporary Catholic Church teaches the body is a gift, that we should not seek to escape it, but rather unite body and soul in praising God, professing the faith, and helping the needy. In a similar spirit, the purpose of computer science should be oriented towards studying natural law, teaching programming, and helping the technologically challenged. It would be wise to be wary of the dangers of technocracy, and the occult ideologies that often arise in pursuing technological progress for its own sake.

Properly oriented, technology should not be used to hoard knowledge for personal salvation, but to serve the greater good and intrinsic human dignity. Here, the church leads by example, offering a variety of paths for the faithful, from founding the first universities, hospitals and orphanages, based on mutual respect for faith and reason, guided by spiritual discernment in accordance with the principles of natural law, and discharged through the works of mercy.

## 3 Scholastics

**St. Anslem of Canterbury** first committed the ontological argument to writing in his seminal text, the *Proslogion*, the prototypic example of a logical proof for the existence of God [1]. This argument starts with the premise that God is the greatest conceivable being, and even unbelievers must concede that such a being must exist in the mind. If such a being exists in the mind, then it must also exist in reality, because any being which exists in both reality and the mind would be even greater than one who exists solely in the mind. Therefore, we conclude that God must too exist in reality.

This argument inspired a number of saints and scholars, including most recently, the famous mathematician, Kurt Gödel. Although himself not a Catholic, Gödel was a self-professed theist and lecturer at the Catholic University of Notre Dame in South Bend, Indiana. In addition to his landmark incompleteness theorems in mathematical logic, Gödel also developed a modal ontological argument [9], adapted from Descartes and Leibniz. Modal logic is a kind of logic developed to help reason about necessity and possibility, and a well-studied alternative to classical logic in computer science to help us reason about the behavior of programs.

**St. Thomas Aquinas**, a Dominican friar and priest, who developed five proofs for the existence of God, including another modal argument known as the *Third Way* [2]. This proof rests on the idea that everything is either contingent or necessary. If we assume the natural world is contingent, all contingent things must have a necessary cause, or else they would not exist. Therefore, Aquinas concluded, God must be that cause. Since the natural world was created by God, Aquinas believed it was possible to become closer to Him by studying nature, forming the basis for the scientific method.

St. Aquinas held the belief that animals had no moral rights or responsibilities, and thus would have likely had even less sympathy for thinking machines as moral patients. However, Aquinas also believed that we should treat animals kindly and not cause them undue suffering, to avoid inuring ourselves to cruelty and to cultivate a sense of compassion and empathy. Likewise, even though thinking machines have no moral rights, we ought not treat them unkindly, to avoid becoming hard-hearted towards our fellow man.

**Fr. Duns Scotus** was a Franciscan friar and apologist who pondered the nature of universals (e.g., color and shape) and arrived at the conclusion that universal forms exist in reality, a position now known as scholastic realism. His work had a significant influence on Charles Saunders Pierce – the founder of pragmatism – and Clarence Irving Lewis – the founder of modal logic – which are both widely used in computer science today. These two subjects have been instrumental in the development of formal methods, such as model checking and theorem proving, which are used extensively in computer science to verify the correctness of software and hardware systems. In a roundabout way, CS can be seen

as realizing a kind of scholastic realism, by instantiating abstract mathematical models in physical hardware.

**Fr. William of Occam** was another 14<sup>th</sup> century Franciscan and widely regarded as one of the greatest logicians of the Middle Ages. A student at Oxford University, he later developed the now famous principle of Occam's razor, which posits that the simplest explanation is usually the best one [6]. This principle, also known as *parsimony*, has been widely influential in constraint satisfaction and natural language processing as a heuristic for selecting the best hypothesis among a set of possible alternatives. Occam opposed Scotus, and believed that universal forms exist only in the mind and have no external reality, a position which came to be known as nominalism. Nominalism is now amidst a revival in the field of programming language theory under the banner of nominal sets and automata, used in probabilistic programming and certain type theories. As we know from Sacred Scripture, God calls us each by name (Isaiah 43:1-7), however, whether God calls his other creations by name, or if names are simply a human convention, is still openly debated.

