

Article

Triadic Relationship of Intelligence, Information, and Complexity

Marcin J. Schroeder 

Faculty of International Liberal Arts, Akita International University, Akita 010-1292, Japan; mjs@gl.aiu.ac.jp

Abstract

This is a companion paper for the short paper published in the proceedings of the 1st International Online Conference of the Journal *Philosophies*, presenting the view that the multiple, diverse manifestations of human, artificial, and natural intelligence do not preclude a uniform characteristic of intelligence, which could be called meta-intelligence, and that this uniform characteristic can be found in the triadic relationship of intelligence, information, and complexity, in which intelligence is considered the capacity to overcome complexity of information. While the conference paper was focused on overcoming the view of a diversity of unrelated forms of intelligence through the lifting of the level of abstraction of the study, and with the use in its background of auxiliary concepts of information and its complexity, the present paper defends, but only indirectly, the same vision of meta-intelligence, but with a shift of the focus to the metaphysical study of a triadic conceptual framework with its relationship to free will, agency, causality, and interaction, and the necessary high level of generality for the concept of information. The arguments used in the conference paper for the conceptualization of intelligence are not repeated here. Instead, this paper provides comprehensive metaphysical and epistemological foundations for the study of the triad of intelligence, information, and complexity, which in the earlier paper were only fragmentary.

Keywords: intelligence; meta-intelligence; information; complexity; agency; causality

1. Introduction

My paper “Intelligence as the Capacity to Overcome the Complexity of Information: Search for Unity in the Diverse Forms of Intelligence” [1], presented at the First International Online Conference of the Journal *Philosophies* [2], and published in its proceedings, formulated and defended the idea that the increasingly popular and common view, identified in multiple publications documented in references to the literature on the subject, that the uniform concept of intelligence (usually called general intelligence) should be separated into multiple intelligences does not preclude the legitimacy of the search for a unified concept of intelligence at a higher level of abstraction, which could be called meta-intelligence, characterizing the entire variety of multi-intelligent systems. This should not be confused with the search for general intelligence, conceptualized as a specific set of behavioral characteristics of objects or systems qualifying them as intelligent, but rather considered as a search for what is common to all diverse forms of multiple intelligences, each with its own characteristics and criteria for detection and assessment. Thus, there may be no object or system that has the characteristics of all forms of intelligence, but there could be a common quality or capacity present in all its manifestations. The present work is



Academic Editors: Rossella
Lupacchini and Soraj Hongladarom

Received: 8 April 2026

Revised: 11 June 2026

Accepted: 24 June 2026

Published: 28 June 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

an extension or companion of the conference paper, which had as its primary objective the justification for searching for such unity and offering an answer based on the association of intelligence with information and its complexity, but without entering the study of the metaphysical and epistemological foundations of such an association.

The paper presented at the conference promoted the view that intelligence, when considered in a broad perspective, can be understood as the capacity to overcome the complexity of information. In this view, the diversity of the forms of intelligence comes from the diversity of manifestations of information and their diverse forms of complexity. For instance, in the early conceptualization of intelligence, understood as proficiency in solving problems, information was identified with its linguistic form, and overcoming its complexity required the use of logical methods involved in reasoning carried out in a language identical with the language of the study of intelligence; emotional intelligence can be traced to the context of issues in human communication going beyond the use of language, convoluted with the information involved in the introspection of our own emotions. Another example, going beyond not only human language but also entirely beyond the human context, can be found in the case of the intelligent behavior of organisms at their diverse levels of evolutionary advancement, optimizing their strategies to survive or reproduce by the selective use of complex information provided by their sensory mechanisms in interactions with their environment.

It is unlikely that the association of intelligence with information and its dynamics could be objected to. Much more exposed to criticism is the way in which intelligence is associated with information through the concept of complexity. All three concepts are subject to never-ending discussions of their diverse conceptualizations. Any discussion of the triadic relationships involving all of them may seem even more difficult. There is nothing wrong with this diversity of conceptualizations. This simply means that all three are philosophically non-trivial, and their triadic involvement makes the discourse deeper. However, this also means that all three require a careful explanation of how they are to be understood and what their metaphysical status is.

This understanding can be achieved by conceptual discipline and the use of logically proper definitions, or, if necessary, by axiomatic formulation when the high level of abstraction requires the involvement of categorial, i.e., primitive, concepts. The most important thing is to avoid the assumption that the familiarity and omnipresence of terms in vernacular, common-sense communications provide an excuse for not identifying their meaning. The present work is intended as a completion of the original proposal, presented in the short conference paper [1], to identify intelligence with the capacity to overcome the complexity of information by shifting focus from intelligence discussed in terms of vaguely sketched concepts of information and complexity to the triadic relation of the three precisely defined concepts, justifying such an identification. Since the main objective of this work is to clarify the metaphysical status of intelligence, the study involves several concepts present in the literature on the subject, such as free will, agency, causality, action, and interaction.

While the present paper presents the triadic relation between intelligence, information, and complexity as a metaphysical and epistemological background for the study of intelligence, it is not intended as a defense of the exclusive role of this triad in a more general philosophical discourse. It asserts the need for transcending the usual perspective of binary relations between the fundamental concepts, for instance, intelligence–information or intelligence–agency, but in this respect there are examples of much earlier philosophical studies in which the binary view of reality was objected to. A distinctive example can be found in the works of Stéphane Lupasco and his followers on the so-called logic in reality, in which the triad of reality, complexity, and the logic of the included middle serves as a

foundation [3]. The present paper is also not intended as a defense of the special role of triadic relations in general philosophical discourse, but rather as an identification of the most fundamental configuration of concepts necessary for the study of intelligence from the perspective of overarching multiple intelligences.

As explained above, this paper is a companion paper to the earlier published paper presented at a conference, which was limited by the required restricted volume to presenting the argumentation for the main thesis that the diversity of the forms of intelligence is necessary, but this does not preclude the study of the unity of their common characteristics that could be considered as meta-intelligence, very different from what is frequently called general intelligence in the context of human and artificial intelligence. The conference paper clearly claimed (using empirical studies as one of the arguments) that it is unlikely that all multiple intelligences can be implemented in a single entity or system. On the other hand, the present paper not only refers to the published conference paper, but it is also written to complement the claims of the earlier one. For this reason, to make the present paper self-sufficient, it is necessary to provide an explanation of the relationship between the two papers.

Because of its directly declared subservient function, there is no reason to include a broader explanation of the reasoning behind the thesis of the conference paper, which is clearly outside of the objectives of the present paper. Here, the objectives are to provide a rationale for bringing together the triad of intelligence, information, and complexity for its use in the study of intelligence and to exclude the concept of causality, which is frequently, but fallaciously, as claimed and explained in this paper, involved in the studies of intelligence through the concept of agency understood as exercising causal influence.

Thus, the content of this paper does not address the thesis of the preceding conference beyond its relatively brief presentation, nor does it repeat any already published arguments. The arguments used in the present paper to demonstrate relevant metaphysical and epistemological interrelations between intelligence, information, and complexity, together with a criticism of the involvement of causality, are twofold. Some of them are diachronic (historical), some are synchronic. Their role is not to provide arguments for the formulation of meta-intelligence as presented and argued in the already published conference paper, but to provide the metaphysical background uniting them into a triad and to illustrate their interrelations by non-trivial examples (obviously not all of them, which are too numerous even for an entire book).

Because all three ideas of intelligence, information, and complexity have various conceptualizations, the present paper includes an explanation of how information and complexity are understood here (including why they should be conceptualized at a high level of abstraction) with references to more extensive and detailed studies. The preceding conference paper has a more extensive explanation for the understanding of multiple intelligences and their gradual emergence from the original, century-old study of human general intelligence, whose uniformity has gradually dissolved. None of this is repeated here.

The paper is organized into five parts (including the Introduction and Conclusions), with the three central sections divided into subsections, each with its own function very briefly explained here.

Section 2 explains methodological aspects of the study and their reflection in the triadic relationship with a focus on intelligence.

Section 2.1 High Generality of Intelligence Requires High Generality of Information

This section explains the methodological rule used in this paper for the necessity of the use of a genuinely higher level of abstraction instead of its declaration. In this paper, this higher level of abstraction is achieved by the use of an appropriately high level of abstraction of the concept of information.

Section 2.2 Study of Information as a Meta-Study of Language

Since intelligence is identified in non-human contexts, it has to be conceptualized with a sufficiently abstract concept of information exceeding the generality of languages, both natural and formal. This means that the study of information has to supervene the study of languages.

Section 2.3 Conditions for Being Intelligent

This section prepares the transition from epistemological concerns to the study of the ontological status of intelligence. Its main purpose is to clarify the question about what can be intelligent, or more precisely, what the general conditions for predication of intelligence are.

Section 2.4 Necessary Conditions for Being Intelligent—Causality, Interaction, and Agency

This section justifies the focus on the triad of intelligence, information, and complexity, but not a tetrad including causality. Causality is shown to be secondary to interaction, which in turn is secondary to information, if information is sufficiently general. This has important consequences for the concept of agency based on causality that loses its meaning. Thus far, agency has been considered the main intellectual tool for the studies of intelligence.

Section 2.5 Intelligence, Agency, and Free Will

This section excludes the rescue of the traditional view of agency in terms of free will.

Section 2.6 Agency and Information

The idea of agency can be reintroduced into consideration when conceptualized in terms of interaction and information.

Section 2.7 Intelligence as a Characteristic of Entities

This section is the next step towards the ontology of intelligence and the prevention of the popular but fallacious claims, giving it a primary ontological status indicated as a noun with reference to an entity, instead of an abstraction from the predicate (adjective) “intelligent”.

Section 3 presents the concept of information as it is conceptualized in this study.

Section 3.1 Information—Level of Conceptual Generality

This section is a more detailed analysis of the need and conditions for the generality of the concept of information that has been identified as the main contender for the primary concept with respect to both intelligence and complexity.

Section 3.2 One–Many Opposition

This subsection presents the role of this opposition with a very long intellectual tradition and an important contemporary role in setting the foundations for the conceptualization of the most general ideas of the exploration of reality, including mathematical set theory. The subsection justifies the exclusion of causality from the foundation of intelligence.

Section 3.3 Information as Identification of Variety and Logic of Information

This section introduces the concept of information in terms of the one–many opposition and set theory.

Section 3.4 Relationship with Other Conceptualizations of Information

This section demonstrates the supervenience of the one-many concept of information over other conceptualizations.

Section 4 provides the presentation of the relations between complexity and the other two members of the triad.

Section 4.1 Complexity vs. Simplicity out of the One–Many Opposition

This section introduces and presents complexity as a predicate of information.

Section 4.2 Paradigmatic Example of the Reduction of Complexity

This section provides an example of the reduction of complexity that can serve as a paradigm for the study of intelligence.

Section 4.3 Examples of Formal Reduction of Information Complexity

This section provides other examples of the formal reduction of complexity of special importance for the study of intelligence. They arise in the context of human intellectual activity, but due to their formality and use in the exploration of reality, they can be used in generalizations beyond human intelligence.

Section 4.4 Other Less Formal Forms of Reduction of Information Complexity

The last part of the paper, Conclusions, returns to the definition of meta-intelligence proposed in the earlier conference paper with some ideas for further study.

2. Methodological Guardrails for Developing a Conceptual Framework

2.1. *High Generality of Intelligence Requires High Generality of Information*

Several instances of confusion regarding the meaning of the term “intelligence” have already been identified in the conference paper preceding this study [1]. There is no need to repeat their description and to search again for the sources of confusion. Instead, this section presents a more general, methodological reflection on potential fallacies in the discussion not only of intelligence, but also of information and complexity.

One of the most common and persistent fallacies in establishing conceptual frameworks for studies of cross-disciplinary subjects is the belief that the concepts can become general simply by a declaration of generality or by using the name “general theory” for the entire conceptual framework. This self-reference error of qualification as general is similar to a much better-known error of self-referential declaration of truth or falsity (the latter leading to the famous Liar Paradox).

The claim that a concept is general when it is defined using concepts of limited generality, or more frequently, using undefined concepts imported from vernacular terminology, which seem obvious and, as such, do not require any definition, is an obvious logical error. Almost always, the commonly used terms that seem obvious have multiple, very different, strongly context-dependent meanings, and each of these meanings has a very limited generality. However, there is a more compelling reason for caution in entering self-referencing predicates into any theory. Nicholas Asher and Hans Kamp reinterpreted Alfred Tarski’s proof of his Theorem on Indefinability of Truth by observing that it involves not just the predication of truth, but any self-referential predicates, and therefore all such predicates cannot be defined within the theory without leading to inconsistency [4]. Moreover, they showed not only this, but also that they cannot be added to the theory without leading to inconsistency. Since they banished self-referential predicates, i.e., predicates that can be applied to the propositions of the theory, as not definable or that cannot be added, this excludes the use of such predicates in the predication of any terms as fallacious.

The expression “general intelligence,” hotly debated in the current discussions of artificial intelligence, is a very good example of such a fallacy. Thus, the use of the qualification “general” in this expression does not make it general, and in practice it is either meaningless or has many meanings, each heavily dependent on its context and the narrow criteria for qualification within it.

On the other hand, overgeneralization can be equally faulty, as was demonstrated in the conference paper [1,5]. In extreme cases of the search for a sufficiently general concept of information, we can find desperate statements of the type “We conclude that we are not

able to say confidently of anything that it could not be information" [6]. This means that the concept of information is spurious and should be eliminated, which, of course, was contrary to the intentions of the authors, who continue to elaborate on its meaning.

2.2. *Study of Information as a Meta-Study of Language*

In this paper, the issue of the proper choice of generality is of special importance, particularly the choice of generality of the concept of information. As explained above, the diversity of the forms of intelligence conceptualized through the concept of information is an obvious consequence of the level of generality of the latter concept. Since the main objective of this work is the search for common characteristics of the diverse forms of intelligence present in multiple contexts, it is necessary to use a very high level of generality for the concept of information. For instance, if intelligence is considered as a characteristic of an artificial system (i.e., a human-designed and -produced system) or a natural system whose functioning is independent of human intervention, the meaning of the concept of information must be free from anthropomorphizing. Both cases require a very high level of generality.

