Can AI Help Us to Understand Belief? Sources, Advances, Limits, and Future Directions

Andrea Vestrucci[1,2]*, Sara Lumbreras[3], Lluis Oviedo[4]

[1] Graduate Theological Union (USA)
[2] University of Geneva (Switzerland)
[3] Pontifical University of Comillas (Spain)
[4] Pontifical University Antonianum (Italy)

Abstract

The study of belief is expanding and involves a growing set of disciplines and research areas. These research programs attempt to shed light on the process of believing, understood as a central human cognitive function. Computational systems and, in particular, what we commonly understand as Artificial Intelligence (AI), can provide some insights on how beliefs work as either a linear process or as a complex system. However, the computational approach has undergone some scrutiny, in particular about the differences between what is distinctively human and what can be inferred from AI systems. The present article investigates to what extent recent developments in AI provide new elements to the debate and clarify the process of belief acquisition, consolidation, and recalibration. The article analyses and debates current issues and topics of investigation such as: different models to understand belief, the exploration of belief in an automated reasoning environment, the case of religious beliefs, and future directions of research.

Keywords

I. Introduction

COGNITIVE science has tried since its inception to offer reliable models on how human mind works. One of these models is built on a representational approach. In this approach, ideas and mental accesses to reality are viewed as representations, and the operations and processes that result in thoughts and decisions are interpreted in computational/algorithmic terms.

Other models present alternative explanations for mind's operations different from a computational model. These other models are broadly labelled as "externalists". They assume that the computational model is not able to fully describe the complexity of mental phenomena. According to this perspective, mental phenomena involve factors that are irreducible to computational/algorithmic terms, and there are better candidates to explain some specific aspects of human cognition – for instance, the so-called embodied, embedded, enacted, and extended ("4e") theories of mind [1].

Some recent developments in robotics and information theory have increased the richness of these perspectives, but also the plurality of interpretations. An interesting example are theories that try to define consciousness as a sort of integration measure in the context of information theory. In the work of Tononi [2], several metrics are proposed to compute consciousness as a measure of how integrated information is in a system. Integrated information theory has provided

formulae for *Phi* or this integration of information, which measures the level of feedback and interaction between the components of a system. This purely mathematical view has been criticized for implying panpsychism (all systems would be conscious to some degree) and for being non-functionalist (the theory does not address the functions of consciousness or their implications).

Another example has involved robots with peripherals such as sensors that receive inputs from the world and actuators that have an impact on it. In some cases, these robots have been able to derive models of themselves from the information received from outside. This has been described as *embodied* and *embedded* systems [3]. However, it is unclear whether embodied cognition or even embodied AI might provide a convincing model to represent human mind and cognition [4].

The process of believing is a good test for computational modelling, and it invites to develop more sophistication models. Believing is a common human experience: everybody holds beliefs of different kind, related to several life contexts, with distinct ranges and applications, from simple ones – such as "I believe the weather today will be good" – to the more engaging and abstract ones – such as "I believe that my life is meaningful".

The process of believing increasingly occupies a central role in the research. This might appear as a change from previous approaches. In fact, although the believing process has long been an object of interest for epistemology, cognitive sciences, and philosophy of mind, it has often been given a secondary, or "lower", status, in contrast to "higher" or "stronger" cognitive attitudes or faculties such as "knowledge", "reason", or "intellect" [5]. This peculiar status of belief is connected with the *probability* of the truth of beliefs. Beliefs are not necessarily

true: they can be true (or false), or they can be more true than false (or more false than true), i.e., they come in various degrees of strength, certainty, and confidence about their truth [6], [7]. In other words, a belief has a truth value that is – more, or less – probable. Thus, since belief is not necessarily veridical, but only probably veridical, this aspect of human cognition has been considered to be secondary to other epistemological notions (e.g., knowledge, especially when knowledge is equated with what justifies belief [8, ch. 9]), or belief is required to satisfy some specific conditions in order to enjoy the same status of other cognitive outputs [9].

However, the probabilistic status of the truth value of belief can be an important resource, because it offers a good environment to test to what extent which computational modelling better helps to understand how beliefs arise, stabilize, and even vanish. Several proposals for computational models of belief present probabilistic estimates for belief truth value. From its very beginnings, Artificial Intelligence (AI) has provided such models. Inference engines apply logical rules to an existing knowledge to deduce new facts. Bayesian networks are able to incorporate the probabilities of facts in order to derive the probabilities of other facts related to them [10]. The present article outlines some analyses and explorations of the computation of the probability of beliefs via the application of automated reasoning programs (section IV); it also discusses whether all beliefs can be object of computation and translated into algorithmic structures, or if there are subsets of beliefs that might not be interpreted in probabilistic terms, for instance believing in values or believing in a transcendent, divine being. Fodor [11] offers some interesting arguments against a computational approach to beliefs in values, due to the character of "generality" of such beliefs in contrast with other processes which are much easier to compute. Concerning religious beliefs, they could be presented in probabilistic terms, but their transcendent status might question such codification [12].

Recent developments in AI, such as deep learning (a type of machine learning based on multiple layered artificial neural networks), have increased the expectations that AI could help us to better understand the functioning of our own minds, and thus to fill some gaps that seems to affect current computational models for cognitive processes. This might include the formation and constitution of beliefs. For instance, AI systems built on pattern recognition and machine self-learning manage to achieve tasks that could come closer to some aspects of the believing process, such as believing in something as consequence of a recurrent pattern of events, or believing as the result of learning from new data. This might purport the idea that AI systems could work in a way that is close to human mental processes, and hence they can assist in discerning about belief formation, development, confirmation, or negation.

The present article aims to explore the hypothesis that such advances in AI could help to discern in a more accurate way to what extent develop the research on the computability of mental activities such as believing process. To do so, the article analyzes recent literature on how, and to what extent, AI systems might contribute to our modelling and understanding of belief. The study of recent computational approaches to believing processes might cast a new light on the philosophical, logical, psychological, and cognitivist perspectives on belief. Section II analyzes different models of belief and believing process. In section III we seek to detect the challenges of the computational approach and to discern the extent of its heuristic potentials. Section IV deepens some recent analyses and explorations of beliefs in automated reasoning environments, with specific reference to the formalization of belief in doxastic logic, the applications of automated theorem provers, and the assessment of the skepticism about the translatability of beliefs into machine syntax. Section V deepens how a specific set of beliefs, religious beliefs, may benefit from a computational approach, and to

what extent these beliefs are irreducible to such approach. Finally, section VI outlines and discusses three directions for the future of the research on the intersections between computational modelling and the extent and improvement of our understanding of belief.

II. Modelling the Process of Believing

In this section we analyze models of belief and believing process that consider developments in computational study. Such models might foster a positive interaction between philosophy of mind and cognitive psychology, limit the risk that beliefs are studied in separate compartments, and increase the communication between fields – for instance, the integration of recent advances in the epistemology of beliefs into cognitive psychology research [13], [14], [15], [16].

At least six models have been proposed in recent years to describe the structure and the dynamic of beliefs. These six models are: the credition or "functions model" by Angel and Seitz [17] [18]; the "stages model" by Connors and Halligan [19]; the "network model" by Castillo *et al.* [20]; the "complex system model" by Lumbreras and Oviedo [21]; the "conversion model" by Smith [22]; and the "dimensions of faith model" by Donaldson [23]. We present a summary of these six models to allow for a brief comparison and to assess their computational features.