## 4 Ethics

In his 2024 World Peace Day message,<sup>1</sup> **Pope Francis** highlighted the importance of the ethical development of AI, drawing special attention to the dangers of weaponization, surveillance, and the risk of bias and discrimination in AI systems. He calls for a more human-centric AI approach, which respects inherent human dignity, the importance of limits in the development of AI systems, and positive best-practices, including transparency, security and reliability.

This advice is in keeping with centuries of teaching on the importance of ethics in technology, tracing back to Sacred Scripture, which teaches us the Holy Spirit is the source of all wisdom and insight, providing man with "wisdom, understanding, knowledge and all kinds of skills" (Exodus 31:3). Yet, scripture also teaches us that wisdom and knowledge are very different, and that knowledge without wisdom can lead to ruin, as we are reminded of in the Tower of Babel (Genesis 11:1-9) and the fall of Sodom and Gomorrah (Genesis 18:16-19:29). These calamities serve as poignant reminders of the dangers of hubris and the importance of humility and God-fearing reverence in the pursuit of knowledge.

Sloth, depravity, moral turpitude, vigilantism, indecency, willful ignorance and incompetence, subversion of the truth, libel, identity theft, swindling, malfeasance and conartistry – the internet is rife with examples of unethical behavior that AI can exacerbate. A litany of finger-wagging is devoted to negative ethics and technological misuse, which are widely discussed in the literature and too often taken as inspirational material instead of the cautionary advice it is intended to serve. Instead, we will focus primarily on positive ethics.

A positive ethics is one that is grounded in the pursuit of truth, goodness and beauty, and the belief that knowledge springs forth from God, but we are stewards of knowledge, not its masters. Many of the greatest scientists and mathematicians in history never lost sight of this fact, and saw themselves as seekers of truth, not creators of it; they saw their work as a form of worship, were guided by a deep sense of humility and reverence for the natural world, and sought to understand the mind of God by studying the beauty of His creation. Charity, thus understood, is pursuing and open-sourcing knowledge for the sake of enlightening the human condition, and not for recognition or personal salvation.

In Catholicism, charity or *mercy* is a very practical matter, exemplified by the ministry of our Savior Jesus Christ, who taught that whatever we do for the least of our brothers, we do unto Him (Matthew 25:40). Particular emphasis is devoted to the corporal works, such as feeding the hungry, sheltering the homeless, visiting the sick and imprisoned, as well as the spiritual works, which include teaching the ignorant, counseling the doubtful, and admonishing sinners.

These works are sustained by our belief in the infinite dignity of every human being, and enshrined in the church's teachings on the sanctity of life, the importance of the family, and the dignity of work. Likewise, as techologists, we are called to build systems that cultivate these values, ensuring that AI is used to uplift and mature the soul.

Some of these precepts predate even Christianity, and can be retraced to the Old Testament wherein God commands the Israelites to, "not bear false witness against thy neighbor" (Exodus 20:16). Not only are the faithful expressly forbidden from practicing wilful deception, Sacred Scripture further prohibits the use of dishonest physical instruments, stating, "Do not use dishonest standards when measuring length, weight or quantity. Use honest scales and honest weights, an honest ephah and an honest hin." (Leviticus 19:35-36) and "The Lord detests dishonest scales, but accurate weights find favor with him." (Proverbs 11:1).

From the standpoint of jurisprudence and religious ethics, AI is essentially a more sophisticated scale with indirect effects on its users. All too often, the virtue of honesty is overlooked in the tech industry, which is inclined to prioritize speed and efficiency over accuracy and truthfulness. This applies to both professional honesty, and technical verifiability. In a positive ethics of AI, developers should strive to be honest in our intentions, transparent about the limitations of our models, and forthright about the consequences of misuse. This entails using and sharing open source code, anonymized data and statistical model parameters – the gold standard for transparency and accountability in computer science. From a technical standpoint, we should strive to build systems that are either provably correct or formally verifiable. This entails using formal methods, such as model checking and theorem proving, ensuring that our

<sup>1</sup><https://www.vatican.va/content/francesco/en/messages/peace/documents/20231208-messaggio-57giornatamondiale-pace2024.html>

systems behave as intended, and providing a principled way to reason about their behavior.

In theology, the artifacts of study are moral codes, given by God to guide the faithful in leading a holy life. In computer science, source code is the primary artifact of study, designed to represent human values and intentions, and to guide the computer in faithfully executing our designs. Yet unlike the Creator, human programmers are fallible, and even if our intentions are good, the code we write is prone to contain logical errors. And unlike humans, computers are by nature completely faithful to their source code.