The requirement of a high level of generality for the concept of information has an additional, rarely considered reason. Whatever the meaning of information considered in the study of information, it must include the paradigmatic case of human informal/natural or formal languages as special cases. Since the study (every study) is carried out in one of such languages, this means that the language of the study has two functions: an object and a means of study. This requires a very careful methodological control of the distinction between the two functions and of consistency between the general statements about information that can be understood as a metatheory and their meaning in special instances of information in the statements of the study [7]. This is not self-reference understood in the traditional way, which is focused on the consistency of the meaning of the lower-level sentence in the object language with the meaning of the higher-level sentence in metalanguage. Here, the situation is different, more complex, as we do not consider the consistency of particular sentences, but rather consider the consistency of theories: the object theory and its metatheory. Unlike the case of self-reference of sentences leading to the Liar Paradox or other paradoxes of self-reference, the relationship between the metatheory and theory is not problematic or paradoxical as long as theories are consistent. Even in the paradoxical mixing of an object sentence and a metalanguage sentence, "This sentence is false," if the word "false" is replaced by "true," the paradox is avoided. Self-reference leads to a paradox because of the inconsistency. However, the control of consistency requires a higher level of generality of metatheory. Asher and Kemp showed that merging two theories is not sufficient and may lead to inconsistency [4]. This is why Tarski's hierarchy of theories based on the distinction of object and meta-languages is necessary [7].

The relationship between the studies of information and language and the role of the former as a meta-study of the latter justifies the focus in philosophy of information on analytical methods. It is not necessary to agree on the status of analytical philosophy or on the division of philosophy into analytical and continental, itself a complex meta-philosophical problem [8], to accept the need for a very careful approach to building a conceptual framework for the study of information reflecting the experience acquired in the study of language and logic. Consequently, in this paper, the concept of information will be considered at a very high level of generality, and its study will involve a wide range of analytical methods, including mathematical formalization. Also, the dangers of the convolution of information and language, with its role as a tool of study, justify the need for methodological guardrails when intelligence is associated with information and its complexity.

2.3. Conditions for Being Intelligent

The early study of intelligence had its motivation in the need for a normative assessment of the mental skills characterizing human beings in performing difficult actions of special importance for society. Thus, the source was in the commonly acknowledged differences between individuals, especially those that could not be explained by social status, extent of education, age, experience, etc. Some of these differences, such as differences in physical strength or skills, were obvious and easy to test. Some others, such as differences in mental skills, did not have overt manifestations and were more difficult to test or evaluate. This means that the general orientation of the initial study of intelligence was more in the direction of differentiation than searching for synthesis. The latter was necessary only for the purpose of establishing a background for performing divisions. The most important division was influenced by the separation of the two realms, that of mind or soul and that of material objects. The latter was easier to comprehend by using our sensory experience, and it influenced the way the latter was studied.

There was a long philosophical tradition, today usually associated with John Locke and his metaphor of the “tabula rasa” (blank slate) with which we are born in our heads and on which our life (more exactly sensory) experience writes the content of our mind [9], but going back to Aristotle [10], who used the same metaphor. In this empiricist tradition, the distinction of intelligence from other mental skills did not have context-independent meaning. The revolution came with the recognition by Immanuel Kant in his *Critique of Pure Reason* of the presumably innate human faculty of reason unifying the effects of experience and being capable of synthetic a priori judgments [11]. Although Kant’s philosophy had both followers and critics, this opened the door to the study of reason (giving the capacity for reasoning) without restriction to the context, ultimately leading to the consideration of intelligence as a general but diverse capacity of an individual reason to organize experience.

Kant’s concept of reason as a human faculty made human beings proactive, at least in comparison to the reactive entities of empiricists, which were shaped exclusively by sensory-mediated experience, i.e., external influence, and it was sufficiently general in his time, but it is not anymore when we want to consider forms of non-human intelligence. The step beyond intelligence as a human characteristic, considered surprisingly early, quite soon after Kant’s revolution, under the influence of Charles Darwin’s revolution of biological evolution, was originally a product of fictional literature in the 1872 novel *Erewhon or Over the Range* written by Samuel Butler [12]. Butler conceived a potential evolution of machines to the level of intelligence and consciousness in analogy to the evolution of life. What was literary fiction in the 19th century has become a realistic possibility in the cellular automaton environment design of John von Neumann’s universal constructor [13], when considered in the context of Alan Turing’s view of machine intelligence [14].

In the earlier sections, the guardrails for the study of intelligence in its relationship to intelligence and complexity have been focused on epistemological concerns. When intelligence is generalized beyond its context of human faculty or capacity, it is necessary to consider its ontological status. If we agree that intelligence can have multiple forms or manifestations and that this multiplicity involves the multiplicity of realizations beyond the human context, then it is necessary to clarify its abstract status as what is common to all its instances. Intelligence is a noun derived from the adjective “intelligent.”

Then it is necessary to ask first which entities can be (not necessarily are) qualified as intelligent. The answer to this question has important consequences for further inquiry. Moreover, in common-sense use of the term intelligence, particularly artificial intelligence, there are frequent problematic formulations suggesting that intelligence (not an intelligent entity) possesses agency [15]. We can see in mass media claims that artificial intelligence drives cars or solves scientific problems. These are obvious examples of terminological

abuse with serious consequences promoting misunderstanding and confusion. Everyone would object to the expressions that human intelligence drives a car, not that a human being is doing it, or that it has solved a problem, not that scientists have solved it, but still, the expression with the word “artificial” replacing “human” is commonly tolerated or even considered natural and meaningful.

In the literature on intelligence, particularly artificial intelligence, there are frequent attempts to avoid such confusion by the use of the concept of an autonomous agent—an entity possessing agency, i.e., the capacity to act, with an internal locus of control. It may seem natural to assume that autonomous agency is a necessary or minimal qualification for being intelligent.

First, we may examine the need for the qualification of agency with the adjective “autonomous”. Is non-autonomous agency possible? Or rather, is it an oxymoron? It is possible to consider the degrees of autonomy for an agent, but this would go beyond our objectives to characterize intelligence. For this reason, in the following, the focus will be on agency understood as autonomous, i.e., with an internal locus of control, without any additional qualification.

However, the concept of agency is not obvious, whether it is qualified by autonomy or not, and its definition has important consequences for the study of intelligence. We can think about diverse forms of agency, understood as a capacity to act, i.e., to make changes in the state of the broadly understood environment, or as a capacity of an actor, or both, with the action initiated and controlled internally. Not every change can be considered a change of state; the difference has commonly been recognized only in the modern era with the development of mechanics (uniform motion of an object along a straight line is a change of position but not a change of its state). Observed change, or alternatively, invariance, may be relative to the observer and not related at all to the object.

2.4. Necessary Conditions for Being Intelligent—Causality, Interaction, and Agency

There was an extensive philosophical discussion of the question of whether change is a local or global phenomenon (nothing changes, some entities are changing while others are invariant, or everything is changing). Even the possibility of any global change was questioned; change may be just an illusion. Moreover, the very idea of change is associated with time.

For a long time, it was considered obvious that change requires some time interval or time difference, i.e., that instantaneous change was impossible. After all, time was most frequently (although not always) understood as a global expression of local changes, all of which, possibly being very fast like a thunderbolt, seemed to require the passing of time. Modern physics made this association questionable, or at least oversimplified. Quantum mechanics provides claims that a change, even a change of state, can be instantaneous (e.g., a change of the state of an electron in the emission of light from an atom, collapse of coherence, etc.)

Special relativity theory eliminated the separation of time from space and, through this, the global character of temporality. Even more dramatic consequences of the progress of physics concern the status of the concept of action, which is frequently associated with causality. The concepts of causality, particularly the cause-and-effect relationship and the directional character of action, have become even more problematic than in their past turbulent history.

Aristotelian and scholastic ideas of the four causes were logical, i.e., the causes of entities or their states characterized the relationships with other ideas, not necessarily entities. Their original traditional exemplification was given by Aristotle in *Metaphysics* (1013^a24–1014^a25) [16], using the creation of a statue by a sculptor in terms of the four

causes, in terms of his agency, as an example, but they were also used in the discussions of the relationships between concepts such as substance, matter, and form, where there was no explicit actor involved. Most likely, the choice made by Aristotle to use an example involving a human agent (sculptor) and his action was pedagogical, as an explanation involving everyday experience was easier to understand than the relationships between very abstract concepts of matter and form, for which the example was a simple explanation. All four causes characterized one entity, in this example, the sculpture.

It is usually claimed that Francis Bacon, in *Novum Organum* [17], eliminated three of the causes and focused exclusively on the effective cause. Actually, Bacon introduced an entirely new concept of the cause-and-effect relationship between events, which served as the subject of inductive reasoning. Inductive reasoning was a method to discover cause-and-effect relationships within reality, independent of human inquiry, not within the realm of concepts framing it.

The problem remained that the use of probability in inductive reasoning made the claim of the existence of a causal relationship questionable. David Hume, in his fundamental *Treatise of Human Nature* [18], objected to both inductive reasoning and causality; the latter was degraded to nothing more than a useful human habit in reasoning. In another turn of the history of causality, Immanuel Kant, in the *Critique of Pure Reason* [11], raised its status to the synthetic a priori judgment. Causality has become more than a reasoning habit, although not an element of reality, but a condition for reality's comprehension, similar to the rules of Euclidean geometry or temporal relations.

However, in a few decades, with the introduction of non-Euclidean geometries, Kant's qualification of geometrical knowledge as synthetic a priori lost its meaning and the interest in other metaphysical concepts reversed to open hostility, especially among scientists. Still, the role of causality as the objective of inductive reasoning was present in interpretations of empirical research. John Stuart Mill sought a system of logic for inductive reasoning, understood as a tool for the discovery of causal relations, analogous to deductive logic of inference [19]. However, his rules, called today Mill's methods or canons, although they have been absorbed into scientific empirical methodology, have never acquired the status of a complete logical system. They functioned more like practical directives for the interpretation of laboratory results, and causality, too, was involved only in interpretations of scientific formalisms. More recently, instead of searching for a logical system to describe causality, the interest of logicians, the most prominent among them being Rudolf Carnap, shifted to inquiries in inductive logic (in contrast to deductive logic) as a logic of confirmation of empirical research, with the focus on probability instead of causality.

This fact was already noticed by Bertrand Russell in 1917: "All philosophers, of every school, imagine that causation is one of the fundamental axioms or postulates of science, yet, oddly enough, in advanced sciences such as gravitational astronomy, the word 'cause' never occurs..." [20]. Scientists in theoretical disciplines refer to causality only in the interpretation of theories, but not within these theories. The actual concept present, and fundamental, in scientific theories has been and continues to be, even now, that of interaction.

Causality is just an expression of the presence of interactions that are responsible for some specific course of events involving two sides and dynamically altering the states of both, which, with the attention of an observer restricted to one side, and with the selected direction of time, manifests as a causal relationship. This "interactivist" view of the irrelevance of causality in physical theories is criticized for being based on physics or the physical sciences. However, the same can be found in non-physical sciences. When we say that bacteria cause a disease, we actually mean that bacteria interact with the cells of the organism, and as a result of these interactions, the entire organism changes its state into

what is called a disease. It is this crossing of the levels of hierarchical complex systems (e.g., from bacteria and cells of an organism at one level to the entire organism at a higher level) that makes the term “causality” convenient, but which also makes the concept of causality spurious in the study of interactions explaining phenomena. Interestingly, it is the idea of causality between different levels of hierarchically organized natural systems (e.g., cell, organ, organism, and population) that has been criticized, not the idea of causality at the same level, which is always secondary to interaction.

A somewhat similar situation (although not directly related to causality but to time irreversibility) arose when Loschmidt’s paradox was used as an argument against Boltzmann’s interpretation of the second law of thermodynamics [21]. The paradox was formulated in terms of the inconsistency between time-reversible mechanics and time-irreversible thermodynamics, arising when we have two levels of description in terms of microstates and macrostates. The second law states that entropy, characterizing the degree of organization of an isolated thermodynamic system, cannot decrease. This excluded the possibility of the reversal of time, as the direction of time had to be consistent with the increase in entropy. On the other hand, in classical mechanics, the description of all phenomena could be reversed in time. Both courses of action were possible. The inconsistency seemed paradoxical if it was assumed that the two theories describe the same reality but at different scales. The paradox has been eliminated by giving the second law a statistical character. The probability of the transition of a complex thermodynamical system to a state of lower entropy decreases rapidly with the number of components, becoming negligible even for a gas of a teaspoon’s volume.

The problem with using the same solution in the case of causality lies in the fact that the statistical interpretation of causality sought through inductive reasoning would destroy its primary feature of being a necessary, objectively present in reality relationship independent of inquiry. Interestingly, the same conclusion against causality as an objective relationship established through inductive reasoning was used by David Hume [18].

The issue is not just of historical interest. From the present-day perspective, the claim of the irreversibility of time is equally problematic, and with it, that of the irreversibility of cause and effect. Of course, the acceptance of the reversibility of cause and effect is bizarre, so the irreversibility of cause and effect requires the assumption of the irreversibility of time. The time-reversibility of mechanics (called in physics *T* symmetry) comes with the third principle of Newtonian mechanics, which eliminates one-way mechanical action of a force, and the presence of a force is a necessary requirement for a change of state (or they are just two different ways to express the same phenomenon). In the first principle (already considered by Galileo), we have a separation of the concept of change (e.g., motion) and change of state. In an inertial reference frame, i.e., a reference frame from which it is impossible to eliminate any observed forces by transition to another reference frame moving with respect to the original frame, in the absence of any forces or when present forces mutually cancel themselves, an object can move in space (change position in an inertial reference frame) with a constant velocity along a straight line without changing its state.

The second principle adds to the picture a rarely invoked assumption that the mass of an object does not change, i.e., that it is conserved. Otherwise, the relation between force and acceleration becomes indeterminate. Within this framework, particularly with the postulation of the third principle, the dual existence of the force of reaction coexisting with the force of action, the choice of which is which, the action or reaction, is arbitrary, and only interaction has an ontological status. Moreover, it entails that the direction of time change in all mechanical phenomena can be considered arbitrary. The only restriction is that the choice of force (action or reaction) is consistent with the choice of the direction of

time (required by the second principle). However, we observe only acceleration, so if we change the direction of time, this changes the acceleration to the opposite, and the force of action changes to the force of reaction. Here, just as a reminder to prevent an objection to the claim that we observe only acceleration—Hooke's force is just a collective expression of electric forces. Thus, the choice of the direction of time is as arbitrary as the choice of whether we call the charge of an electron positive or negative, and which of a pair of always coexisting forces is the action and which is the reaction.