A. Creditions as Processes of Believing

The first model proposes a self-organizing system, with strong neurological roots, based on four functions: "enclosure" or integration of some basic units – perceptions, ideas – into an existing network in which they are accommodated; "converter function", which establishes belief's application range or its influences on an action course; "stabilizer function", able to keep some constant and reliable appearance despite environment changes; "modulator function", regulating the interplay between cognition and emotions.

The credition function [24] allows the individual to trust her inner probabilistic representations, and acts at two different dimensions: cognition and emotion. Credition has the key function of guiding action by means of reciprocating feedback, which involves exploration and learning. Credition is therefore an essential cognitive process to understand the human mind and behavior. It is important to underline that credition belongs not only to the realm of cognitive processing but also to the domain of subjective experience. Thus, credition is not only calculated but also experienced, as will be discussed later.

Further developments in the original pattern [25] have proposed an integrated model that considers, first, the broadly assumed dual schema that distinguishes between perceptual (or immediate) and evaluative (or analytic) processes; and, second, the "hierarchical structure of belief representation" that distinguishes between a physical, an interpersonal, and a social level. These two dimensions can be integrated to provide a structure where perception gives place to representations through action and evaluation or selection; this happens at the three mentioned levels: physical, interpersonal, and social, in a nested hierarchical process.

The described model is complex as it integrates a former and a latter schema, or reads the former four functions through two dimensions: the first describes the dynamics that link perception to selection in forming beliefs; the second records different levels influencing that configuration. In this last version, the four basic functions can be followed through some cognitive processes taking place along two distinct axes: one axis moves from perception to belief formation as somewhat an internal process of cognitive elaboration; the second axis accounting for more external influences or interaction. In this way, general functions are represented or accomplished in these two dynamic sets or components.

B. Stages of Belief Formation

Based on studies on delusion, Connors and Halligan [19] assume a functional stance to explain beliefs in terms of representations that help a subject to regulate her own behavior. They establish a five stages model of belief formation:

1. Precursor or proto-belief, which may be triggered by sensation or event, by external information or social communication, and by introspection on pre-existing beliefs.

2. Search for meaning, which "involves explaining or accounting for the experienced precursor and accommodating it within the existing web of beliefs" and avoiding cognitive inconsistency [19, p. 7].

3. Candidate belief evaluation, a process that scrutinizes new possible beliefs, after testing their explanatory power and their congruence with formerly acquired beliefs; such process is often subjected to the influence of many biases and affective states.

4. Accepting or holding the belief, or conscious assumption of the new belief as true after all the required tests, thus giving the belief a relative stability.

5. Consequential effects of holding the belief, on how new beliefs influence the world observation and judgments; this happens through a re-configuration of the existing "web of beliefs" and their fixation in memory, and determines decisions and an action course.

This schema appears as quite lineal and consistent with psycho-cognitive observation, together with neurological data. A logical process is followed from the first step, when a "proto-belief" is forming in one's mind, through its filtering and confirmation, after unconscious or conscious tests and influenced – or even biased – by cognitive mechanisms present in human mind.

C. Beliefs as Self-Sustaining Networks

Castillo *et al.* [20] describe a belief as "a network of perceptual experiences that have something in common." Inspired by complex networks, in a similar way as how ecological systems transfer energy, they intend that a belief links experiences through a transfer of meaning. In that sense, such experiences can be elaborated as perceived changes that influence or help to solve a task. Their functions are recognized to be adaptive: beliefs help to make predictions, to constrain attention, and to bridge interruptions due to variability.

Also in this model, relevant experiences need to be coupled to former ones giving rise to beliefs. The ecosystems analogy is further explored to find similar dynamics in beliefs processes: autocatalysis, or positive feedback between different system levels; circular causality, or mutual influence between single elements and final outcome, reinforcing all the system; and centripetalism, "the idea that a network will attract resources into its circuit to sustain itself" [20], amplifying its relevance to confirm its acquired positions and to exclude conflicting ones.

This model allows to explain how the system reaches some stability despite fluctuations in the environment and some aspects characteristic of beliefs, such as storage, retrieval, and apparent agency. At the same time, this schema helps to understand how beliefs emerge and change. In analogy to how new systems emerge from dissipating gradients and reaching new equilibria, beliefs emerge to gain new meaning after different experiences interact and become coupled. Change is the result of perceived mistaken beliefs, often too locally linked. However, due to the self-enforcing nature of belief networks, conflicting evidence is not enough to explain change, which happens only when the old belief stops to be perceived as beneficial for the entire system, giving rise to a new network and more efficient coupling. This process points to a degree of order present in such systems, which can be conceptualized, as in other similar networks, always as something taking place spontaneously, i.e., as order emerging from chaos in living systems.

This new model takes advantage of cybernetic dynamics that count with a consistent tradition. As such, a belief is always seen as an ordered set able to link or couple experiences, and, by the same token, broader beliefs can be conceived as sets that are coupled in a systemic or coherent way, addressing some adaptive tasks, or covering some function in relationship with their environment.

D. Beliefs as Complex Systems

Recently, beliefs have also been likened to complex systems [21]. A complex system is a self-organized ensemble of reiterated, uncontrolled multiple interactions between a plurality of components. Examples of complex systems include Earth climate system, ant colonies, and the Web [26].

This analogy integrates some of the principles of belief systems as networks, and expends them to make sense of the dynamics of belief within the wider scope of complex systems. As mentioned, a complex system is an entity composed of many interacting components. Even if the components are relatively simple, the behavior of the system is difficult to predict due to the emergence of new phenomena in the system. Examples of complex systems can be found in a wide variety of context, from engineering to biology or finance, but, regardless of their specific context, all complex systems share the same properties, such as nonlinearity, emergence, spontaneous order, adaptation, and feedback loops [27]. The power of complex system studies lies in how general these properties are, and how they help explain very different phenomena. For instance, from a complex system perspective, the behavior of a flock of birds trying to advance and protect itself while avoiding internal collisions could be linked to that of a group of firms trying to develop their companies according to established business models while avoiding excessive competition.

From the complex system perspective, the main properties of belief are: being goal-oriented, openness, complexity, spontaneous order, and adaptation. The goals of belief have been studied in the literature under differing lights [28]. Belief serves at least three different purposes: it provides a model of the world and anticipates the consequences of action; it filters new evidence and establishes priorities for the decisions; and it defines what is important, what is a priority, what should or should not be done. Beliefs are open because they receive inputs from the interaction with the environments or with other individuals ([29], [30], [31]). They are complex because of the number of factors that influence them: personal features such as analytic cognitive style [32], feelings of superiority [33] or even parenting styles [34] have been shown to influence belief formation and change. Belief networks, as complex systems, are subject to nonlinear phenomena. Nonlinearity means that the same stimuli does not lead always to the same response. For instance, it takes more information to change beliefs than to confirm them. In belief networks, change is generally difficult but, in times of crisis, the change in one belief can spread to a large number of them. This is true for personal and for social beliefs, where the crisis dynamic could be explained as a paradigm shift [35]. Emergence means that new properties and structure originate from the system. In addition, beliefs organize in more or less consistent and related spheres of influence. Consistent belief systems have been described by some as "attractors", states towards which a system tends to evolve [36]. Finally, the adaptation property is displayed by belief systems as they evolve to fulfil their objectives of providing a model of the world, filtering experience and guiding action.

This model can be useful to integrate many of the properties of belief and to anticipate some of the phenomena that might emerge in a manner that can be subject to empirical testing. Still, this model is in its early steps and its development could lead to interesting insights.