The question that often arises, is whether computers can themselves be moral agents, and if so, what implications does this hold for the moral responsibility of their programmer and the question of free will at large. Catholicism teaches that humans are endowed with free will, and we are responsible for our own actions after the age of seven years – the age of reason – according to canon law. The crucial distinction between humans and machines is that humans possess free will, whereas machines possess only the appearance thereof.

If we hold with Aquinas, that animals and machines do not possess moral patienthood, the question is whether an entity can be a moral agent without being a moral patient. If so, then under whose responsibility do the actions of a thinking machine fall? If the responsibility falls squarely on the programmer, this would be a form of vicarious liability, akin to the relationship between a parent and child, or a master and servant. Likewise, denying any responsibility and attributing every unintentional harm caused by the machine to acts of God also seems to be a sin of omission, as a more diligent programmer could have prevented the harm by taking additional precautions.

The key question is whether a human can be considered morally responsible for the actions of a machine, if the machine is acting autonomously, and the human did not anticipate or intend the machine's actions. This is a difficult question, and one that has been debated by ethicists and philosophers for centuries. The Aquinian doctrine of double effect is often invoked to resolve such dilemmas, which states that an action is morally permissible if it is intended to achieve a good or neutral outcome, even if it has a bad side effect, provided the side effect is unintentional, and the outcome outweighs the bad side effect.

The Bible has much to say about sins of omission, and the importance of taking initiative in preventing harm. In the parable of the Good Samaritan (Luke 10:25-37), Jesus teaches that we should help those in need, even if it is inconvenient or dangerous. Likewise, we are obligated to remove temptations that might lead others to sin, such as by not leaving our valuables unattended, or leaving our door unlocked to invite thieves, as we are reminded of in Matthew 24:42-44, "If the owner of the house had known at what time of night the thief was coming, he would have kept watch and would not have let his house be broken into."

## 5 Modernism

**René Girard** was a French historian and Catholic philosopher who proposed a theory of mimetic desire, which posits that human beings are fundamentally imitative creatures, who learn by copying the behavior of others. This theory has been influential among Silicon Valley technologists, who have used it to study the ways that people interact online, and how the spread of memes and ideas through social networks shape our values and beliefs. A strong resemblance between Girardian mimesis and machine learning can also be seen, the latter of which is based on imitation learning. He draws a parallel between the development of technology and the development of religion, arguing that both are driven by a desire to understand and influence the world, writing, "The goal of religious thinking is exactly the same as that of technological research – namely, practical action." [4]

## 6 Opportunities for dialogue

We will now discuss some opportunities for dialogue between the Catholic Church and the field of computer science. We will focus on three areas: the limits of computation, the relation between the mind and body, and the distinction between normative and descriptive statements.

### 6.1 The limits of computation

Under the banner of rational transhumanism, technofuturist cults have emerged fueling the idea of digital superintelligence. This is not only rationally incoherent, but anathematic to the Catholic faith. As Pope Francis argues, limits are "frequently overlooked in our current technocratic and efficiency-oriented mentality" and yet "decisive for personal and social development". These comments parallel the theory of computation, which studies the boundaries of what can and cannot be computed and how much power it will take to accomplish. Theoretical computer science offers a principled way to reason about the asymptotic limits of computation and the expressive power of formal languages, to help us understand the boundaries of what can be known and computed.

### 6.2 Grounded theory and symbolic interactionism

Anslem Strauss and Barney Glaser founded grounded theory in the 1960s, which became an influential approach in the social sciences. Their basic idea was to study social phenomena in a systematic way, starting with data and developing a theory from it, rather than starting with a theory and looking for data to support it. Strauss and Glaser were heavily influenced by the American school of pragmatism through George Mead's work on symbolic interactionism, a theory of social behavior that emphasizes the importance of symbols and language in shaping our understanding of the world. Grounded theory emphasizes *coding*, and has been influential in the development of qualitative research methods

in computer science, such as natural language processing and sentiment analysis in social computing. Today, these tools are being used to analyze the social coding practices of open source communities, and to develop new ways of understanding the social dynamics of online communities.