Further development, enriching physics with electrodynamics involving interactions between electric and magnetic fields, required more general symmetry (PT symmetry) involving not only time inversions (T), but also parity ("handedness") reversal (P). Finally, the discovery of nuclear forces (viz., the weak interaction breaking P symmetry) and the concept of antimatter led to even more general symmetry involving, additionally, the reversal of charge (CPT symmetry). The entire development of physics was convoluted with the increase in the role of symmetry, including CPT symmetry.

The experience of making mechanics and thermodynamics compatible tells us that action cannot be ascribed to simple systems, which in their dynamics must be symmetric with respect to CPT reversals, but each, P, C, and T, can be reversed, under the condition that this reversal is paired with one of the two others. However, this symmetry can be broken in the case of highly complex systems, but the symmetry violation is highly context-dependent and may have only a statistical character; in principle, the symmetry, for instance, T symmetry, can be restored if a sufficiently long time is considered.

Also, we must be aware that the illusory violation of symmetry in interactions, which is quite common in folklore models of reality, may be a consequence of confusion. For instance, most people would claim that the Earth is revolving around the Sun and that the gravitational force of the Sun is causing this revolution. Is not this what the Copernican revolution was about? Actually, it is not true, and the symmetry of the interaction is preserved. Both the Sun and the Earth are revolving around their common center of mass. Because the ratio of masses of the two bodies is so large (one third of a million), this center is close enough to the center of the Sun to be of negligible interest when we deal with our daily matters, because the trajectory of Earth is a large ellipse, and this motion has a great impact on our lives, when the Sun is just wobbling, and its motion is barely detectable by astronomers. However, the presence of such wobbling is the most effective method in discovering exoplanets in other planetary systems, which proves the error of the heliocentric model breaking the symmetry.

Finally, there are some consequences of the symmetry of interactions of great importance that escape popular attention because they do not influence our daily lives. For instance, most people believe in the unlimited precision of the scientific predictions regarding the motion of the objects accessible to our perception. Actually, this apparent high precision is surprisingly limited because of the complexity induced by the mutual character of interactions. Newtonian mechanics allows us to find exact solutions of equations for two interacting bodies, but not for three or more. All we can do to describe the dynamics of many-body systems is to find approximate solutions. If only one or two components of a multi-component system have large masses, the approximation may give us practical solutions to predict the configuration in the not-too-distant future. Even for just three bodies of comparable mass, long-term exact predictions are not possible. If the masses are identical, for some configurations, no predictions are possible at all. For instance, we do not know whether our Solar System is stable or will scatter at some time in the future, fortunately distant at the scale of human life.

For the objectives of this paper, it is important that, at the level of mechanical phenomena, and through the fundamental role of mechanics for all other physical theories, we can

only consider interactions, not actions in physics, unless there is justification for broken symmetry. However, this does not eliminate the concept of agency, only agency defined through causality. We can reformulate the agency not as exercising the capacity to cause changes, but as a capacity to engage in internally controlled interactions that alter the state of an agent and the agent's environment.

The story of the search for a general theory of the phenomena in physics that we intuitively perceive as causality has not finished, even if, as Russell observed, it remains in physics only in interpretations of physical theories as an abbreviated expression of the presence of interactions. The reason for the persistent presence of the term "causality" is a characteristic of interactions that has puzzled physicists since the time of Isaac Newton. How can interactions (such as gravitational ones) occur between spatially separated objects? Rene Descartes tried to escape the mystery by developing, in 1644, a mechanics in which only direct, local interactions on contact are possible and the possibility of the existence of a vacuum is excluded (as something void of properties could not exist) [22]. Newton, who was influenced by Descartes, in his own *Principia* published in 1687 [23], postulated action at a distance in the context of gravitation, but in the second edition admitted that he could not identify the properties of gravitation ("gravity") from phenomena and refused to explain how it is possible with his famous exclamatory statement "hypotheses non fingo" (I frame no hypotheses) [24].

The mystery of action at a distance became even deeper when it was empirically confirmed that a vacuum can exist and that interactions may be between spatially separated objects across the void space. Additionally, in the formalism of Newtonian mechanics, interactions at a distance were instantaneous. The only hope for sanity in the natural world was the discovery made by the astronomer Ole Rømer in 1676 that light travels at the finite speed of about 226,663 km per second, fast enough to seem infinite, but slow enough to have its finite value detected at astronomical distances. This was not sufficient to save the intuitive vision of reality in physics. When 19th-century physics justified the claim that light is a wave, there was a question about the medium for it, which was called the ether. It was expected to have strange properties, but even more strange were the negative results of the attempts to detect its presence.

The revolution came with the recognition of the issue of the lack of meaning of simultaneity, recognized by Henri Poincaré and independently by Albert Einstein. The solution, called Special Relativity, was to eliminate the separation of spatial and temporal relations. Instead of considering how the position of an object changes as a function of time (a curve in three-dimensional space parametrized by a one-dimensional variable of universal time), motion becomes a matter of the choice of a curve in four-dimensional space-time restricted by the condition expressed as the maximum speed equal to the speed of light. The universal, global time of Newtonian mechanics became distributed in local observer-dependent variables. With each point of four-dimensional space-time, there was associated a three-dimensional cone (called the light cone) on which lines represented possible trajectories of light. All trajectories of objects equipped with mass had to be in the interior of light cones, and all interactions (electromagnetic, i.e., light or other) had to be in or on the cones. With the introduction of Special Relativity, separate spatial and temporal relationships were merged into relationships between points of space-time, identified as events.

Within this model of reality as described by space-time, its dynamics is very different from the model based on Newtonian mechanics. There are no changes (understood as continuous transformations in universal time), but differences, and these differences between events have meaning only if they are invariants of the specific group of transformations of space-time (the Lorentz group). In hindsight, we can identify in Newtonian mechanics

an analogical group of transformations (called the Galilean group) acting separately on space and time. This requires the separation of space and time, and making time intervals and space intervals the invariants of the group transformations. As a consequence, in the classical model of reality, there were spatial differences, temporal differences, and dynamic changes linking spatial differences with temporal ones. These distinctions lost their meaning in Special Relativity.

Instead of the study of the dynamics of changes of the states of objects, in Special Relativity, it is more natural to study geometric, topological, or algebraic properties of space-time. The early study from 1964, referring to causality, “Causality Implies the Lorentz Group,” by Christopher Zeeman [25], had as its objective the study of the partial order (poset) of space-time introduced by the quadratic form defining the structure of the Minkowski space of space-time, which essentially orders events in space-time according to the conditions in which events CAN cause which ones, not those in which they DO. There was nothing there about causation, it was exclusively about the exclusion of interactions between events separated by light cones, i.e., alternatively, the requirement that one event is inside the light cone of the other, plus making that inside greater in the order (“later”). The main result of the paper did not add anything to the study of causality, but it was very interesting and surprising that relativity can be simply introduced in space-time by defining an algebraic structure of a poset of events, and the Lorentz group emerges as its invariance group. In 1977, Wojciech Cegla and Arkadiusz Jadczyk identified a related but richer algebraic structure of the partial order type in the Minkowski space of space-time [26,27]. In this case, too, this structure does not describe causality but rather a logical structure of the relativistic space-time analogical to quantum logics [28–31]. Here, too, like in Zeeman’s work, the structure of an ortholattice on the subsets of space-time, generalizing the concept of a Boolean algebra of propositional logic, was derived not from the causal relation, but from orthogonality, formalizing causal independence, i.e., impossibility of interactions.

The most striking result regarding the ortholattice describing the logical constraints on interactions in space-time was that the lattice is orthomodular, i.e., has the same fundamental property possessed by quantum logics. Antonio Zecca proved that these confusingly called causal logics have some properties different from quantum logics. The most important result was that causal logics are atomic but not atomistic and do not have the exchange property [32]. This means that, unlike quantum logics, they can be reducible to direct products and therefore may not admit unrestricted superposition of states. However, the similarities of the logic type structures in quantum mechanics and relativity, even in the presence of some differences, are interesting as a bridge between the theories [32,33].

Summarizing, causal logics do not describe causality but the logical constraints on interactions within space-time, understood not as dynamic changes but rather static differences between objects understood as space-time trajectories. This is not surprising when we realize that these logical constraints come from the symmetric condition of separation between events by light cones. The symmetry of the condition of separation defining the logical structure can only produce a symmetric structure that describes the condition for interaction, not causation.

2.5. Intelligence, Agency, and Free Will

This is not the end of conceptual problems in the description of agency. We could eliminate causality as its defining concept and instead use the capacity for engagement in internally controlled interactions. But this is not a sufficient condition for agency. The concept of agency is close to the minefield of the studies of free will, determinism, and mind–body duality. There are cases in which it is easy to exclude agency. A dropped stone is falling (“doing” something) in free fall, its state is changing, but this fall is strictly

determined by external conditions (the presence of a gravitational interaction), and it is easy to exclude its agency. This does not mean that external determination provides clear criteria to exclude agency. Human beings are determined in many ways in their behavior. Culture can be defined as a system of non-physiological or non-biological determinants of human behavior acquired through communication within a human collective (society, ethnic group, family, etc.). At the individual level, our behavior is determined by our personality traits or habits.

In both cases, determination is not (or may not be) absolute, but rather a matter of degree, and our actions, which usually follow such determinants, can be controlled by our will and may be altered. However, this does not exclude the possibility that this control is illusory, that we do not follow the cultural determinants because some other stronger determinants independent of our control force us to do so. Even if the strong form of complete determinism is accepted, compatibilists claim that free will can be exercised as it is unconstrained by determinants. Our behavior, or our choice of behavior, we may believe, can be altered, for instance, by reasoning, and our actions may follow a different path or alternatively can be forced to follow the constraints of determination, but against our will. Thus, according to the compatibilist view, even when fully determined in our actions, we still can have free will, though this can be suppressed externally so as not to be manifested.

The dominating belief about free will continues to be objected to by many psychologists and philosophers, for instance, Robert Sapolsky in his celebrated book *Determined: A Science of Life Without Free Will* [34]. Thus, at least at present, free will is a matter of personal belief, free choice of philosophical position, not empirically or theoretically proven, and it is still a subject of open inquiry. The key point is the distinction between the choice in intention vs. the actual realization of the choice of action that can be inconsistent with intention. Interestingly, this leads to the so-called Sapolsky Paradox [34]. If we really have a free choice about our philosophical position on the determination of choice, then this contradicts determination in its strict incompatibilistic form, allowing us to exercise free choice. Of course, there is nothing paradoxical in the belief in free will.

More relevant than a speculative general reflection on human free will were failures of empirical attempts to demonstrate the illusory character of autonomy in making decisions initiated by highly publicized experiments carried out by Benedict Liebet [35]. Liebet and his followers claimed that they could detect the preparation in the brain for actions before the conscious decision to perform them [35–37]. More recent research invalidated these claims by showing that the earlier experiments did not consider the differences between deliberate and arbitrary choices in making decisions [38]. However, the fact that the earlier claims were not justified does not entail support for the opposite claims of the capacity for complete autonomy in making decisions.

Another direction of the critical analysis of the meaning of free choice is in the studies contrasting physical determination with randomness. An example is the so-called Free Will Theorem of John H. Conway and Simon B. Kochen [39,40], which states that if observers (e.g., human observers) have free will in the sense that their choices are not determined in the past, then, under specific assumptions drawn from quantum mechanics and relativity, so must elementary particles equipped with spin. That is, human experimenters' free will, understood as the ability to make choices not entirely determined by prior events, entails that some elementary particles must also exhibit a corresponding form of indeterminacy.

The theorem, although supported by a formal proof, with the assumptions weaker than those in physics, suggesting universality, does not negate the existence of free will in human observers. It only demonstrates yet another exotic consequence of the inconsistency between the understanding of macro- and microlevels of nonlocality. If we accept that human measurement determines the choice of properties of quantum mechanical objects

from the superposition of multiple states, each describing a different property, producing indeterminacy of properties in the absence of measurement, then it is a natural consequence that the free choice by a human observer of one of several measurements will be reflected in the indeterminacy of properties. This does not entail that the quantum object (elementary particle) has free will, and the argument against the free will of observers based on the claim that the free will of particles seems odd loses its validity.

2.6. Agency and Information

This overview of issues involved in the study of the status of human free will, a very restricted part of the study of agency, shows that the involvement of the question about human free will does not help us in our search for the necessary conditions for intelligence. Still, it is not clear whether the possession of agency, and the type of agency, is a necessary condition for intelligence. The tragic example of one of the greatest physicists of modern times, Stephen Hawking, who for decades was physically almost completely incapacitated by disease but continued his extraordinary contributions to science that required a very high level of intelligence, demonstrates that agency, as a condition for intelligence, has to be qualified. What type of agency is necessary for intelligence? There were many attempts to answer this question in the context of human intelligence. For instance, the following examples are rather randomly chosen, with hundreds, if not thousands, of similar attempts having been made with similar results. In the early discussion of the meaning of intelligence in 1921, J. Peterson proposed that “[Intelligence is] a biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior” [41]. Seventy years later, the research group chaired by Ulric Neisser proposed the answer to the same question that the most fundamental characterization of intelligence requires four abilities: “Intelligence is an ability to understand complex ideas, to adapt to the environment, to learn from experience, and to engage in reasoning to overcome obstacles” [42].

The answer was intended for the context of human intelligence and, consequently, is excessively context-dependent if we want to consider a more general concept of intelligence. For this purpose, we can look at what is common to all four abilities distinguished by Neisser’s research group and Peterson’s characterization. Each of the four is related to information, its use, or processing. Understanding complex ideas is about the ability to identify the meaning of information, including that which is complex (obviously, someone who understands complex ideas would understand simple ones). Both adaptation to the environment and learning from experience are about the reception of information from the environment, the identification of its meaning, and its use for controlling one’s own action. Finally, engaging in reasoning to overcome obstacles, the most human-related ability out of the four, is similar to the first one. Engaging in reasoning is a mode of overcoming the complexity of acquired information, and overcoming obstacles is a human view of resolving complexity. Thus, we can generalize the concept of intelligence proposed by Neisser’s team as the ability to acquire information, to reduce its complexity, and to identify its meaning. Not very different is Peterson’s view of intelligence as a mechanism by which the effects of complex stimuli are brought together and given a somewhat unified effect in behavior. Stimuli are a form of information, and bringing them together to affect behavior is information processing, simplifying complexity.