E. Beliefs as Functional Maps

Aaron Smith [22] proposes a model more akin to religious beliefs and inspired on sequential models of religious conversion. His foundation lies in a functional or adaptive view of beliefs attuned to social connection, risk detection, and life navigation. He devises five components in that process: "concepts" including relevant representations and forms of recall; "computation" that works like an "engine of belief" or mechanism that process information to produce inferences; seven "iterative mechanisms", from "personal identification" through repetition and practice to reasoning and confirmation; "commitment" or personal assimilation, which includes unconscious reinforcement and rewards by promises and expectations; and, finally, "consequences" or positive effects in terms of social benefits or cultural innovation by trust.

This model is complex and not just sequential, since some loops and mutually enforcing means are present, especially at the model's core, i.e., the "computation" component, which includes several cognitive mechanisms working together to prompt beliefs, including religious beliefs.

F. Dimensions of Faith

For Steve Donaldson [23], faith is a general psychological attitude that consists in attributing a probability to the existence of something or to an event or expectation. As mentioned, believing means considering the probability for something to be true. This is a central cognitive function that contributes importantly to give sense to our world and to make decisions. Clearly, beliefs aim to be true. In order to confirm the truth of belief or to present evidence for it, several mechanisms may contribute (e.g., personal introspections, emotions of fear or pleasure, rational claims), and several factors may endanger the process, such as the lack of rational interest.

Religious beliefs are viewed as a type of belief that shares common elements or similar structures with other cognitive systems [37]. As such, by studying beliefs we acquire a heuristic tool that applies to a broader spectrum, from economy to emotions, from science to religion.

Donaldson classifies beliefs according to three levels of observation: primary or immediate; secondary or mediated by other means; and tertiary or resulting from reflection. He then establishes a scale to determine the levels of certainty and how beliefs that are "known for sure" are different from those that represent values.

The general idea in the six models that we mentioned above is that beliefs follow their specific logic and can be modelled based on some basic characteristics and relationships. All six proposals assume a functional stance: beliefs help us engage with the world and with our relationships in and with it. All six models apply a cognitive psychological framework, and in some cases the computational aspects emerge as a part of a complex system that includes many other dimensions or components. For instance, in both "network model" and "conversion model", cybernetics and computation appear as one stage of a global process. In these models, computation can play a role in beliefs formation, but believing cannot be reduced to just computational means or processes, or to algorithmic elaboration of collected information. The critical point is to what extent AI developments could help to shed more light on this aspect or could reveal some aspects still hidden in our ability to form and hold beliefs. This topic is analyzed in the next section.

III. Making Place for AI in the Believing Process

In this section we attempt to see if advances in the field of AI and its applications may help the research on belief and provide insight into the believing process, and whether AI is in competition, or in consonance, with the previously-analyzed models of belief. The first step is to distinguish for what AI might be relevant, and for what it might not play any role. This might beg the question about what lies beyond the realm of computation. Nevertheless, it might be useful to advance a proposal built on available views that clarify the affinities and distinctions AI and human cognition.

A good starting point is the recent book by Brian Cantwell Smith [38]. The book attempts to discern, through an in-depth knowledge of AI systems, what is specifically human in our way to know and to deal with the world. The central point is the distinction between the concepts of reckoning and judging. Reckoning is the type of calculation provided by current AI systems; judging is the human kind of decision making, based on the evaluation of circumstances, events, knowledge, and also beliefs. This "judging" is built on ethical values, existential insights and projects, wisdom, and the distinction between what is actual and what is possible, what is real and what is apparent.

According to Cantwell Smith, the cognitive capabilities of humans and AI systems are different, and they can hardly overlap. AI systems deal with discrete representations of the world that are processed through algorithms. On the other hand, humans engage directly with the world through non-linear representations and projects, and by formulating propositions that have non-dualistic truth value – that is, propositions that are neither 100% true nor 100% false. Amongst those propositions or statements there are the so-called "beliefs", statements that have a truth-value that is probabilistic.

Although some have proposed that embedded AI bridges the gap between what could be associated to Cantwell Smith's concepts of "reckoning" and "judging" [39], the mere inclusion of sensor inputs on a process might not be equivalent to cover the full spectrum of elements that constitute human believing process [40]. Could increasing the number of sensors lead to anything qualitatively different? Along the same lines, fuzzy logic has been used to formalize the probabilistic truth value of belief [41]. However, fuzzy logic might be considered to be not fully able to capture all traits of human believing process. In such process, the probability of a belief might also be connected to emotional states or prior beliefs. It would be needed a much more sophisticated fuzzy logic than, for instance, the one currently used in engineering contexts.

What lessons can we learn from this "cognitive" distinction between human and machine? Let us imagine a "believing machine", an AI system able to generate beliefs from a large number of data or information.

This hypothetical believing machine can be helpful to better assess some hidden processes in human belief formation. If we ask how this machine would work, we can formulate the following points:

- The believing machine would be a system able to collect all relevant information or inputs through pattern recognition. The machine would be able to distinguish between what is relevant and what is not for a belief formation. It would filter inputs and prior beliefs based on rules of contiguity and causality.

- The believing machine would decide through the right algorithms an output from the totality or a selection of the collected information. This task would need specific statistic tools, such as probability calculation. Hence, the system would act as a predictor, by using prior information to predict the outcomes of actions. As such, the machine would be a support for decision-making.

- The process of decision-making could make use of machine learning systems guided through positive and negative feedbacks from the application of acquired beliefs. Thus, the believing machine would filter new evidence and would evolve the belief system to fitting with prior beliefs (including the emotional investment of the subject) and new information.

All these processes could also be constitutive parts of human believing process.

However, contrary to the human situation, this machine would not engage with the beliefs it generates; it would be indifferent for it to believe in a belief or in a different one. On the other hands, we humans put much at stake in our beliefs, we are committed to (some of) them and such commitment influences the way we think, judge, and act in the world. Believing in the righteousness of a value *x* (for instance, eco-responsibility, or gender equality) affects our life in a very different way than not believing in the righteousness of this value *x*.

Moreover, our beliefs are constantly changing: they are confirmed or dismissed, enforced or discredited by ideas, experiences, relationships, and introspection. They demand reformulation and adaptation to new life situations. The process of recalibration of beliefs involves impressions, emotions, cultural and social circumstances, personal commitments, etc. In other words, beliefs are experienced by a subject and have an object – believing is a subjective experience, analogously to "qualia" (the subjective and conscious experiences, such as the sensation of cold or heat). As such, the believing process is deemed to be simultaneously a cognitive and an emotional process – as the credition model underlines. This makes difficult to compare humanly experienced belief to mechanical processes undertaken by an AI. Furthermore, the case made by Cantwell Smith regarding the distinction between machine and human cognitive capabilities might even introduce a second level of belief: believing that we humans and machines are radically different precisely as far as beliefs are concerned, and that this difference is positive and good for human flourishing.

Probably the believing machine would also need external assistance. For instance, pattern recognition requires a previous work of tagging by an operator who identifies and labels relevant objects or information. Even if statistical methods could be applied in its most sophisticated way, interpretation of the results would demand a further consideration and judgment. Moreover, feedbacks can be ambiguous and complex: some beliefs could result in double effect actions, and again some discernment based in judgment and a broader view would be needed – something close to what Cantwell-Smith calls "wisdom".

A machine able to generate beliefs could become a good heuristic tool. This machine could even improve at the point to incorporate a number of functions to the point of coming close to human believing process. However, it might be hard to imagine an algorithmic translation of the aspects of interpretation, judgment, and commitment that qualify human beliefs.