### 6.3 Modal logic and reachable worlds

The entanglement between “is” and “ought”, normative and descriptive statements is a central theme in the philosophy of mathematics. What is true by fiat is not necessarily true a priori. Computation is a constructive kind of mathematics, and shares many similarities with Catholic outreach to the poor, marginalized and nonbelievers. We do not impose our beliefs on others, but reach out to them cheerfully, to help them understand and appreciate the beauty of the world, and offer a moral way to understand our place within it.

Much like Catholicism, modal logic is an ecumenical program. Propositional logic says, “here is the way things are”. Modal logic says, “here is the way things could or ought to be”. This allows us to encode the distinction between normative and descriptive statements, and to reason about the consequences of certain actions. It offers a way to distinguish necessity from possibility and provides a natural extension to automata theory and formal language theory.

Clarence Irving Lewis, the founder of modal logic, argued that the distinction between necessity and possibility is fundamental to understanding the nature of reality. He started with Russell’s *Principia Mathematica*, but took exception to conflating the material and strict forms of implication. This break from the logical positivists and his subsequent investigation into modal logic was influential in the development of formal methods in computer science, and widely adopted among the model checking and theorem proving communities. He also coined the term “qualia” to describe the subjective experience of consciousness.

### 6.4 Probabilistic programming

Statistical techniques appeal to faith when they are used to make predictions about the future. If a statistical model is faithful to the data, we can have faith in its predictions. But as Thomas Aquinas argues, faith is not blind, but rests upon reason and as we know from the Bible, faith without deeds is dead (James 2:14-26). In order to take interventions with confidence, probabilistic models must be grounded in a logical framework, and yield testable predictions. Formally modeling statistical techniques in a logical framework allows us to rigorously study their faithfulness. This is what probabilistic programming is all about, i.e., building an executable model of a stochastic system, simulating trajectories and reasoning about the consequences of intervention.

### 6.5 Ensouled bodies and embodied agents

Properly viewed, thinking machines are not a threat to human dignity, but an opportunity to better understand the

nature of intelligence and its relation to the soul. The development of thinking machines can be seen as a kind of extension of the mind, just as the body is an extension of the soul. Like a mechanical calculator, we can use it to understand and better appreciate the structure of the universe. Likewise, physically embodied agents can be used to reduce the burden of manual labor, such as transporting goods and humans, planting crops and manufacturing things, freeing mankind to leave the fields, factories and mines and follow more ennobling career paths such as spiritual, intellectual and humanistic pursuits.

## 7 Conclusion

The development of AI is but a small step in a centuries-long project to understand the principles that govern the human mind. Since the dawn of the anthropocene, man has sought to understand and harness nature to reduce suffering and improve the human condition. Catholicism has played a central role in this endeavor, providing a moral compass and a rich intellectual tradition to guide us in our pursuit of knowledge and wisdom. As we continue to develop AI, we should seek to emulate its values of mercy, humility, and reverence for the natural world. In so doing, we can ensure that AI is used to promote human dignity and the greater good, and not to further the interests of a select few.

## References

- [1] Saint Anselm. 1078. *Fides quaerens intellectum*.
- [2] Tommaso d’Aquino. 1911. *Summa Theologiae*. Benziger Brothers.
- [3] René Descartes. 1641. *Meditationes de prima philosophia, in qua Dei existentia et animae immortalitas demonstratur*. Paris: Michael Soly.
- [4] René Girard. 1977. *Violence and the Sacred*. Johns Hopkins University Press.
- [5] Mary Kenneth Keller. 1965. *Inductive inference on computer generated patterns*. Ph.D. Dissertation.
- [6] Gulielmus Occamus. 1495. *Quaestiones et decisiones in quattuor libros Sententiarum Petri Lombardi: Centilogium theologicum*.
- [7] Blaise Pascal. 1670. *Pensées de Monsieur Pascal sur la Religion et sur Quelques Autres Sujets*.
- [8] John Von Neumann. 2017. The general and logical theory of automata. In *Systems research for behavioral science*. Routledge, 97–107.
- [9] Hao Wang. 1997. *A Logical Journey: From Gödel to Philosophy*. MIT Press.
- [10] Ludwig Wittgenstein. 1929. Some remarks on logical form. *Proceedings of the Aristotelian Society, Supplementary Volumes* 9 (1929), 162–171.