If, as we can see in the examples above, the key concept for the conceptualization of intelligence is information, there is a legitimate question about the involvement of agency. Do we need agency as a prerequisite for intelligence? No, if agency is associated with physical action; yes, if agency is understood as the ability to operate on or to process information. Of course, the development and implementation of the latter type of agency

(processing information) may require physical processes and physical actions. Information is always encoded in a physical medium [43–46], so the processing of information always involves physical processes, e.g., the dynamics of the states of the medium of information. However, the qualification of a system as intelligent can be expressed exclusively in terms of information, as it is the physical implementation of intelligence, not intelligence itself, that is determined by the physical implementation of information. For instance, judging the low intelligence of apes because they cannot produce human language would have been an obvious error.

2.7. *Intelligence as a Characteristic of Entities*

One of the frequent confusions in the use of the adjective “intelligent” is a result of its use to qualify the results of actions of intelligent entities without recognition of the distinction between an intelligent agent and the product of such an agent’s action. This is a consequence of the fact that intelligence has a normative function in the assessment of behavior. The action of an intelligent agent, in this normative sense, is intelligent, and its products are considered intelligent. Alternatively, something good, complex, and not easy to produce must have been produced by an intelligent agent.

This obviously faulty reasoning is based on the false assumption that what is random and chaotic is simple, while it is exactly the opposite. One of the greatest technological challenges of the present time is to implement randomness in computers. Thus, the arguments from Creationists for the necessary existence of God, according to which someone had to create the world that “works” in such complex ways, are faulty. Against their intention, this feature of the world could be equally well, and sometimes is, used as an argument for God’s nonexistence. Since the world “works” perfectly according to the relatively simple rules accessible to human reason, God’s existence is not necessary.

This was the reason for the revolutionary effect of Charles Darwin’s work on biological evolution. His theory was not only so rational that evolution seemed intelligent, but at the same time, it showed that this intelligence comes from a simple process that did not require any divine intervention. Obviously, the claim that biological evolution entails the non-existence of God is similarly faulty, as we can only conclude that the existence of highly complex life, or more specifically, of human intelligence, does not require the existence of God but does not exclude it.

Thus, the use of the adjective “intelligent” to qualify the results of actions of intelligent entities, frequent in the vernacular, should be avoided or considered just an abbreviation, unless its context identifies the intelligent entity that produced that result. The use of the adjective “intelligent”, when it qualifies the result of an action without reference to the agent or action, does not have any general meaning.

There is some indirect analogy to the problem with the use of the adjective “effective,” for instance, in the context of the conceptualization of intelligence. To qualify an action as effective, we have to refer to the outcome and to the means available to the agent, or more exactly, the circumstances of the action. Efficiency is not an innate characteristic of an agent but of the relation between the means and the outcome. It is strongly context-dependent. What is efficient in one context may be inefficient in another.

Yet another analogical confusion is in the use of the noun “resources”. There is no general definition of the concept of resources. It must be relativized to the action and its outcome. The choice of what is a resource and what is not for an action depends on the choice of the outcome but also on the choice of the realization of the action. In some ways, to implement an action, something can be a resource, for another it could be a neutral but not important condition, and for yet another it could be an obstacle. Moreover, the terms “efficient” and “resource” have some positive, normative connotations that can

be confusing; for instance, a sentence like “The oil fields are rich resources for effective pollution of the environment” seems bizarre, unless the misuse of the terms in this context is intentional to express sarcasm.

This leaves us with a still open question about what or who can be intelligent. It has been recognized in all human civilizations in their early stages that humans differ in their mental abilities, and these abilities were important in assigning the roles held by individuals in their collectives. This made it necessary to develop criteria for the assessment of such abilities and to establish scales of intelligence qualification.

The criteria and scales of individual abilities have always been culture- and context-dependent, typically established by the observation of the effectiveness of some actions in achieving culturally determined and distinguished goals with adherence to cultural norms. As such, they do not have much value for our conceptualization of intelligence transcending human contexts. They have more value for the study of cultural relativization of different manifestations of intelligence than for searching for synthesis. We must remember the diversity of mental abilities that should be distinguished. What in the past or in different cultures was considered uniform mental agility is at present diversely conceptualized; the concept of intelligence has become distinct from wisdom, knowledge, education, talent, etc. It has a more fundamental character as a condition for acquiring the other abilities, or, in reverse, it describes the potential for the other abilities that can be its realization in social, collective reality.

There is one more use of the adjective “intelligent” that requires scrutiny in expressions of the type “intelligent society” or “intelligent collective” corresponding to the idea of “collective intelligence.” Michael Levin not only defends ascribing intelligence to collectives but gives compelling arguments that all instances of intelligence are intelligent collectives [47]. This view is natural from the perspective of the claim of the conference paper [1], of which the present paper is an extension, that intelligence is the capacity to overcome the complexity of information. In the conference paper [1], there is an argument referring to the Law of Requisite Variety introduced by W. Ross Ashby, that high-level intelligence can be achieved only by systems of high-level complexity. Thus, there is no problem with considering collective intelligence, except that in such a context, the adjective “collective” is redundant and may be confusing. The dangers of such confusion are parallel to the dangers in the use of the expressions “collective action” or “collective agency” discussed extensively in the literature, in the context of human collectives, in which additional complexity arises with their social and cultural organization and the involvement of intentionality at the individual and collective levels [48].

The danger in using a separate term, “collective intelligence,” lies in the possible but false interpretation of the existence of separate types of intelligence—of individual and collective intelligence. It is possible that in a complex hierarchical system, intelligence can be manifested at different levels of a hierarchy, but this does not mean that the distinction of the levels of manifestation must be reflected in, or give any reason for, the distinction of the levels or types of intelligence. In particular, the manifestation of intelligence at the higher hierarchical level of a complex system does not entail a higher level of intelligence for its components. Similarly, the fact that a high form of intelligence is manifested at the lower level of the hierarchical complex system does not entail that higher intelligence, or intelligence at all, will be manifested at the higher level.

For this reason, the expression “intelligent society,” used in some futuristic designs, is very problematic, even when it is applied to hypothetical visions of a society whose members are very intelligent or a society with advanced technology equipping its members and organizations with artificial intelligence systems of a high level of intelligence. The problem arises from the unjustified extension of the qualifications of the members of the

collective to the entire collective. The only situation in which the qualification of a society as intelligent would make sense is where the society as a whole satisfies the criteria of intelligence, for instance, when in its entirety it is processing information and reducing its complexity, or, in a more general perspective, without the focus on information, where we can claim that the society has agency. This, however, is never considered in the plans for future societies, where the focus is exclusively on technological advances. Moreover, in all existing societies, actions are taken by individual representatives or individuals holding power, not by societies.

Thus, considering the content of the preceding and present sections, in the following, the adjective “intelligent” will be applied to some entities (individual or collective) that can be considered agents (i.e., possessing agency) in the restricted sense of being capable of engaging in interactions through the reception, transformation, and emission of information. This means that intelligence will be considered a qualification of the capacity of an entity to engage in the dynamics of information.

3. Information

3.1. Information—Level of Conceptual Generality

As explained earlier, there are two reasons for using a very general concept of information in this study. One of them is that any restriction of generality should be made cautiously, as it may exclude some instances of intelligence in which intelligent (even highly intelligent) entities engage in the dynamics of information, but where the information is excluded due to a too narrow conceptualization. For instance, any definition of information restricting the concept of information to its manifestation in human languages restricts the concept of intelligence to human beings or, at most, to entities that can engage in processing at least one human language. This is, of course, unacceptable impoverishment. We want to consider intelligence in its much more general manifestations, including the intelligence of animals or other forms of life, systems processing formal, non-human languages; technological systems involved in computation; or even potential extraterrestrial entities capable of intelligence. Thus, we already need a very high level of generality.

The second reason for the need for a high level of generality in the concept of information is the fact that the theory of information must be a metatheory, in the sense introduced by Alfred Tarski [49], of the medium of this study, which is already a language of high level in Tarski’s hierarchy. Tarski’s Theorem on Indefinability of Truth in Semantically Closed Theories requires the distinction between a metalanguage and an object language to formulate, in a richer and more general metalanguage, the semantics for theories formulated in the object language. This leads to a hierarchy of languages, in which the theory of information, a concept more general than linguistic information, should play the role of metatheory for the theory of language. This can be achieved in a manner analogous to the relationship between mathematics and metamathematics [6,49–53].

In this study, the point of departure is a theory combining the predicate calculus of first-order logic (understood as a theory of a restricted, linguistic form of information) and the theory of sets in its Zermelo–Fraenkel axiomatization with the Axiom of Choice, ZFC [54]. Of course, we could not use only ZFC as a departure point because ZFC requires predicate calculus (PC). This paper does not require a high level of formalization, so this information is provided only to justify the use of concepts and rules of logic and the use of concepts such as set, belonging to a set, etc. We will need a few algebraic concepts (e.g., a Boolean algebra), but they can be defined in PC + ZFC. On the other hand, we will assume that our understanding of a language is quite restricted to the simplest possible form. We do not assume much more than logical or basic set-theoretical terminology, so the starting point is in a relatively simple, well-known theory. It would be much more

difficult to include any form of modal logic or any form of theoretical system describing natural languages. Thus, in the formulation of our overarching theory (or rather theories) of generalized information, we will have to make sure that it is consistent with PC and ZFC and nothing more.

3.2. One–Many Opposition

There will be only one categorial (undefined) concept, “one–many opposition”, used here mainly to provide a philosophical background. It will be replaced by its set-theoretical interpretation in the ramifications into diverse forms of information in different information systems. The one–many opposition is probably one of the oldest and most fundamental ideas in the philosophical traditions of all major civilizations, but its formulations and understandings are more diversified than those of any other ideas. There has probably been no philosopher in the entire European tradition, from Parmenides [55] or Tales [55] to Plato [56,57] and Aristotle [16], Kant [11], William James [58–60], or currently active philosophers, who has not addressed at least some of its aspects. One–many opposition is also fundamental in the tradition of Chinese philosophy [61], and it is present in the philosophy of the Ancient Middle East [62]. Not always, but very frequently, the opposition is expressed explicitly as one–many.

In the Vedantic tradition of Indian philosophy, the issue of the one–many relation has, for instance, an expression in the discussion of the unity of *brahman* and *atman*. Both concepts are the results of the search for unity in the varieties of phenomenological (external) and psychological (internal) experiences of human intellect. The orthodox Vedantic position makes the ultimate step towards absolute unity of only apparent variety in the form of identity of *brahman* and *atman*. The discussion has continued, leading to deviations from orthodoxy to some forms of duality. The relationship between the phenomenological multiplicity of the world and the unity of the intellect was identified with knowledge or information (these concepts have not been differentiated clearly).

In Ancient Greek tradition, pre-Socratic philosophy was preoccupied with overcoming the variety associated with ignorance in the search for unity associated with wisdom. In a moderate form, the resolution has been achieved through the reduction of the myriad of phenomenological experiences to the few basic elements, which, when mixed in appropriate proportions, could produce a variety of objects of experience.

In the more radical views of Parmenides, popularized by his disciple Zeno of Elea, the variety has been given the status of illusion. True reality, or rather “being,” was identified with absolute unity [55].

For Plato, the one–many relationship was important enough to be the main subject of a separate dialogue, *Philebus* [56], in which Socrates presents and surprisingly does not oppose the position of Parmenides: “We say that the one and many become identified by thought, and that now, as in time past, they run about together, in and out of every word which is uttered, and that this union of them will never cease, and is not now beginning, but is, as I believe, an everlasting quality of thought itself, which never grows old.” [56]. The Eleatic tradition and the explicit presence of the one–many topic can also be found in another of Plato’s dialogs, *Parmenides*, in which Socrates talks to Parmenides about his views and the views of Zeno: “For you, in your poems, say the all is one, and of this you adduce excellent proofs; and he on the other hand says that there is no many; and on behalf of this he offers overwhelming evidence. You affirm unity, he denies plurality.” [56].

Aristotle, whose influence on medieval philosophy was immense, can be considered ultimately responsible for the word “information” that was introduced in Scholastic philosophy. In his twofold characterization of a substance as matter and form, it is the form that is responsible for our comprehension of reality [16]. Although form itself cannot exist

separately without being carried by the indefinite, unknowable matter that gives the object realization of its identity (object's materialization in the transition from potential to actual existence), form's priority in our acquiring knowledge about the object is clear. To know about the object, its form must be transmitted to give form to our thought, and our mind becomes "informed." This medieval etymology of the word "information" was very likely responsible for the persistent confusion of the concepts of information and knowledge. Also, it was the reason for the attempts to identify information with form [63].

The classical period of Greek philosophy and medieval Scholastic philosophy made so many different contributions to the discussion of the one-many relation that going further in the search for typical views would be pointless. Instead, it is more relevant, especially in the context of looking at the foundations of the definition of information, to consider the role of the one-many opposition in the early development of set theory.

Today, the question of in what sense we can say that a set exists, particularly an infinite set, is rarely asked. But at the end of the 19th and the beginning of the 20th century, it was a central theme of philosophical discussion. The discussion began with the view expressed by George Cantor, whose work set the firm foundation for set theory as a separate mathematical theory, on understanding a set as "[...] any multiplicity which can be thought as one, i.e., any aggregate of determinate elements which can be united into a whole by some law" [64]. Strong criticism from the side of mathematicians such as Leopold Kronecker, who believed that for the existence of a set it is necessary to establish a definite construction in a finite number of steps which incorporates elements one by one into a whole, brought the issue of how to understand the unification of the variety of elements into unity to the attention of mathematicians and philosophers alike. Edmund Husserl participated in the discussion, writing from the philosophical perspective: "Any talk of sets or multiplicities necessarily involves the combination of the individual elements into a whole, a unity containing the individual objects as parts. And though the combination involved may be very loose, there is a particular sort of unification there which would also have to have been noticed as such since the concept of set could never have arisen otherwise [...]. If our view is correct, the concept of set arises through reflection on the particular [...] way in which the contents are unified together [...] in a way analogous to the manner in which the concept of any other kind of whole arises through reflection upon the mode of combination peculiar to it" [65].