IV. Computation of Belief Via Doxastic Logic

In section II we analyzed some theories of belief that are conceived by using terminology borrowed from Computer Science such as network, modelling, system, etc. In section III we outlined some distinctions between human and machine cognitive capabilities by focusing on the formulation of beliefs and on how those beliefs impacts human life.

Now, it is time to analyze current attempts of positive interactions between belief and computer programs. Those attempts concern the application of automated theorem provers to assess the epistemic value of beliefs, that is, to calculate the probability for a belief to be true (or false), and thus to modify the belief in order to improve its probability of being true. The application of such programs to beliefs requires an intermediate step: the translation of a belief into a formal language that can be understood by the syntax of the machine. This formal language is provided by doxastic logic, the logic that deals with opinions and beliefs.

Doxastic logic is a subset of modal logic, the logic that formalizes possibility and necessity. Possibility and necessity are the "modes" of the truth-values of a proposition: a proposition can be necessarily true (or false), for instance once the proposition is demonstrated, or a proposition can be possibly true (or false), for instance before that the proposition is demonstrated. The link between modal logic and belief is the following: as stated in sections I and III, beliefs are expressed in propositions that have a truth-value that is probabilistic. Moreover, as already mentioned, the probabilism of beliefs' truth-value is precisely what distinguishes beliefs from other human cognitive attitudes and faculties. Given that a belief is expressed by a proposition that is probably true (or false), then it is possible for this proposition to be true (or false). Therefore, the proposition expressing a belief can be translated into the formal language of modal logic, since this is the logic that studies the possibility or necessity of the truth-values of propositions. "Doxastic logic" is the name of the field of modal logic that studies the formalization of beliefs.

The study of the logical formalization of beliefs dates back to the 1950s. One of the early most famous work on this is [42]. In this seminal work Hintikka applied possible world semantics to the logical study of knowledge and belief. Possible world semantics interprets possibility and necessity (the two operators of modal logic) as quantifiers over possible worlds: necessary is a proposition that is true in all possible worlds, and possible is a proposition that is true in some possible worlds. Thus, the application of possible world semantics to knowledge and beliefs is another way of saying that epistemic logic (the logic of knowledges) and doxastic logic (the logic of beliefs) are subsets of modal logic.

Since then, the scientific community has witnessed an exponential growth of the research on the extent and the limits of the logical study of beliefs [43]. The objects of this research are multiple: how to formalize the connection between beliefs and their premises or presuppositions; if, and how, the statements about beliefs can be axiomatized; the extension of the logical investigation from the mere content of belief to other connected topics, such as the purpose of belief, the consequences of believing something, and the justification of beliefs; (connected to the previous point) the way to recalibrate and correct beliefs via the interaction with other agents (believers) or the acquisition of new information; the logical treatment of the lack of belief, for instance in the sentence "there is something that I neither believe nor disbelieve" [44].

The logical study to beliefs has the worth of refining our understanding and insight on beliefs. This includes the clarification of some logical issues affecting belief or defining its epistemic specificity, for instance if belief is compared to knowledge. One of those logical issues concerns the possibility (or even the necessity) for beliefs to be inconsistent, that is, to entail a contradiction with its premises. For an overview of the varieties of inconsistency that occur in beliefs, see [45]. Moreover, the research has deepened the logic at the basis of the improvement, awareness, or progressive resolution of such logical issues of belief from the standpoint of the *believer* itself; this is a kind of dynamic epistemic (or doxastic) logic [46].

The development of the logical study of belief and of the processes of producing, enriching, and modifying beliefs provides a formalization of beliefs and believing processes. This formalization uses the operator *B* for "belief" and variables for subjects and objects of belief. A standard formalization is the following formula: $Ba\gamma$, which reads: "The subject *a* believes γ". This formalization is in first-order doxastic logic, because the operator *B* ("belief") is applied to an object. However, doxastic logic can also be of higher-order, in case the operator *B* is applied to itself, as in formulas that are built on a nested doxastic operator *B*, for instance the formula $Ba\gamma \rightarrow BaBa\gamma$, which reads: "If the subject *a* believes γ, then *a* believes to believe γ". Thus, high-order doxastic

logic involves believing about beliefs. Higher-order doxastic logic is useful to study problems concerning the inconsistency of a belief, or to formalize processes of epistemic reflection over one's own belief (usually called "introspection").

The formalization of beliefs and believing processes in higher-order modal language can be used as input for automated reasoning systems in order to compute the probability of truth-value of such beliefs and believing processes. Recent attempts in this direction have used the following higher-order automate theorem provers: LEO-II, TPS, Stallax, and IsabelleP[1]. Such experiments have shown the different strengths of the four theorem provers, given that not all epistemic and doxastic problems were solved in the automated reasoning environment [47]. This would lead to a potential improvement of the theorem provers.

Another example is the computation of degrees of plausibility/possibility of beliefs via a ratio between sets of possible worlds [48]. In this case, the degree of plausibility of a belief was computationally checked via the application of the model checker Mc-COGWED on belief translated in the language of a specific logic, the COmputationally Grounded WEighted Doxastic logic (COGWED).

These experiments present examples of positive interaction between (formalization of) beliefs and computation. On the one hand, they show that beliefs and believing process – when correctly formalized in the language of high-order doxastic logic – can indeed be translated into algorithms and, thus, be computed. As such, these experiments provide useful insights on the logical consistency of beliefs, the degree of probability of their truth value, and the extents and mechanisms of modifying and improving the beliefs. On the other hand, the different degrees of solvability and complexity of doxastic problems involving beliefs invited to develop more effective and higher-performance theorem provers. Since higher-order theorem provers represent a fundamental field in AI research [49], the applications of theorems provers to formalized beliefs strongly invite to an interdisciplinary cross-fertilization between three areas of research: research on AI, philosophical research on beliefs, and research on the epistemology and logic of belief.

However, as already hinted in the previous sections, it is possible to detect an apparent limit in the computational approach to beliefs: the idea that the formalization of beliefs in doxastic logic is possible only via a simplification of the actual believing process. This simplification concerns the fact that beliefs are understood as propositions of which it is possible to calculate the degree of plausibility and to check the veracity, and, thus, to evaluate on which extent beliefs can count as knowledge.

This criticism harkens back to the distinction – hinted in the previous section – between the cognitive capabilities between human and AI systems as far as beliefs are concerned. In fact, the simplification of belief might invite to purport that the machine is confined to operate only upon a limited number of beliefs, the ones that can be expressed in formal language and translated into machine syntax [50]. This would imply the exclusion of all beliefs which content is not limited to events that can or cannot be, and therefore that cannot be reduced to a calculation of the probability of their truth value. These beliefs include, again, the belief in values, emotions, personal virtues and weaknesses. Religious beliefs pertain to this domain too: they are beliefs on beings

that are transcendent, uncaused, independent from our mind, and that nevertheless play an important role in our existences, affecting our choices, actions, and lives.

In sum, the skepticism about the translatability of beliefs into machine syntax focuses on the idea that formalizations of beliefs might not distinguish between specific contents of belief – since formalization treats such contents as variables. However, in our life some beliefs are more important than others, last longer than others, or have more important consequences or a stronger impact than others, precisely for their specific content. For instance, a belief such as "John is sick" can be different depending on the relationship between John and the subject. Such criticism would invite to disregard all attempts to translate beliefs into computational language since such attempts seem to consider only a simplification of the complexity of beliefs.

However, it is useful to consider at least two counterarguments to these criticisms.

The first counterargument concerns the fact that doxastic logic deals not only with contents referring to events, but also with self-awareness. This is the above-mentioned case of introspection, formalized by formulas built on nested doxastic operators ($Ba\gamma \rightarrow BaBa\gamma$). As such, (higher-order) doxastic logic studies beliefs that have mental states – and not only events – as their objects. Higher-order doxastic logic could represent the difference in importance or complexity of beliefs in terms of the degree of "nestedness" of the doxastic operator. Moreover, dynamic doxastic logic deals with complex forms of believing processes such as belief change, belief revision, and complex forms of belief such as collective belief. The complexity of beliefs is precisely the material for the current advances in doxastic logic [46].

The second counterargument refers to the specificity of religious beliefs. We deepen it in the next session.

V. The Case of Religious Beliefs

As stated in the previous section, a criticism that questions the relevance of the application of automated reasoning programs to the understanding of belief concerns the risk of losing the specificity of religious beliefs as beliefs in transcendent entities, i.e., in entities that are abstract, uncaused, and whose existence is independent from human mind.

There is a counterargument against this criticism: religious beliefs are indeed beliefs in something, that is, they have an object as much as beliefs in events. As such, nothing impedes to present a formalization also of such beliefs.

In fact, the γ (the content of belief) in the formula $Ba\gamma$ can easily be interpreted as a religious content, "Vishnu's existence", "God's omnipotence", "deity x" or "property y of the deity x". Nothing impedes the computation of such belief. This includes the computation of arguments in support of religious beliefs – i.e., arguments that claim to prove the validity of the attribution of the property y to the deity x object of a religious belief.

The computational translation of arguments in support of a religious belief (based upon a formalization of such arguments) can be important from the point of view of the epistemic introspection of the religious believer because it might help to distinguish between what is strictly necessary and what is not necessary in the logical structure of the argument. In other terms, the computational translation of an argument supporting a religious belief might help to detect what is redundant in the non-computational version of the argument, thus clarifying the belief itself. The consequence of this operation is the increase in self-awareness of a belief, and, thus, the possibility to improve the consistency of the argument supporting the belief.

---

[1] Concerning IsabelleP, see [47, p. 122]: "The higher-order proof assistant Isabelle/HOL is normally used interactively. In this mode it is possible to apply various automated tactics that attempt to solve the current goal without further user interaction. Examples of these tactics are blast, auto, and metis. It is also possible to run Isabelle from the command line, passing in a theory file containing a lemma to prove. Finally, Isabelle theory files can include ML code to be executed when the file is processed. While it was probably never intended to use Isabelle as a fully automatic system, these three features have been combined to implement a fully automatic Isabelle/HOL, called IsabelleP".

We present two experiments. The first experiment is proposed by Oppenheimer and Zalta, on the wake of the program in computational metaphysics [51]. Oppenheimer and Zalta applied the theorem prover Prover9 to their axiomatization and formalization of Anselm's ontological argument for the existence of God [52], [53]; the result of this application is the discovery that Prover9 needs less lemmas and premises to prove the argument than the ones required in the humanly-formalized version of the argument. This discovery is fundamental to assess the logic of the ontological argument, and, thus, to deepen the extent and limits of the soundness, plausibility, and justification of the belief of such existence.

The second experiment focuses on the work led by Christoph Benzmüller: it consists in the application of high-order theorem provers to a formalization of Gödel's ontological argument [54], [55]. Also in this case, the application of automated reasoning programs led to a simplification of the logical structure of the argument. This discovery provides an incomparable help to deepen the soundness, plausibility, and meaning of believing in an entity (called "God") that possesses all positive qualities at the highest degree.

However, it is possible to question whether these experiments truly address the issue of the specific content of religious belief. As stated, this specific content are entities that are transcendent, i.e., that are abstract, uncaused, and that exist independently on human brains, and that nevertheless affect human lives. It might seem that the aspect of "affecting human life" is completely missed in the two experiments mentioned above. In light of this impact that religious beliefs have on the life of the believers, religious beliefs can be considered part of the big family of "existential beliefs", i.e., beliefs that provide meaning and purpose of human existence. The specificity of religious belief is precisely to have as object a transcendent entity that is source of existential meaning.

Let us harken back to the believing machine of section III. This machine would be able to generate religious or spiritual beliefs, in the same way in would generate beliefs about events, people, politics, economy. The difference would be that the formulation of religious beliefs would imply the distinction between transcendence and immanence. Is this distinction just a minor issue, something that could be easily programed, or is it something unassailable for a machine?

It seems to be hard to conceive a system which tags an event or information as "transcendent" or "supernatural". But it seems to be even harder to conceive a system which recognizes the existential value of transcendence, in the same way as a religious or a spiritual mind is able to do. In sum, what seems to be difficult is to build a *self*-transcending machine. In fact, according to Cantwell Smith, AI systems cannot refer to something external, even less if this "externality" is radically external, i.e., beyond the physical world, "transcendent". Such machine could only assist discerning when something moves to this transcendent level, thus requesting more information about what this transcendent level is about. And anyway, in no case the machine would be able to grasp the existential meaning of this transcendent thing, i.e., the connection between this transcendent with the existence of the machine itself. It would seem that the capacity of transcendence marks a limit for AI systems, and adds a new entry to the list of specific human cognitive features described by Cantwell Smith.

However, it is important to underline a possible ambiguity with the term "transcendence". We can understand it in two ways: as a term that refers to something that lies beyond the physical realm, i.e., something abstract; or we can understand "transcendence" as referring to something that lies beyond the limits of human intelligence, and, thus, beyond the limits of human language.

In the former case, it is worth mentioning again the program in "computational metaphysics". Computational metaphysics is "the

implementation and investigation of formal, axiomatic metaphysics [...] in an automated reasoning environment" (http://mally.stanford.edu/cm/). Axiomatic metaphysics is an axiomatic theory of abstract objects [56], [57]. Thus, if we understand "transcendence" as "set of abstract objects", then our understanding of these abstract objects can indeed be computed, and the experiment by Oppenheimer and Zalta supports this.

On the other hand, if "transcendence" refers to something that lies beyond the limits of language, then there are two options: either there is no possible linguistic formulation of this transcendence, or this transcendence shows the limit of language. In the first case, the object of belief cannot be expressed by language, then our belief in such transcendence is void because it cannot be formulated. In the second case, the limit of language is still stated by language, e.g. in the sentence "The transcendent object x marks the limit of language". The linguistic formulation of the limit of language implies the distinction between object language and metalanguage: a metalanguage is a language that speak about another language called "object language". Now, to be coherent with the definition, transcendence shows the limit not only of a given object language, but of every possible metalanguage. Therefore, the discourse on this transcendence (a discourse called "theology") is a discourse on the structure of the relationship between object language and metalanguage – a relationship that is at the basis of any possible logical endeavor. Given that this discourse is in principle formalizable [58], nothing impedes that the "belief" in this metalanguage-limiting transcendence is formalizable in the syntax of a machine.

There is also an alternative way to conceive a positive interaction between religious belief and computation. This approach conceives religious beliefs from a decision-making perspective. Rather than focusing on the epistemic aspect of belief, this approach concerns the practical aspect of belief: the modifications and improvements of one's course of action in light of the influx that a specific belief has on the determination of future actions. In this practical approach, the focus switches from "believing what" or "how/why believing what" to "believing, and then doing what". In other words, this approach defines the specificity of religious belief not by referring to a specific (more or less satisfactory) connection with the epistemic requirements of belief, but by referring to the aspects of commitment, decision, choice of action that are the manifestations, outputs, or expressions of one's faith [59], [60]. This approach would contribute to the interaction between machine and belief by connecting the believing process to the research on the computation of decision-making processes [61].