According to historians studying his intellectual biography, Husserl became convinced that the process of integration is of a psychological character: "The concept of collection in Brentano's sense, Husserl explained, was to arise through reflection on the concept of collecting. Sets, he thus reasoned, arose out of collective combination, in being conceived as one. This combining process involved when objects are brought together to make a whole only consists in that one thinks of them 'together' and was obviously not grounded in the content of the disparate items collected into the set. It could not be physical, so it must be psychological, a unique kind of mental act connecting the contents of a whole" [66].

The issue became more puzzling when Bertrand Russell, in his famous paradox of a collection of sets that are not elements of themselves, showed that the assumption that every definable collection is a set must be false. If the process of integration is purely psychological, why can some collections be integrated into sets, but not others?

An axiomatic approach to set theory removed the problem of paradoxes, but the issue of justification for the infinite selection of elements for sets (the Axiom of Choice) remained a source of discussion until a long period of everyday mathematical practice had made its paradoxical consequences (e.g., Banach–Tarski Paradox) more acceptable than the poverty of the mathematical toolkit.

Thus, the problem of integration of a variety of elements into the unity of a set has become rather a subject of study for historians of the philosophy of mathematics, being eliminated from the agenda of philosophical discussions, not through a decisive resolution but because of the decline in the interest of both parties, mathematicians and philosophers. The former group lost interest because of the (false) popular belief that once some axiomatic set theory has been provided and its consistency is not challenged, there is no reason to inquire into the status of sets, as it is irrelevant for mathematical practice.

Zermelo–Fraenkel axiomatic set theory with the Axiom of Choice (ZFC), the most popular version of set theory in applications to mathematics, is clearly one of the possible ways to formalize the one–many opposition and to solve the problems that puzzled philosophers for millennia. However, it should be remembered that this solution comes at a price. In ZFC, as in the majority of versions of set theory, there are only three primitive (i.e., undefined or categorial) concepts: the binary predicate of equality “=” inherited from the predicate calculus; the binary ZFC-specific predicate “belongs to,” or, alternatively, “is an element of” (with symbol “ $\in$ ”); and variables representing sets indicated by letters, small or capital.

Typical confusion among those who do not have much interest in set theory is caused by the traditional notation, in which sets on the left side of the symbol “ $\in$ ” are written with small letters and sets on the right side of the symbol “ $\in$ ” with capitals, where those on the left side written with small letters are called “elements”. The term “element” is meaningless in ZFC; there is only a binary predicate “is an element of” relating sets. The only objects in ZFC are sets. The only set that does not have any elements is the empty set, so the position on the side of the symbol cannot serve as a distinction for the qualification of sets.

If we identify many from the one–many opposition with being a set, what is the meaning of one or oneness? In what sense is the empty set many, if the identity of it that follows from the fact (the ZFC theorem) that the empty set exists and is a set is associated with oneness? However, the opposition seems present in the fact that for every set x , $x \neq \{x\}$ is a theorem of ZFC and the set $\{x\}$ has cardinality 1, while the set x can have arbitrary cardinality, e.g., 0 or 2. Another strange consequence of the association of many with being a set is the question about how sets of cardinality 1 (they must exist; for instance, $\{\emptyset\}$) are many. Obviously, oneness in the one–many opposition is unrelated to the cardinal value of 1, or at least, the relation is not simple. We can see here the same ambiguity that puzzled Plato’s character of Socrates in *Parmenides*, mentioned above—the coexistence of one and many.

The one–many opposition cannot be explained purely in terms of the concepts of set theory, such as ZFC, because set theory considers only existence, not identity, and it is identity that is fundamental to the discussions of the one–many relationship. At this point, we must be careful to prevent a terminological confusion arising from the use of the unrelated term “identity” for an equation that has the entire domain of the equation as its solution set. Although the homonymy is not accidental (the sides of the equation can be considered not just equal but identical), the term is used here in a different meaning. More exactly, in ZFC, the idea of identity is not completely absent but reduced to existence and is expressed in terms of the language of the predicate calculus of the first order.

We can find it in ZFC in the Axiom Schema of Reduced Comprehension (sometimes called the Axiom Schema of Separation) [54]: Given any set S , there exists a set A (a subset of S) such that, given any set x , x is an element of A (i.e., $x \in A$) iff x is an element of S and $\varphi(x)$ holds (is true) for this x , where $\varphi(x)$ is a formula of set theory with a free variable x . The axiom schema tells us that to identify a subset of a set S , we need to use a set from which we select elements (i.e., all elements of the collection have to be elements of some

predefined set S to avoid Russell's paradox), and then we can identify the elements of our set by the description of its properties in the language of the set theory.

In this axiom schema, the set-theoretical formula φ plays a central role. Intuitively, it is information that allows us to select elements for the set A , formulated as the property that identifies A . However, it is formulated in the language of the predicate calculus as a statement about the existence of set A produced by the formula $\varphi(x)$, not as the identification of the set A —it does not tell us anything about the properties of A , not even the cardinality, which is the most general property of any set. We start with the formula $\varphi(x)$, and then we claim that there exists a corresponding set A . Here, the connection with the one–many opposition is lost or at least obscured. It is no longer about the status of set A as one or many.

In the approach to conceptualization of information presented in this paper, which had its mathematical formalism introduced in the 2011 paper [67] and developed in a long series of papers [68–73], information was defined as an identification of variety, understood as a resolution of the one–many opposition in terms of transition from many to one, where one is identified with identity. Thus, information gives identity to multiplicity, and this identity can but does not have to be achieved with linguistic means.

An identification of variety can be realized in two main ways. We can select one out of the many or we can bind the many into one whole, typically the latter by equipping it with some structure. The selection of one out of many is a selective manifestation of information; the binding of many into, or with, a structure is its structural information. These are the two manifestations of information, not two distinct types, as they always coexist.

3.3. Information as Identification of Variety and Logic of Information

We want to consider models of information that are independent of the restrictions of linguistic encoding of information, and, additionally, we want to have information characterized relative to the diverse types of its encoding, i.e., diverse information systems, which include, as a special case, the ZFC set theory with the predicate calculus. Moreover, we want to consider multiple levels of information related to complete or partial identification. At the same time, we want to have a theory in which we can interpret at least some more context-dependent theories of information. For this purpose, it is necessary to have a mathematical conceptual framework of very high generality but developed with the use of ZFC set theory.

These conditions justify the use of the theory of closure spaces, which is a direction within general algebra, but which has served as a foundation for the formalisms of most mathematical theories, including logic (metamathematics), topology, geometry, algebra, probability theory, and some theories of mathematical physics. The most important mathematical concepts will be defined or explained here, but the details can be found in the literature, e.g., in the monograph on lattice theory by Garret Birkhoff [74].

The structure of a closure space on a set S has an additional distinctive and unique characteristic. It is autological (as opposed to Weyl's heterology), in the sense that it can be used at a higher level of generality as a closure space on the power set 2^S to describe all closure spaces on S . This follows from the fact that the set of all Moore families on S is a Moore family on 2^S (the set of all subsets of S). Since we are interested in raising the level of generality in the transition from ZFC + PC to a theory of information that has linguistic information as its special case, the autological character is of particular importance [75].

The key concept for the formalism of the theory of closure spaces is a set S with a closure operator $f: 2^S \rightarrow 2^S$, which is a function which satisfies the following conditions: (i) $\forall A \subseteq S: A \subseteq f(A)$; (ii) $\forall A, B \subseteq S: \text{If } A \subseteq B, \text{ then } f(A) \subseteq f(B)$; (iii) $\forall A \subseteq S: f(f(A)) = f(A)$.

Every closure space can be defined in an equivalent (cryptomorphic) way by a Moore family of subsets of S , i.e., a family closed with respect to arbitrary intersections and

including the set S . Every Moore family $\mathcal{M}$ defines a transitive operator: $f(A) = \bigcap \{M \in \mathcal{M} : A \subseteq M\}$, and in turn the family $f\text{-Cl} = \{M \subseteq S : f(M) = M\}$ is a Moore family defining the closure f .

We will consider a closure space $\langle S, f \rangle$ with a closure operator f and its corresponding Moore family $\mathcal{M}$ of closed subsets as an information system. The specific choice of closure space depends on the choice of the type of information system. For instance, we can consider geometric, topological, logical information, etc. The family of closed subsets $\mathcal{M} = f\text{-Cl}$ is equipped by set inclusion with the structure of a complete lattice $\mathcal{L}_f$, which we consider to be the logic of an information system $\langle S, f \rangle$. It plays a role in the generalization of traditionally understood logic in the linguistic context to (not necessarily linguistic) more general information systems, although, in general, it does not have to be a Boolean algebra, as it usually is in linguistic logic, and it may be just a complete lattice. In many cases, it maintains all the fundamental characteristics of traditional logical systems.

The use of the term “information system,” preceding the description of the manifestation of information, is an expression of the recognition of the relative character of the concept of information, which can be manifested in multiple forms, in analogy to the concept of a structure.

Now, we can proceed to the description of the manifestation of information within an information system $\langle S, f \rangle$. An encoding of information is a distinction of a subfamily $\mathfrak{F}$ of $\mathcal{M}$, such that it is closed with respect to the (pairwise) intersection and is dually hereditary, i.e., with each subset A of S belonging to $\mathfrak{F}$, all subsets of S including A belong to $\mathfrak{F}$ (i.e., $\mathfrak{F}$ is a filter in the lattice $\mathcal{L}_f$). The Moore family $\mathcal{M}$ can represent a variety of substructures of a structure of a particular type (e.g., geometric, topological, algebraic, logical, etc.) defined on S . This corresponds to the structural manifestation of information. Filter $\mathfrak{F}$, in turn, in many mathematical theories associated with localization, can be used as a tool for identification, i.e., selection of an element within the family $\mathcal{M}$, and under some conditions within the set S . Thus, filters can be interpreted in both the structural and selective way. The formalism based on filters reflects structural characteristics of information, where filters represent a state of some universe of inquiry (quite frequently simply called “possible worlds”). Certainly, the universe of inquiry should not be confused with the world, as the filters can be defined only after a particular logic of information is chosen, i.e., after the type of information is established.

There is a natural question about the reason for the use of closure spaces as models for information systems and filters in the lattice of closed subsets as models for encoding information. In our experience of the use of the term information, it can be associated with the methods of encoding, storage, and processing. We know about a great variety of distinct methods or ways in which information can be encoded by human intervention or by natural processes. In each case of encoding, we have some method or mechanism of its realization, and this is called here an information system. Closure spaces can be found in virtually all mathematical theories either as an original fundamental concept (e.g., in the case of topology) or as a fundamental concept of a derived, alternative formulation (e.g., in the finite geometries). Of course, nobody can claim that the choice of closure spaces for the description of information systems is the only one possible or even the best, but the reasons for such a choice are many, some of which, more compelling, are presented above.

The second question about the choice of filters of closed subsets as instances of encoded information is more complex. The original reason for this choice was to formulate a theory of information for which we can develop semantics. The most typical view of semantics in the context of traditional linguistic logic is exclusively the association of symbols with their denotations, objects that are represented by these symbols. We could see its reflection in the Axiom Schema of Reduced Comprehension mentioned earlier in this paper: Given

any set S , there exists a set A (a subset of S) such that, given any set x , x is an element of A (i.e., $x \in A$) iff x is an element of S and $\varphi(x)$ holds (is true) for this x , where $\varphi(x)$ is a formula of set theory with a free variable x . The formula $\varphi(x)$ is nothing else but a complex symbol, and set A is its denotation within S .

However, there is a dual logical structure on the sets of formulas based on the ordering of the logical consequence closure operator. In the early form, this dual structure, called connotation, considered by John Stuart Mill [19], was restricted to the set of universals organized by Aristotelian syllogistics, in which the formulas are exclusively of a simpler type, “every A is B ”, “some A is B ”, “no A is B ”, and “some A is not B .” Also, the rules of inference in syllogistics are different. In this case, Mill was not focusing on syllogistics, but on the order of universals. However, Mill’s connotation can be easily extended to predicate logic and ZFC. In the original syllogistic version, the relevance to semantics is obvious due to the use of the ordering of universals into the genus–species order used in classical definitions. Obviously, the classical genus–species definition has as its function establishing the meaning of universals, so connotation is a semantic tool. On the other hand, when we look carefully at the ordering structure of connotation, the definitions are associated with filters [76,77].

Finally, this formalism makes a clear distinction between information and its encoding, which is completely missing in all works stimulated by Claude Shannon’s famous work [78]. Interestingly, this distinction was already present in the paper by R. S. V. Hartley [79], preceding Shannon’s work by two decades. This lack of distinction was the reason why Shannon had to exclude from his study the meaning of information. As was noticed by Hartley, the same information can be encoded in multiple different but equivalent ways. The meaning of information is a class of abstraction of the equivalence of its encodings, in the same way as the meaning of a concept can be defined by different but equivalent definitions. The meaning of information is one of its most important characteristics, unfortunately ignored by Shannon and his followers, although the reason is obvious. The entropy of different encodings of information, even those with the same meaning, may be very different. Therefore, it does not tell us anything about the meaning of encoded information.

3.4. Relationship with Other Conceptualizations of Information

E.C. Shannon’s entropy is prized for providing a quantitative methodology in the study of information transmission [78]. Shannon never defined the concept of the meaning of information in his famous work on the transmission of information [78,80]. The complete, programmatic dissociation from the meaning of information is natural because the meaning of information is irrelevant to the issues of communication studies.

Is the quantitative methodology lost when we model information as a filter in the lattice of closed subsets? At a higher level of generality, especially in infinite information systems, quantitative description becomes more complex and diverse. However, at the level corresponding to Shannon’s finite case (i.e., when the set S is finite with n elements), and when we do not make any restriction on the subsets corresponding to properties, i.e., the closure is trivial ($\forall A \subseteq S: f(A) = A$), the task is quite simple. First, we have to recall that in the finite case every filter is principal, i.e., for every filter $\mathfrak{F}$, there is a subset A of S such that this filter consists of all subsets that include A (i.e., $\mathfrak{F} = \mathfrak{F}_A = \{B \subseteq S: A \subseteq B\}$). Let us assume that A has k elements. Then we can use reasoning similar to the reasoning used by Hartley [79] in establishing his quantitative description of information. Since S has n elements and A has k elements, the least number of elementary units of any type within an auxiliary set used for encoding that is necessary to encode all elements of S and A respectively is the rounding to the nearest integer of $\log_2 n$ and $\log_2 k$; thus, we can consider $H_A = \log_2 n - \log_2 k = \log_2 n/k$ to be a quantitative description of information

represented by $\mathfrak{S}$ when it is understood as the degree of selection. The maximum value of $\log_2 n$ is for A with one element only, and the minimum value of 0 for $A = S$.

We do not use probability here at all; Hartley never referred to probability either. In his paper, he derived his quantitative description of information using purely combinatorial reasoning about different encodings of information, similar to that above. Also, following Hartley, we do not restrict our reasoning to letters, characters, or messages. Moreover, we do not assume here that encoding has to be in the form of ordered sequences of a distinguished set of units (a positional system).

The transition to the probabilistic formulation in the finite case can be facilitated by the interpretation of the above reasoning in terms of characteristic functions of subsets. The subset A is associated with its characteristic function $\chi_A(x)$ on S such that $\chi_A(x) = 1$ when $x \in A$, and $\chi_A(x) = 0$ when $x \notin A$. It is easy to see that $\mathfrak{S}_A$ consists of all subsets B of S such that $\forall x \in S: \chi_A(x) \leq \chi_B(x)$. The characteristic function $\chi_A(x)$ can be interpreted as a random variable X with a Bernoulli distribution with parameter $p = 1$. In the next step, we can consider a more general case with the parameter p belonging to the interval $[0, 1]$. Then, we have that the probability $P(A)$ of A considered as an event is p and for all sets $B \in \mathfrak{S}_A$, the probability $P(B) \geq p$. However, $P(B) \geq p$ does not entail $B \in \mathfrak{S}_A$. The necessary and sufficient condition is that the conditional probability $P(B|A) = 1$.

This leads us to the next step. If we have a family $\mathfrak{S}$ of events for the finite probability space S with a probability measure P defined by $\mathfrak{S} = \{B \subseteq S: P(B) = 1\}$, then $\mathfrak{S}$ is a filter as $p(A \cap B) = 1$ when $p(A) = p(B) = 1$, and $\forall A, B \subseteq S$: If $A \subseteq B$ and $p(A) = 1$, then $p(B) = 1$. This applies to every conditional probability generated by P , too. Then, we have a correspondence linking filters in 2^S and probability measures in the probability space S with probability measure P , which determines conditional probabilities $P(\bullet|A)$ for every A such that $P(A) \neq 0$. Therefore, information in the form of a filter can be defined by the filter $\mathfrak{S}_A = \{B \subseteq S: P(B|A) = 1\}$.

Finally, when we have the very special case of the uniform (classical) probability distribution for P , our magnitude of information considered above, we have the value $H_A = \log_2 n/k = -\log_2 P(A)$. Shannon's entropy $H(X) = -\sum p_i \log_2 p_i$ for the random variable X partitioning S into subsets A with the values $x_1 = -\log_2 p_1, x_2 = -\log_2 p_2, \dots, x_m = -\log_2 p_m$ is simply the expected value for the random variable X with values $x = H_A$, whenever A s are elements of some partition of S .

This was just an example that shows that in the very special and simple case of the trivial closure ($\forall A \subseteq S: f(A) = A$), the corresponding model of information in the form of a filter in the logic of information is not very far from the approach of Shannon's information theory. However, the general case is much more complex. For instance, there are non-prime filters in the infinite case, and the logic of information (the complete lattice of closed subsets) may be non-Boolean.

4. Complexity of Information and Its Reduction

4.1. Complexity vs. Simplicity out of the One–Many Opposition

The study of complexity requires additional caution because of the influence of the vernacular use of the terms “complexity” and “simplicity,” carrying some normative characteristics. In everyday use, particularly in the context of communication, it is a very frequent request to “make things simpler,” but rarely to “make things more complex.” On the other hand, there are complaints about “oversimplification” as an error calling for correction, but not for “complexification.” Additionally, in vernacular usage, simple is associated with what is easy. Not only in English, but in many other languages, “simple” and “easy” are either synonymous or frequently used as synonyms. In the following, there will be an explanation for such a normative understanding of the opposition of complex

vs. simple as an expression of the pragmatic preference in the use of information made popular by the principle of parsimony. The danger of the influence of this normative characteristic is related to the fact that the distinction between what is complex and what is simple is not as obvious as it may seem. Once again, we encounter here a shadow of the one–many opposition.

In the one–many opposition discussed earlier in the context of the definition of information, the focus was on the oneness understood as identity. However, in the millennia-long discussion of the opposition, oneness was sometimes invoked in the meaning one-in-number, following the understanding of the one–many opposition in terms of the set-theoretical cardinality. We could see that this understanding leads to an oversimplification of the one–many opposition, as the same set can be interpreted as many, since its cardinality can be higher or lower than one, or one when we consider it this way, reflected in set theory as a transition from set x to set $\{x\}$, the latter always with cardinality 1.

We considered oneness interpreted as identity in the definition of information as that which makes one out of many (either through selection or by a binding structure). The cardinality view remains present as a characteristic of the result of the transition from many to one. However, surprisingly, this does eliminate the issue of whether one is simple or many.

For instance, the level or degree of complexity of information can be understood as a qualification of information in terms of its quantitative characteristics arising from cardinality (the number or measure of the components of information) and qualitative characteristics (the structure of these components and the degree to which these components are bound together) [81–83].

The qualification of simplicity may depend on which of the two ways of understanding is in focus in a given study. In the algebraic study of finite groups, one of the most successful long-term mathematical research programs on a large scale was the classification of simple finite groups. Simple groups are groups that are not isomorphic to direct products of smaller groups. Every finite group is either simple or can be constructed as a product of simple groups. Thus, simple groups are building blocks from which all groups are built, and the classification of these building blocks was a natural problem. However, simple groups are by no means easy to study, and their cardinality can be immense. For instance, a simple group called the Monster Group has ca. 10^{53} elements (more exactly, 808,017,424,794,512,875,886,459,904,961,710,757,005,754,368,000,000,000).

On the other hand, the field of complex numbers has in its name the adjective “complex” because it can (and usually is) presented as a direct square, i.e., a direct product of the two copies of the field of real numbers in which multiplication is altered. Thus, we have simplicity involving huge cardinalities and complexity of the product with just two factors.

4.2. Paradigmatic Example of the Reduction of Complexity

Probably the most paradigmatic example of the role of the reduction of complexity in mathematics is in elementary linear algebra, where vector spaces of infinite cardinality, but finite dimensions, are presented in a way that allows finitary methods. It is not only the first and fundamental step in all introductory textbooks on the subject, providing conceptual tools for study, but also a key reason for the great importance of linear algebra in its applications outside of mathematics.

Let R be any number field (the symbol here is R because it is most frequently the field of real numbers) and $V(R)$ be any vector space over the field R , with vectors in $V(R)$ indicated here by the underline, e.g., $\underline{v}$ or when vectors are indexed $\underline{v}_k$, and scalars in R by the Greek letters, e.g., α or when indexed α_k .

A finite set of vectors $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$ in a vector space $V(R)$ is independent if for every linear combination of the elements of this set of vectors $\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_k \underline{v}_k$ such that $\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_k \underline{v}_k = \underline{0}$, we have $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$. This condition is equivalent to a more intuitive condition that for every i from 1 to k and any choice of scalars, $\alpha_1, \alpha_2, \dots, \alpha_k \in R$ $\underline{v}_i \neq \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_{i-1} \underline{v}_{i-1} + \alpha_{i+1} \underline{v}_{i+1} + \dots + \alpha_k \underline{v}_k$, i.e., none of the vectors in the set is a linear combination of others. An infinite set of vectors is independent if every one of its finite subsets is independent.

A finite set of vectors $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$ in a vector space $V(R)$ generates the space if for every vector $\underline{v}$ in $V(R)$ there are (not necessarily different) elements $\alpha_1, \alpha_2, \dots, \alpha_k$ of R such that $\underline{v} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_k \underline{v}_k$.

A set of vectors $B \subseteq V(R)$ is a base of $V(R)$ if B is independent and it generates the vector space $V(R)$. If $V(R)$ has a finite base with k elements, it always has many other possible bases, but these bases always have the same cardinality k , called the dimension of $V(R)$; in short, it is reflected in the symbol $V^k(R)$. It turns out that a base can be characterized equivalently, either as a maximal independent set or a minimal set generating the space.

Every vector space $V(R)$ of dimension k is isomorphic to (i.e., mathematically indistinguishable from) the vector space $R^k(R)$ built on the direct product (power) R^k of R . It is easy to construct a function (linear mapping) $\varphi: V^k(R) \rightarrow R^k(R)$ which is invertible, i.e., is an isomorphism. For this purpose, let $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$ be any base of $V^k(R)$. Define the function $\varphi: V^k(R) \rightarrow R^k(R)$ on the vectors from the basis by $\varphi(\underline{v}_i) = (0, 0, \dots, 1, \dots, 0) \in R^k$ where the only non-zero coordinate is on the i -th position of a k -tuple $(0, 0, \dots, 1, \dots, 0)$ and on the other vectors $\underline{v} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_k \underline{v}_k$ in $V(R)$ by $\varphi(\underline{v}) = (\alpha_1, \alpha_2, \dots, \alpha_k)$. The function $\varphi(\underline{v})$ is a linear morphism, and even more, an isomorphism, and we can identify the vector $\underline{v}$ with a sequence of scalars $(\alpha_1, \alpha_2, \dots, \alpha_k)$ from $R^k(R)$.

Here we arrive at the crucial but dangerous point. Our choice of the bases in $V^k(R)$ and $R^k(R)$ from among infinitely many possible ones was arbitrary, and it was dictated by convenience, as it did not make any difference for our purposes. Thus, the identification of the vector $\underline{v}$ with a sequence of scalars $(\alpha_1, \alpha_2, \dots, \alpha_k)$ from $R^k(R)$ is heavily dependent on these arbitrary choices. For this reason, the popular erroneous habit of considering many sequences of objects or numbers as “vectors” is mathematical nonsense, unless the structure of a vector space is present and justified and the choices of the bases are declared. We must remember that the base $B = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$ is a set, not a sequence, and the listing of its element vectors is arbitrary. On the other hand, the construction of the isomorphism is dependent on the order of arguments, and the order in the sequence $(\alpha_1, \alpha_2, \dots, \alpha_k)$ is crucial. Moreover, all vectors in $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$ are different, but the scalars in $(\alpha_1, \alpha_2, \dots, \alpha_k)$ can be equal.

The reason why the possibility of representing vectors by finite sequences of numbers is so important is the reduction of the cardinal complexity of the infinite vector space $V^k(R)$ to a finite number of vectors from the base $B = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$. Whatever we do with vectors in $V^k(R)$, assuming the linearity of such operations, is being reduced to operating on a finite number of vectors $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$ from the base and returning to the $V^k(R)$ using the coordinates $\alpha_1, \alpha_2, \dots, \alpha_k$ and the representation $\underline{v} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 + \dots + \alpha_k \underline{v}_k$.

Here, we have a dramatic reduction of complexity, from infinity to the finite case, in the sense of cardinality, but there is also a reduction of structural complexity. The vector space $V^k(R)$ is reduced to the aggregate of independent components, which are one-dimensional subspaces generated by each of the vectors in the base $B = \{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k\}$. Each of the subspaces can be identified with R . The independence of vectors in B makes the operations on the component subspaces independent.

If the field R is built on real numbers, we can go further and enrich the structure of a vector space $V^k(R)$ by considering an additional concept of a scalar product (in this case, the

so-called dot product) defined in terms of coordinates as follows. For all pairs of vectors, $\underline{v} = (v_1, v_2, \dots, v_k)$ and $\underline{u} = (u_1, u_2, \dots, u_k)$, $\underline{v} \bullet \underline{u} = v_1 u_1 + v_2 u_2 + \dots + v_k u_k$, where v_i and u_i are scalars representing the i -th coordinates of the two vectors, respectively. With the scalar product, we can describe the orthogonality of vectors ($\underline{v} \perp \underline{u}$ if $\underline{v} \bullet \underline{u} = 0$), and the reduction of structural complexity goes much further if we use orthogonal bases, i.e., bases in which all vectors are pairwise orthogonal.

This type of reduction of complexity can be seen outside of mathematics, for instance, in the applications in mechanics, where motion in three-dimensional space can be considered in each dimension independently, or, more generally, not necessarily in physics, in the configuration or phase spaces of any dynamic systems, reducing their complexity by the distinction of their degrees of freedom and considering the changes of the state of the system independently in separate dimensions. Frequently, the non-linearity of some systems is blamed for the complexity of the methods of their study or for their absence. The main reason is that without linearity, i.e., without the vector space structure engaged in the description of such systems, there is no other similarly powerful method of reducing complexity.

Thus far, we have considered only finite-dimensional vector spaces, i.e., vector spaces that have at least one finite base with k vectors (and therefore many other bases of the same cardinality). In infinite-dimensional vector spaces that do not have finite bases, to achieve a reduction of complexity as described above, we have to have an additional structure of a scalar product (more general, not the dot product above, which can be defined only when bases are finite), allowing the introduction of the limits of infinite sequences of vectors, and an additional condition of completeness securing the existence of such limits. The vector spaces equipped with a scalar product with respect to which they are complete are called Hilbert spaces, which play a fundamental role in quantum mechanics.

Interestingly, infinite-dimensional vector spaces devoid of any form of scalar product or any other additional structure are almost absent in mathematics. Their properties are quite exotic and very different from familiar properties in the finite-dimensional case. On the other hand, infinite-dimensional Hilbert spaces are in many respects similar to finite-dimensional vector spaces. This seems natural, as in all finite-dimensional vector spaces, we can introduce the dot product as a special case of the scalar product, and with respect to the dot product, all finite-dimensional vector spaces are Hilbert spaces.

Naturally, the reduction of cardinal complexity in the infinite-dimensional Hilbert spaces cannot be from infinite to finite cardinality, but still we can reduce cardinality from continuum (cardinality of the set of real numbers) to countable cardinality (cardinality of the set of natural numbers). However, the description would have been too long and too technical to include here, and it is not relevant to our objectives.