VI. Directions of Future Research

In light of what analyzed, we see at least three directions of future research:

1. The first direction concerns fostering the exploration of the complexity of beliefs in an automated reasoning environment. This includes several points: 1.1. Applying automated reasoning programs to different forms of belief might encourage the dialogue between, on one hand, the research in doxastic logic and dynamic epistemic logic and, on the other hand, philosophy of mind and cognitivist psychology: this interdisciplinary dialogue would better assess what aspects and types of belief have yet to be formalized in doxastic terms. 1.2. (connected to the previous point) Developing the investigation of beliefs in an automated reasoning environment helps to better clarifying what precisely is the "existential" aspect of belief, e.g., what are its epistemological specificity, and what is the specific practical impact of "existential beliefs" on our decision-making processes. 1.3. This first direction of research would also improve the understanding of the distinction of different types

and subtypes of beliefs, for instance as a development of what was presented by Hadley [62]; this would contribute to intersect cognitive science and AI on the topic of belief, and it would be a good starting point for integrating computational modelling in the research on the epistemology of specific kinds of belief – such as religious beliefs [63].

2. The second direction of future research focuses precisely on religious beliefs and the interactions between religious statements (as expressions of religious beliefs) and automated reasoning programs. This includes presenting other applications of theorem provers to other arguments issued from religious beliefs (e.g. *a posteriori* arguments, theological paradoxes, deontic arguments on divine justice, etc.). This direction is simultaneously close and distinct from some recent contributions in (and on) analytic theology [64]: analytic theology aims to "press philosophical tools into theological service" [65, p. 475], while this direction of research aims to apply computational tools for theological service. Such "theological service" consists in detecting redundancies, improving coherency, and reassessing the validity of theological arguments within an axiomatic framework. On the wake of the program in computational metaphysics, this direction of research is called "computational theology" [66]. Moreover, recent research focuses on the relationship between magic and technology [67]; it will be useful to deepen the use of AI in the sociological context of magic as a way to clarify the distinction and analogies between religion and magic.

3. The third direction of research focuses on how the study on the extents and limits of interactions between AI systems and beliefs can contribute to the current debate on the definition of belief systems understood as collections of beliefs with different contents. One example of belief system is religion. The limits that affect all competing definitions of religion – substantive/ontological, functionalist [68], [69], etc. – can be better framed via the deepening of the specificity of logical and computational aspects of belief, including the computational understanding and clarification of the arguments in support of such beliefs. This might have positive applications to the recent discussion on the consonances between religious belief and mathematical realism [70].

These three directions of research might even open to advances in AI developments. The challenge to apply automated reasoning programs to doxastic problems might encourage the development and improvement of these programs themselves [47]. Moreover, the three directions of research might provide elements for fostering the question of the place of belief in scientific research, and the research on the relationship between religion and science.

Will it be possible to write algorithms able to express the complexity of our believing activities and processes? Or will the richness of the spectrum of beliefs, and in specific religious beliefs, prove to be a limit to computability? Whatever the answer might be, as far as it is not tested, it is only a matter of opinion – better, it is only a matter of *belief.* Thus, the best course of action is to foster the multidisciplinary interactions and consonances between the research in AI and the investigation on believing processes, so to provide strategies to test our hypotheses, and to come up with conclusions that are at least provisionary.

References

[1] A. Newen, L. de Bruin, and S. Gallagher, Eds., *The Oxford Handbook of 4e Cognition*, Oxford, U.K.: Oxford University Press, 2018, doi: 10.1093/oxfordhb/9780198735410.013.45.

[2] G. Tononi *et al.* "Integrated Information Theory: From Consciousness to Its Physical Substrate," *Nature Reviews Neuroscience*, vol. 17, pp. 450-461, 2016, doi: 10.1038/nrn.2016.44.

[3] M. Hoffmann and R. Pfeifer, "Robots as Powerful Allies for the Study of Embodied Cognition from the Bottom Up," in *The Oxford Handbook of 4e Cognition*, A. Newen, L. de Bruin, and S. Gallagher, Eds., Oxford, U.K.: Oxford University Press, 2018, pp. 841-862, doi: 10.1093/oxfordhb/9780198735410.013.45.

[4] R. Manzotti, "Embodied AI beyond Embodied Cognition and Enactivism," *Philosophies*, vol. 4, no. 3, 2019. Accessed: Aug. 1, 2021. [Online]. Available: https://www.mdpi.com/2409-9287/4/3/39.

[5] M. Ayers and M. R. Antognazza, "Knowledge and Belief from Plato to Locke," in M. Ayers, *Knowing and Seeing: Groundwork for a New Empiricism*, Oxford, U.K.: Oxford University Press, 2019, ch. 1, pp. 3-33, doi: 10.1093/oso/9780198833567.003.0001.

[6] L Eriksson and A Háyek,"What Are Degrees of Belief?," *Studia Logica*, vol. 86, pp. 183-213, 2007, doi: 10.1007/s11225-007-9059-4.

[7] F. Huber, "Belief and Degrees of Belief," in *Degrees of Belief*, F. Huber and C. Smith-Petri, Eds.,Berlin/Heidelberg, Germany: Springer, 2009, ch. 1, pp. 1-33, doi: 10.1007/978-1-4020-9198-8.

[8] T. Williamson, *Knowledge and Its Limits*, Oxford, U.K.: Oxford University Press, 2002, doi: 10.1093/019925656X.001.0001.

[9] M. Schulz, "Strong Knowledge, Weak Belief?," *Synthese*, 2021. Accessed: Aug. 1, 2021. [Online]. Available: https://link.springer.com/article/10.1007/s11229-021-03180-x.

[10] F. V. Jensen, *An Introduction to Bayesian Networks*, London, U.K.: UCL Press, 1996.

[11] J. Fodor, *The Mind Doesn't Work that Way*, Cambridge, MA, USA: MIT Press, 2000.

[12] R. Swinburne, *The Existence of God*, 2nd ed., Oxford, U.K.: Oxford University Press, 2004, doi: 10.1093/acprof:oso/9780199271672.001.0001.

[13] W. Spohn, *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford, U.K.: Oxford University Press, 2014, doi: 10.1093/acprof:oso/9780199697502.001.0001.

[14] M. Smith, *Between Probability and Certainty: What Justifies Belief.* Oxford, U.K.: Oxford University Press, 2016, doi: 10.1093/acprof:oso/9780198755333.001.0001.

[15] H. Leitgeb, *The Stability of Belief: How Rational Belief Coheres with Probability.* Oxford, U.K.: Oxford University Press, 2017, doi: 10.1093/acprof:oso/9780198732631.001.0001.

[16] L. Moretti, *Seemings and Epistemic Justification: How Appearances Justify Beliefs*, Cham, Switzerland: Springer, 2020, doi: 10.1007/978-3-030-43392-5.

[17] H.F. Angel, "Religiosität als menschliches Potential. Ein anthropologisches Modell der Religiosität im neurowissenschaftlichen Horizont," in *Religiosität: Anthropologische, theologische und sozialwissenschaftliche Klärungen*, H. F. Angel *et al.*, Eds., Stuttgart, Germany: Kohlhammer, 2006, ch. 5, pp. 62-89.

[18] H. F. Angel, L. Oviedo, R. F. Paloutzian, A. L. Runehov, and R. J. Seitz, *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, Cham, Switzerland: Springer, 2017, doi: 10.1007/978-3-319-50924-2.

[19] M. H. Connors, P. W. Halligan, "A Cognitive Account of Belief: A Tentative Roadmap," *Frontiers in Psychology*, vol. 5, 2015. Accessed: Aug. 1, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01588/full.

[20] R. D. Castillo, H. Kloos, M. J. Richardson, and T. Waltzer, "Beliefs as Self-Sustaining Networks: Drawing Parallels Between Networks of Ecosystems and Adults' Predictions," *Frontiers in Psychology*, vol. 6, 2015. Accessed: Aug. 1, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01723/full.

[21] S. Lumbreras and L. Oviedo, "Belief Networks as Complex Systems," *Limina: Grazer theologische Perspektiven*, vol. 3, no. 2, pp. 92-108, 2020, doi: 10.25364/17.3:2020.2.5.

[22] A. Smith, *Thinking about Religion: Extending the Cognitive Sciences of Religion*, Basingstoke, U.K.: Palgrave MacMillan, 2014, doi: 10.1057/9781137324757.

[23] S. Donaldson, *Dimensions of Faith: Understanding Faith Through the Lens of Science and Religion*, Eugene, OR, USA: Wipf & Stock, 2015, doi: 10.1111/heyj.13085.

[24] H. F. Angel and R. J. Steiz, "Process of Believing as Fundamental Brain Function: The Concept of Credition," *Research Bulletin of the Sigmund Freud PrivatUniversität Wien*, vol. 3 no. 1, 2016, doi: 10.15135/2016.4.1.1-20.

[25] M. Sugiura, R. J. Seitz, H. F. Angel, "Models and Neural Bases of the Believing Process," *Journal of Behavioral and Brain Science*, vol. 5, no. 1, pp. 12-23, 2015, doi: 10.4236/jbbs.2015.51002.

[26] J. Ladyman, J. Lambert, K. Wiesner, "What is a Complex System?," *European Journal for Philosophy of Science*, vol. 3, pp. 33–67, 2013. doi:10.1007/s13194-012-0056-8.

[27] M. Mitchell, *Complexity: A Guided Tour*, Oxford, U.K.: Oxford University Press, 2011.

[28] J. D. Frank, "Nature and Functions of Belief Systems: Humanism and Transcendental Religion," *American Psychologist*, vol. 32, no. 7, pp. 555–559, 1977, doi: 10.1037/0003-066X.32.7.555.

[29] J. Crocker, S. T. Fiske, and S. E. Taylor, "Schematic Bases of Belief Change," in *Attitudinal Judgment*, J. R. Eiser, Ed., Springer Series in Social Psychology, New York, NY, USA: Springer, 1984 pp. 197-226, doi: 10.1007/978-1-4613-8251-5_10.

[30] N. Rodriguez, J. Bollen, and Y. Y. Ahn, "Collective Dynamics of Belief Evolution Under Cognitive Coherence and Social Conformity," *PloS One*, vol. 11, no. 11, 2016. Accessed: Aug. 1, 2021. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0165910.

[31] B. Sodian, D. Zaitchik, and S. Carey, "Young Children's Differentiation of Hypothetical Beliefs from Evidence," *Child Development*, vol. 62, no. 4, pp. 753-766, 1991, doi: 10.1111/j.1467-8624.1991.tb01567.x.

[32] G. Pennycook, J. A. Cheyne, P. Seli, D. J. Koehler, and J. A. Fugelsang, "Analytic Cognitive Style Predicts Religious and Paranormal Belief," *Cognition*, vol. 123, no. 3, pp. 335-346, 2012, doi: 10.1016/j.cognition.2012.03.003.

[33] K. Toner, M. R. Leary, M. W. Asher, and K. P. Jongman-Sereno, "Feeling Superior is a Bipartisan Issue: Extremity (Not Direction) of Political Views Predicts Perceived Belief Superiority," *Psychological Science*, vol. 24, no. 12, pp. 2454-2462, 2013, doi: 10.1177/0956797613494848.

[34] T. Ruffman, J. Perner, and L. Parkin, "How Parenting Style Affects False Belief Understanding," *Social Development*, vol. 8, no. 3, pp. 395-411, 1999, doi: 10.1111/1467-9507.00103.

[35] K. Jones, "Some Epistemological Considerations of Paradigm Shifts: Basic Steps Towards a Formulated Model of Alternation," *The Sociological Review*, vol. 25, no. 2, pp. 253-272, 1977, doi: 10.1111/j.1467-954X.1977.tb00289.x.

[36] B. Goertzel, "Belief Systems as Attractors," in *Chaos Theory in Psychology and the Life Sciences*, R. Robertson and A. Combs, Eds., New York, NY, USA: Psychology Press, 1995, ch. 9, pp. 123-134, doi: 10.4324/9781315806280.

[37] A. Fuentes, *Why We Believe: Evolution and the Human Way of Being*, New Haven, CT, USA: Yale University Press, 2019.

[38] B. Cantwell Smith, *The Promise of Artificial Intelligence: Reckoning and Judgement*, Cambridge, MA, USA: MIT Press, 2019, doi: 10.7551/mitpress/12385.001.0001.

[39] D. McDermott, "Artificial Intelligence and Consciousness," in *The Cambridge Handbook of Consciousness*, P. D. Zelazo, M. Moscovitch, and E. Thompson, Eds., Cambridge, U.K.: Cambridge University Press, 2007, ch. 6, pp. 117-150, doi: 10.1017/CBO9780511816789.007.

[40] S. Lumbreras, *Respuestas al transhumanismo: cuerpo, autenticidad y sentido*, Madrid: Digital Reason, 2020.

[41] M. Daňková and L Běhounek, "Fuzzy Neighborhood Semantics for Multi-agent Probabilistic Reasoning in Games," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, IPMU 2020, Communications in Computer and Information Science, vol 1239, M.-J. Lesot *et al.*, Eds., Cham, Switzerland: Springer, 2020, pp. 680-693, doi: 10.1007/978-3-030-50153-2_50.

[42] J. Hintikka, *Knowledge and Belief: An Introduction to the Logic of the Two Notions*, 2nd ed., V. F. Hendriks and J. Symons, Eds., London, U.K.: College Publications, 1962.

[43] R. Rendsvig, J. Symons, "Epistemic Logic," *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., 2021. Accessed: Aug. 1, 2021. [Online]. Available: https://plato.stanford.edu/archives/sum2021/entries/logic-epistemic/.

[44] Y. Ding, "On the Logic of Belief and Propositional Quantification," *Journal of Philosophical Logic*, 2021, doi: 10.1007/s10992-021-09595-8.

[45] D. Perlis, "The Role(s) of Belief in AI," in *Logic-Based Artificial Intelligence*, J. Minker, Ed., Boston, MA, USA: Springer, 2000, pp. 361-374, doi: 10.1007/978-1-4615-1567-8_16.

[46] A. Baltag, B. Renne, "Dynamic Epistemic Logic," *Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., 2016. Accessed: Aug. 1, 2021. [Online]. Available: https://plato.stanford.edu/archives/win2016/entries/dynamic-epistemic/.

[47] C. Benzmüller, "Combining and Automating Classical and Non-Classical Logics in Classical Higher-Order Logics," *Annals of Mathematics and Artificial Intelligence*, vol. 62, no. 1, pp. 103-128, 2011, doi: 10.1007/s10472-011-9249-7.

[48] T. Chen, G. Primiero, F. Raimondi, and N. Rungta, "A Computationally Grounded, Weighted Doxastic Logic," *Studia Logica*, vol. 104, pp. 679–703, 2016, doi: 10.1007/s11225-015-9621-4.