For us, the example of the simpler case of finite-dimensional vector spaces is sufficient, as it is this case that has become a paradigm for many mathematical techniques of the reduction of complexity, including that of the reduction of information complexity, which will be presented in this paper.

4.3. Examples of Formal Reduction of Information Complexity

The example of the reduction of complexity in finite-dimensional vector spaces is very elementary and very specific. How could it provide a paradigm for the general methods of the reduction? The lifting of generality began with the connection between geometry and vector spaces, which was recognized very early, but mainly through the use of linear algebra as a tool for algebraizing geometry in the spirit of what is now called analytic geometry, which is taught in high school, but which goes far back to the time of Descartes. It was more of a service of linear algebra for geometry.

One of the most influential mathematicians who reversed this relationship was Reinhold Baer in the context of projective geometry. In his very influential 1952 book *Linear Algebra and Projective Geometry* [84], Baer studied the “internal” formalization of the projective geometry structure in terms of the lattice of its subspaces, without any reference to vector spaces, and provided a formal description of how this lattice can be represented in terms of subspaces of a vector space over the field of (real) numbers.

The construction that he presented in his book initiated a long series of fundamental results achieved by others in representing lattices, including lattices of closed subsets in some closure spaces, in the structures arising from vector spaces over some fields or division rings. Such representations are of great practical importance, as numerical fields, especially of real or complex numbers, are common tools in empirical research. Baer’s goal was the coordinatization of the abstract structure of projective geometry, but his method of construction of the division ring (a slight generalization of the concept of a field) of automorphisms of the lattice was instrumental in the work on the coordinatization of quantum logics, which are closure spaces (possibly described equivalently as complete lattices) with closed sets identified with the closed subspaces of a Hilbert space [31,85–91]. The results of these works were important for the study of the formalisms of quantum theory, but also for the more general goal of identifying the conditions for general closure spaces to be represented by the structure of substructures, e.g., subspaces of a vector space, closed subspaces of a Hilbert space, etc.

These results are of great importance for the study of information when information is conceptualized within the theory of closure spaces. One of the benefits is the possibility of lifting some concepts from familiar, less general mathematical theories to the level of the high generality of this theory. However, as often happens in the processes of generalization, it is important to face possible ramifications. Equivalent concepts at the lower level of generality may become separated into essentially different ones at the higher level.

Many concepts present in the paradigmatic example of finite-dimensional vector spaces can be formulated in terms of the theory of closure spaces in such a way that they maintain their original vector space meaning in their vector space representations.

For instance, in a closure space $\langle S, f \rangle$, a subset A of S is independent if for every $x \in A$, $x \notin f(A \setminus \{x\})$. A subset A of S generates S if $f(A) = S$. Finally, a subset A of S is a base if it is independent and generates S . While the first two concepts are frequently used in the theory of closure spaces, the third one, introduced because of its important role in vector spaces and in closure spaces derived from them, loses its importance in more general cases. Many closure spaces do not have bases at all; if they have bases, the bases may have different cardinalities; finally, it may be impossible to determine closed subsets using bases or their elements. For this reason, I introduced a concept of a frame in a closure space $\langle S, f \rangle$ [92,93].

In general, frames are very different from bases, but they share some of the roles of bases in vector spaces, and this is why this concept is useful in the reduction of the complexity of information. In the following, there is a summary of relevant basic properties of frames for this paper. Let $\langle S, f \rangle$ be a closure space and B its subset (not necessarily proper). B is a frame for $\langle S, f \rangle$, if $(*) \forall A \subseteq S \exists B_A \subseteq B: f(A) = f(B_A)$. The condition $(*)$ is equivalent to: $\forall A \subseteq S: f(A) = f(B \cap f(A))$. Frame B is proper if B is a proper subset of $S \setminus f(\emptyset)$; it is a minimal frame if there is no proper subset of B that is a frame.

Minimal frames have the missing property of bases considered above in the description of vector spaces (all subspaces are generated by subsets of frames). Frames always exist, but they may be trivial (improper), equal to S or $S \setminus f(\emptyset)$. Obviously, in finite closure spaces, there are always minimal frames, which, however, can be improper. Infinite closure spaces may have many (even infinitely many) frames, but they may have no minimal frames at all.

If B is a frame in $\langle S, f \rangle$, then we have a canonical closure space $\langle B, f_B \rangle$, in which the closure f_B is a restriction of f to subsets of B , well defined due to the condition defining frames: $\forall A \subseteq S: f(A) = f(B \cap f(A))$.

The concept of a frame of $\langle S, f \rangle$ is different from the concept of a generating subset (it is more restrictive) or base (generative and independent) in a general closure space $\langle S, f \rangle$; however, in the closure space defined earlier in a finite-dimensional vector space, every base is a minimal frame. Thus, in this very special case, the concept of a base and minimal frame coincides.

It is easy to see that B is a frame for $\langle S, f \rangle$ if either one of the equivalent conditions is satisfied:

- $\forall A \subseteq S \exists B_A \subseteq B: f(A) = f(B_A)$ (the defining condition).
- $\forall A \subseteq S: f(A) = f(B \cap f(A))$.
- $\forall C, D \subseteq X: f(C) \cap B = f(D) \cap B \Rightarrow f(C) = f(D)$.

We can see that the action of the closure operator f is determined by the action of f on subsets of frame B . This means that these are frames, not the closure space bases that are actual generalizations of bases in vector spaces.

We can now consider a reversed process, where we start from a closure space $\langle B, g \rangle$ and a set S , such that $B \subseteq S$ and B is a frame in $\langle S, f \rangle$. Then, we call $\langle S, f \rangle$ a frame extension of $\langle B, g \rangle$ if the action of g on B is a restriction of the action of f to subsets of B . It turns out that the extensions are not unique, but they always produce closure spaces with identical logics, i.e., their lattices of closed subspaces are the same. The differences between the extensions are similar to the differences in vector spaces arising from the arbitrary choice of the order of bases, although in general, they may occur for additional, but equally marginal reasons, such as the choice of the closure of the empty space. The crucial point is that they all have the same logic.

With the concept of a frame in closure space, we can consider the reduction of information complexity for any information system $\langle S, f \rangle$ in which there are proper frames, which is achieved by the transition from the closure space $\langle S, f \rangle$ to $\langle B, f_B \rangle$ described above. If an information space has a minimal frame, the reduction of complexity is maximal. However, in general, minimal frames may have different cardinalities, so it may be maximal, but not maximum reduction.

4.4. Other Less Formal Forms of Reduction of Information Complexity

The use of frames is only one of many ways to reduce complexity. It is the closest to the paradigmatic example of vector spaces. Also, it requires rather limited involvement of the mathematical formalism of closure spaces. The other forms of reduction formalized in my earlier works use the decomposition of the logic of an information system into direct products of component information systems [94] or the partitioning of the information system into the orbits of the action of symmetry groups [95,96]. However, they require quite extensive mathematical preliminaries, which could obscure the purpose of this paper, which is not an elaboration of the diverse forms of reduction, but rather explaining how it is understood in the context of intelligence. Instead, the remaining part of this paper provides less formalized but more intuitive examples of the reduction of information complexity.

We can find an early example of the reduction in Euclid's *Elements*, one of the most influential works of Mediterranean Antiquity that shaped the style of inquiry in Western civilization. Although Euclid's presentation of planar geometry in its first six books of the *Elements* was impressively extensive and advanced, the influence came not with its content (Euclid included there geometric knowledge accumulated by several generations of geometers) but with the way it was organized in what we now call the axiomatic theory, starting from the five geometric "postulates" and five "common notions" (today we would

call them geometric axioms and general axioms) and deriving from them, using logical inference formulated as a construction with compass and ruler, all 173 theorems. At that time, it was a commonly accepted view that the axioms, being rather obvious, did not require proof of truth, and the more complex theorems derived from the axioms inherited their truth automatically.

Euclid was not the first to use this methodology. We can find its earlier use in Aristotle's *Analytics*, where Aristotle derived, using seemingly obvious "rules of conversion," the validity of the remaining 11 valid syllogisms (syllogisms that from true premises lead to true conclusions) from the three perfect syllogisms, which, being self-evident, seemed obviously valid. However, the *Elements* made the method easy to understand and follow, which made it the primary tool in education of the methods of inquiry and reasoning in the next two millennia.

Euclid's *Elements* has remained an ideal to follow not only in education but also in intellectual practice. In Antiquity, the same methodology was used by Archimedes in his treatise on statics, *On the Equilibrium of Planes or on the Centers of Gravity on Planes*. Outside of what we now call mathematical sciences, the use of the method was more difficult, but there were many attempts to follow it. In the Middle Ages, Thomas Aquinas used it in his *Summa Contra Gentiles*. In the Modern Era, Baruch Spinoza used it for *Ethics*, and Thomas Hobbes presented his theory of politics in his 1660 *Leviathan or The Matter of the Forme of a Commonwealth Ecclesiasticall and Civil*. The closest to the original was the fundamental 1687 work *Principia* of Isaac Newton, who explicitly patterned it on the *Elements*.

The example of Euclid's *Elements* leads to the explicit general formulation of logic using concepts presented in the preceding section. Alfred Tarski developed his very influential description of propositional logic using as its central concept the closure operator $Cn(X)$ on the set of sentences, assigning to every set of well-formed sentences X , the set extending X by all sentences which can be inferred from sentences in X [97]. Based on Tarski's closure space formalism, many more general closure space formalisms of logic have been developed by multiple authors. Using the formalism developed by Tarski, we can see that axiom systems are exactly frames within theories, and the axiomatic method is exactly an expression of the reduction of information complexity presented above.

The most popular example of the reduction of complexity is William Ockham's principle of parsimony, known as Ockham's Razor. In time, it has acquired the frequently quoted ontological parsimony form "One should not multiply entities beyond necessity—*Entia non sunt multiplicanda sine necessitate*". However, this formulation of the principle is not to be found in any of Ockham's extant writings [98]. Instead, it most likely was originally an epistemological directive: "For nothing ought to be posited without a reason given, unless it is self-evident (literally, known through itself) or known by experience or proved by the authority of Sacred Scripture" [99].

For centuries, it served as a directive for effective inquiry in a slightly abbreviated and general, but less specific form, stating that of the two ways to explain phenomena, the simpler one should be chosen. With its popularity and frequent adaptations to particular contexts, the understanding of simplicity has become diverse, including, for instance, the principle of minimal action [100]. We can interpret it, concluding from its diverse formulations, as an expression of the view that a simpler explanation means one that requires less information, the meta-principle justifying it being the reduction of complexity.

Thus far, the examples have been limited to the reduction of information complexity in its human contexts. There is another example of the reduction of the complexity of information in the context of computation. In this case, complexity or simplicity has a very clear meaning and is not related to human understanding. Andrei Kolmogorov and, independently, Gregory Chaitin proposed a measure of information encoded on a tape

of the Turing Machine as the length of the shortest program (i.e., the number of cells on the tape taken by the program) for a Universal Turing Machine which, when run, would produce this encoding [101–103]. This measure was later interpreted as a measure of algorithmic complexity. Obviously, this measure of algorithmic complexity reflects how much the original complexity can be reduced. Interestingly, such a measure is not computable, i.e., a Turing Machine cannot find such a shortest program, and therefore a computer with the standard architecture cannot do it [104].

5. Conclusions

The short paper to which this paper is a companion proposed a concept of intelligence overarching the concepts developed for all its multiple manifestations. This paper has as its main objective a metaphysical and epistemological analysis of the triad of involved concepts, information, complexity, intelligence, and their mutual relationships. The first and second concepts are discussed more extensively, while the discussion of the third one, intelligence, to avoid repetition of the content of the earlier conference paper, is focused more on the concepts usually considered along with it, action, agency, and causality. The paper provides extensive argumentation for the elimination of the concepts of action and causality, focusing on interaction and reformulation of agency in terms of interaction and information.

Naturally, while information and complexity have been considered in a very general way in this paper, in the discussion of intelligence, most examples of their manifestation are related to some degree to their human context due to the involvement of information encoded in the form of human-type languages. Very little has been said about the non-human reduction of complexity. This is natural, because in the relevant literature there are many reports about empirical research on information and complexity, but far fewer reports on the reduction of complexity and almost none about intelligence as the reduction of complexity, as this idea is rather new. After all, most of the intelligence research is based on the involvement of non-human agents in human-type behavior or the use of human-type languages.

However, this does not mean that there is no empirical research on the intelligence of animals expressed as a reduction of the complexity of information. Unfortunately, almost always, the interpretations of the results involve overstatements, such as claims that the behavioral observations confirm that the animals perform what is called reasoning, which suggests a human-type cognitive process. All that is confirmed in such empirical studies is the presence of some cognitive process in which animals are capable of simplifying the information coming from their environment and adapting their behavior to this simplified information. Calling it reasoning suggests a process in at least some ways similar to human reasoning, which is, or at least can be, under conscious control, but the mechanism is still unknown. The best example of such simplification is the so-called “transitive reasoning” in which animals, for instance, chimpanzees, identify the orderings of objects of many different types that are transitive (if $a < b$ and $b < c$, then $a < c$). Then they use this discovery to achieve their goals (typically to get more snacks or more water). That this involves reduction of the complexity of information is very clear from the experiments, but whether the mechanism is related in any way to human reasoning, nobody knows at present [105].

I believe that the theoretical, abstract design of intelligence described in the previous and present papers, free from anthropomorphization, opens a new, very broad field of empirical study. As I wrote earlier, the separation of meta-intelligence into separate manifestations depends on the way complexity and its reduction are understood. The search for examples of intelligence in less traditional contexts requires an identification of suitable forms of reduction as the first step in research.

Thus, finally, let us consider the legitimate question of how we can describe intelligence as a reduction of the complexity of information in the case of general non-human agents, which are not only non-human but very different from humans. Here, I would propose the following answer.