[49] K. Bansal, S. M. Loos, M. N. Rabe, C. Szegedy, and S. Wilcox, "HOList: An Environment for Machine Learning of Higher-Order Theorem Proving," in *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, CA, 2019. Accessed: Aug.1, 2021. [Online]. Available: http://proceedings.mlr.press/v97/bansal19a/bansal19a.pdf.

[50] J. Leach, *Mathematics and Religion: Our Languages of Sign and Symbol*, West Conshohocken, PA, USA: Templeton Press, 2010.

[51] B. Fitelson and E. N. Zalta, "Steps Towards a Computational Metaphysics," *Journal of Philosophical Logic*, vol. 36, pp. 227-247, 2007, doi: 10.1007/s10992-006-9038-7.

[52] P. Oppenheimer and E. N. Zalta, "A Computationally-Discovered Simplification of the Ontological Argument," *Australasian Journal of Philosophy*, vol. 89, no. 2, pp. 333–350, 2011, doi: 10.1080/00048401003674482.

[53] P. Oppenheimer and E. N. Zalta, "On Anselm's Ontological Argument in *Proslogion* II," submitted for publication.

[54] C. Benzmüller and D. Fuenmayor, "Computer-Supported Analysis of Positive Properties, Ultrafilters and Modal Collapse in Variants of Gödel's Ontological Argument," *Bulletin of the Section of Logic*, vol. 49, no. 2, 2020, doi:10.18778/0138-0680.2020.08.

[55] C. Benzmüller and D. Fuenmayor, "Can Computers Help to Sharpen Our Understanding of Ontological Arguments?" *Mathematics and Reality, Proceedings of the 11th All India Students' Conference on Science Spiritual Quest*, IIT Bhubaneswar, Bhubaneswar, India, 6-7 October, 2018, pp. 195-226, doi: 10.13140/RG.2.2.31921.84323.

[56] E. N. Zalta, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel, 1983. Available: http://mally.stanford.edu/abstract-objects.pdf.

[57] E. N. Zalta, *Principia Logico-Metaphysica*, 2021, unpublished. Accessed: Aug. 1, 2021. [Online]. Available: http://mally.stanford.edu/principia.pdf.

[58] A. Vestrucci, "Metalanguage and Revelation: Rethinking Theology's Language and Relevance," *Logica Universalis*, special issue "Theological Discourse and Logic", S. Krajewski and M. Trepczyński, Eds., vol. 13, no. 4, pp. 551–575, 2019, doi: 10.1007/s11787-019-00236-y.

[59] L. Buchak, "Faith and Steadfastness in the Face of Counter-Evidence," *International Journal of Philosophy of Religion*, vol. 81, pp. 113–133, 2017, doi: 10.1007/s11153-016-9609-7.

[60] L. Buchak, "Can It Be Rational to Have Faith?," in *Probability in the Philosophy of Religion*, J. Chandler and V. S. Harrison, Eds., Oxford, U.K.: Oxford University Press, 2012, ch. 12, pp. 225-246, doi: 10.1093/acprof:oso/9780199604760.003.0012.

[61] M. Calder *et al.*, "Computational Modelling for Decision-Making: Where, Why, What, Who and How," *Royal Society Open Science*, 2018. Accessed: Aug. 1, 2021. [Online]. Available: https://royalsocietypublishing.org/doi/10.1098/rsos.172096.

[62] R. F. Hadley, "The Many Uses of 'Belief' in AI," *Minds and Machines*, vol. 1, pp. 55-73, 1991.

[63] L. Oviedo and K. Szocik, "Religious–And Other Beliefs: How Much Specificity?" SAGE Open, January 2020. Accessed: Aug. 1, 2021. [Online]. Available: https://journals.sagepub.com/doi/full/10.1177/2158244019898849.

[64] O. D. Crisp and M. C. Rea, Eds., *Analytic Theology: New Essays in the Philosophy of Theology*. Oxford, U.K.: Oxford University Press, 2009, doi: 10.1093/acprof:oso/9780199203567.001.0001.

[65] O. D. Crisp, "Analytic Theology," *The Expository Times*, vol. 122, no. 10, pp. 469–477, 2011, doi: 10.1177/0014524611408533.

[66] A. Vestrucci, "Computational Theology and Natural Theology," presented at the 2021 conference of the Ian Ramsey Center *Natural Theology in the 21st Century*, University of Oxford, Oxford, U.K., 15-17 July 2021.

[67] L. Obadia, "Moral and Financial Economics of 'Digital Magic':

Explorations of an Opening Field," *Social Compass*, vol. 67, no. 4, pp. 534-552, 2020, doi:10.1177/0037768620950237.

[68] V. S. Harrison, "Internal Realism and the Problem of Religious Diversity," *Philosophia*, vol. 34, no. 3, pp. 287-301, 2006, doi: 10.1007/s11406-006-9029-5.

[69] K. Schilbrack, *Philosophy and the Study of Religions: A Manifesto*, New York, NY, USA: Wiley Blackwell, 2014.

[70] V. S. Harrison, "Mathematical Objects and the Object of Theology," *Religious Studies*, vol. 53, no. 4, pp. 479-496, 2017, doi: 10.1017/S0034412516000238.

Andrea Vestrucci

Ph.D. in ethics from the University of Lille SHS and the University of Milan, and Th.D. in systematic theology from the University of Geneva. He was a professor of ethics and logic at the Federal University in Fortaleza, Brazil, and currently he is a researcher in systematic theology and ethics for the Graduate Theological Union, Berkeley. He is also a privatdozent for the University of Geneva. He is a laureate of the Academic Society of Geneva. He was the recipient of the Australian Award by the Australian Federal Government. He also taught and was researcher at Monash University, the University of Münster, and the University of California Berkeley. His current research focuses on the ethics of Artificial Intelligence, the interactions between religion, logic, and technology, and the ethical and methodological aspects of interreligious dialogue. Amongst his recent publications: the book Theology as Freedom: On Martin Luther's De servo arbitrio (Mohr Siebeck, 2019), the paper "Recalibrating the Logic of Free Will" (Theology and Science, 2020), the edition Religions and Languages: A Polyphony of Faiths (Sophia, Springer, 2021).

Sara Lumbreras

Sara Lumbreras holds a PhD and a MSc Eng from Universidad Pontificia Comillas. She is a professor at the Institute for Research in Technology and teaches at the Industrial Management Department at the ICAI School of Engineering and the Financial department at the ICADE School of Business and Law. She is currently Deputy Director of Research Results at the Institute of Research in Technology. Her research focuses on the development and application of decision support techniques for complex problems, mainly in the energy sector and in particular in grid design, but also in finance and the health sector, where she applies Artificial Intelligence to the prediction of disease evolution. She works with classical optimization techniques (such as Benders' decomposition), metaheuristics and Machine Learning. She has five years of experience in the private sector (JPMorgan London). In addition to her technical research, she studies transhumanism and the implications of Artificial Intelligence in anthropology.

Lluis Ovideo

Lluis Oviedo is a full Professor for Theological Anthropology at the Pontifical University Antonianum of Rome, and Fundamental Theology at the Theological Institute of Murcia (Spain). He is team member of research group on Creditions, based in Graz University (Austria). He edits a book series on "New approaches to the scientific study of religion" (Springer) and the bibliographic bulletin ESSSAT News & Reviews. Research interests focus on the dialogue between theology and sciences, including the new scientific study of religion, and more recently to the interaction between religious belief and AI, and to how religious beliefs and behaviors are related to personal and social wellbeing.