The intelligence of non-human agents (i.e., systems capable of entering into internally controlled interactions with their environment that alter the agent or the environment, or typically both) can be understood as a qualification assessed by considering the information complexity in the environment and the reduction of complexity brought by the agent's interactions. In anticipation of the objection that this is very different from intelligence in a human context, in the examples above I can see a lot of similarities. For instance, for Euclid, the (cultural) environment was the accumulated body of geometric knowledge consisting of a multitude of theorems inherited from several authors, some with proofs, some without. We will never know how many of his 173 theorems were of his authorship, most likely only a few. However, he reduced all that he inherited to five principles and the methodology of deriving the theorems from these principles. This highly structured cultural "geometric" environment served many generations of others. Similar interactions with cultural, social, economic, and natural environments can be searched for analogous reduction by non-human agents with the use of the very general, not necessarily linguistic formalism of information and reduction of its complexity described in the paper.

Funding: This research received no external funding.

Institutional Review Board Statement: This study did not include humans or animals and did not require ethical approval.

Informed Consent Statement: Not applicable.

Data Availability Statement: This is a purely conceptual work that did not collect or analyze any data.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ZFC Zermelo–Fraenkel axiomatic set theory with the Axiom of Choice
PC Predicate calculus of the first order

References

- Schroeder, M.J. Intelligence as the Capacity to Overcome the Complexity of Information: Search for Unity in the Diverse Forms of Intelligence. *Proceedings* **2025**, *126*, 14. [\[CrossRef\]](#)
- The 1st International Online Conference of the Journal Philosophies. In Proceedings of the Intelligent Inquiry into Intelligence-Contributing to the 2025 IS4SI Summit, Online, 10–14 June 2025.
- Brenner, J.E. *Logic in Reality*; Springer: Dordrecht, The Netherlands, 2008.
- Asher, N.; Kamp, H. Self-Reference, Attitudes, and Paradox. In *Properties, Types, and Meaning I*; Chierchia, G., Partee, B.H., Turner, R., Eds.; Kluwer: Dordrecht, The Netherlands, 1989; pp. 85–158.
- Liebert, M.A.; Benner, S.A. Defining Life. *Astrobiology* **2010**, *10*, 1021–1030. [\[CrossRef\]](#) [\[PubMed\]](#)
- Buckland, M.K. Information as Thing. *J. Am. Soc. Inf. Sci.* **1991**, *42*, 351–360. [\[CrossRef\]](#)
- Schroeder, M.J. Is Using Language in the Study of Information Bootstrapping? In Proceedings of the 7th International Conference on Philosophy of Information (ICPI) contributing to the 2025 International Summit of IS4SI, Varna, Bulgaria, 24 June 2025; p. 8.
- Glock, H.-J. *What Is Analytic Philosophy?* Cambridge University Press: Cambridge, UK, 2008.
- Locke, J. *An Essay Concerning Human Understanding*; Winkler, K.P., Ed.; Hackett Publ. Co.: London, UK, 1996; pp. 33–36.
- Aristotle. *De Anima (On the Soul)*; Hett, W.S., Translator; Heinemann: London, UK, 1936.
- Kant, I. *Critique of Pure Reason*; Guyer, P.; Wood, A.W., Translators; Cambridge University Press: Cambridge, UK, 1999.
- Butler, S. The Book of the Machines. In *Erewhon or Over the Range*; Trubner: London, UK, 1872; pp. 190–223, Chapters XXI–XXIII.
- Von Neumann, J.; Burke, A.W. *Theory of Self-Reproducing Automata*; University of Illinois Press: Urbana, IL, USA, 1966.

14. Turing, A.M. Computing Machinery and Intelligence. *Mind* **1950**, *59*, 433–460. [CrossRef]
15. Editorial Nature Review Physics. How to edit anthropomorphic language about artificial intelligence. *Nat. Rev. Phys.* **2023**, *5*, 263. [CrossRef]
16. Aristotle. *Metaphysics*. In *The Basic Works of Aristotle*; McKeon, R., Ed.; Random House: New York, NY, USA, 1941; pp. 689–926.
17. Bacon, F. *Novum Organum*; Collier: New York, NY, USA, 1902.
18. Hume, D. *A Treatise of Human Nature*; John Noon: London, UK, 1739. Available online: <https://davidhume.org/texts/t/> (accessed on 24 March 2026).
19. Mill, J.S. *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*; John, W., Ed.; Parker: London, UK, 1843.
20. Russell, B. *Mysticism and Logic, and Other Essays*; Unwin: London, UK, 1963.
21. Wu, T. Boltzmann's H Theorem and the Loschmidt and the Zermelo Paradoxes. *Int. J. Theor. Phys.* **1975**, *14*, 289–294. [CrossRef]
22. Descartes, R. *Principia Philosophiae*; Roger, V.; Miller, R.P., Translators; D. Reidel: Dordrecht, The Netherlands, 1983.
23. Newton, J.S. *Philosophiae Naturalis Principia Mathematica*; S. Pepys Reg. Soc. Praeses: London, UK, 1686. Available online: <https://cudl.lib.cam.ac.uk/view/PR-ADV-B-00039-00001/9> (accessed on 31 March 2026).
24. Newton, I. General Scholium appended to 13g. In *The Mathematical Principles of Natural Philosophy*; Motte, A., Translator; Motte: London, UK, 1729; Volume 2, p. 392. Available online: <https://www.gutenberg.org/cache/epub/76404/pg76404-images.html> (accessed on 31 March 2026).
25. Zeeman, E.C. Causality Implies the Lorentz Group. *J. Math. Phys.* **1964**, *5*, 490–493. [CrossRef]
26. Cegła, W.; Jadczyk, A. Causal Logic on Minkowski Space. *Commun. Math. Phys.* **1977**, *57*, 213–217. [CrossRef]
27. Cegła, W.; Jadczyk, A. States on Causal Logic. *Lett. Math. Phys.* **1979**, *3*, 109–112. [CrossRef]
28. Casini, H. The logic of causally closed spacetime subsets. *Class. Quantum Grav.* **2002**, *19*, 6389. [CrossRef]
29. Cegła, W.; Florek, J. Ortho and causal closure as closure operation in the causal logic. *Int. J. Theor. Phys.* **2005**, *44*, 11–19. [CrossRef]
30. Cegła, W.; Jancewicz, B. Non-modular Lattices Generated by Causal Structure. *Int. J. Math. Phys.* **2013**, *44*, 11–19.
31. Jauch, J.M. *Foundations of Quantum Mechanics*; Addison-Wesley: Reading, MA, USA, 1968.
32. Zecca, A. Non-Standard Aspects of Minkowski Causal Logic. *Found. Phys. Lett.* **1996**, *9*, 295–300. [CrossRef]
33. Mayet, R. Orthomodular Lattice in Relativistic Non-quantum Mechanics. *Int. J. Theor. Phys.* **1995**, *34*, 1595–1599. [CrossRef]
34. Sapolsky, R. *Determined: A Science of Life Without Free Will*; Penguin Press: London, UK, 2023.
35. Libet, B. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* **1985**, *8*, 529–539. [CrossRef]
36. Soon, C.S.; Brass, M.; Heinze, H.J.; Haynes, J.D. Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.* **2008**, *11*, 543–545. [CrossRef] [PubMed]
37. Soon, C.S.; He, A.H.; Bode, S.; Haynes, J.D. Predicting free choices for abstract intentions. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6217–6222. [CrossRef] [PubMed]
38. Maoz, U.; Yaffe, G.; Koch, C.; Mudrik, L. Neural precursors of decisions that matter—An ERP study of deliberate and arbitrary choice. *eLife* **2019**, *8*, e39787. [CrossRef] [PubMed]
39. Conway, J.; Kochen, S. The Free Will Theorem. *Found. Phys.* **2006**, *36*, 1441–1473. [CrossRef]
40. Conway, J.; Kochen, S. The Strong Free Will Theorem. *Not. AMS* **2009**, *56*, 226–232.
41. Thorndike, E.L. Intelligence and its measurement: A symposium. *J. Educ. Psychol.* **1921**, *12*, 3–24. [CrossRef]
42. Neisser, U.; Boodoo, G.; Bouchard, T.J.; Boykin, A.W., Jr.; Brody, N.; Ceci, S.J.; Halpern, D.F.; Loehlin, J.C.; Perloff, R.; Sternberg, R.J.; et al. Intelligence: Knowns and Unknowns. *Amer. Psychol.* **1996**, *51*, 77–101. [CrossRef]
43. Landauer, R. Information is Physical. *Phys. Today* **1991**, *44*, 23–29. [CrossRef]
44. Landauer, R. The Physical Nature of Information. *Phys. Lett. A* **1996**, *217*, 188–193. [CrossRef]
45. Landauer, R. Information is a Physical Entity. *Phys. A* **1999**, *263*, 63–67. [CrossRef]
46. Landauer, R. Information is Inevitably Physical. In *Feynman and Computation: Exploring the Limits of Computers*; Hey, A.J.G., Ed.; Perseus: Reading, MA, USA, 1999; pp. 77–92.
47. Levin, M. Collective Intelligence of Morphogenesis as a Teleonomic Process. In *Evolution "On Purpose": Teleonomy in Living Systems*; Corning, P.A., Kauffman, S.A., Noble, D., Shapiro, J.A., Vane-Wright, R.I., Pross, A., Eds.; MIT Press: Cambridge, MA, USA, 2023; pp. 175–197.
48. Tuomela, R. Actions by Collectives. *Philos. Perspect.* **1989**, *3*, 471–496. [CrossRef]
49. Tarski, A. The Concept of Truth in Formalized Languages. In *Logic, Semantics, Metamathematics*; Woodger, J.H., Translator; Corcoran, J., Ed.; Hackett: Cambridge, MA, USA, 1933/1983.
50. Rasiowa, H. *The Mathematics of Metamathematics*; Sikorski, R., Ed.; Państwowe Wydawnictwo Naukowe: Warszawa, PL, USA, 1963.
51. Albert, H. *Treatise on Critical Reason*; Princeton University Press: Princeton, NJ, USA, 1985; Chapter I, Section 2.
52. Zermelo, E. Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundam. Math.* **1930**, *16*, 29–47. [CrossRef]

53. Gödel, K. The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis. *Proc. Natl. Acad. Sci. USA* **1938**, *24*, 556–557. [CrossRef] [PubMed]
54. Jech, T. *Set Theory: The Third Millennium Edition, Revised and Expanded*; Springer Monographs in Mathematics; Springer: Berlin, Germany, 2003.
55. Barnes, J. *Early Greek Philosophy*, 2nd ed.; Penguin Books: London, UK, 2001.
56. Plato. *Philebus*; Jowett, B., Translator; Project Gutenberg: Salt Lake City, UT, USA, 2008. Available online: <https://www.gutenberg.org/files/1744/1744-h/1744-h.htm> (accessed on 31 March 2026).
57. Plato. *Parmenides*; Project Gutenberg: Salt Lake City, UT, USA, 1687. Available online: <https://www.gutenberg.org/ebooks/1687> (accessed on 31 March 2026).
58. James, W. *The Principles of Psychology*; Holt: New York, NY, USA, 1896; Volume 2.
59. James, W. The One and the Many. In *Some Problems of Philosophy*; James, W., Ed.; Longman's Green and Co.: New York, NY, USA, 1911.
60. James, W. The One and the Many. In *The Writings of William James: A Comprehensive Edition*; McDermott, J.J., Ed.; Random House: New York, NY, USA, 1967; pp. 258–270.
61. Chan, W.-T. *A Source Book in Chinese Philosophy*; Princeton University Press: Princeton, NJ, USA, 1973.
62. Dalley, S. *Myths from Mesopotamia: The Creation, the Flood, Gilgamesh, and Others*; Oxford University Press: Oxford, UK, 1989.
63. Young, P. *The Nature of Information*; Praeger: New York, NY, USA, 1987.
64. Cantor, G. Über unendliche, lineare Punktmannigfaltigkeiten, 5. *Math. Ann.* **1883**, *21*, 545–586. [CrossRef]
65. Husserl, E. *Philosophie der Arithmetik*; Pfeffer: Halle, Germany, 1891.
66. Ortiz Hill, C.; Rosado Haddock, G. *Husserl or Frege, Meaning, Objectivity and Mathematics*; Open Court: Chicago, IL, USA, 2000.
67. Schroeder, M.J. From Philosophy to Theory of Information. *Int. J. Inf. Theory Appl.* **2011**, *18*, 56–68.
68. Schroeder, M.J. Semantics of Information: Meaning and Truth as Relationships Between Information Carriers. In *Proceedings of IACAP 2011, Aarhus University, 4–6 July 2011*; Ess, C., Hagengruber, R., Eds.; Monsenstein und Vannerdat Wiss.: Münster, Germany, 2011; pp. 120–123.
69. Schroeder, M.J. Concept of Information as a Bridge between Mind and Brain. *Information* **2011**, *2*, 478–509. [CrossRef]
70. Schroeder, M.J. Ontological Study of Information: Identity and State. *Kybernetos* **2014**, *43*, 882–894. [CrossRef]
71. Schroeder, M.J. The Difference that Makes a Difference for the Conceptualization of Information. *Proceedings* **2017**, *1*, 221. [CrossRef]
72. Schroeder, M.J. Theoretical Study of the Concepts of Information Defined as Difference and as Identification of a Variety. In *Philosophy and Methodology of Information—The Study of Information in a Transdisciplinary Perspective*; Dodig-Crnkovic, G., Burgin, M., Eds.; World Scientific Series in Information Studies; World Scientific: Singapore, 2019; Volume 10, pp. 289–314.
73. Schroeder, M.J. Theoretical Unification of the Fractured Aspects of Information. In *Understanding Information and Its Role as a Tool: In Memory of Mark Burgin (in 2 Parts)*; Schroeder, M.J., Hofkirchner, W., Eds.; World Scientific: Singapore, 2025; Chapter 6; pp. 131–183.
74. Birkhoff, G. *Lattice Theory*, 3rd ed.; American Mathematical Society Colloquium Publications, AMS: Providence, RI, USA, 1967; Volume XXV.
75. Schroeder, M.J. Search for an Algebraic Structure of Not-Necessarily Algebraic Structures. In *Algebras, Logics, Languages and Related Areas*; Kunimochi, Y., Ed.; RIMS Kôkyûroku, Kyoto, Research Institute for Mathematical Sciences, Kyoto University: Kyoto, Japan, 2018; No. 2096; pp. 59–68. Available online: www.kurims.kyoto-u.ac.jp/kyodo/kokyuroku/contents/2096.html (accessed on 31 March 2026).
76. Schroeder, M.J. Relationship between Rota's Logic of Information Based on Partitions and Logic of Information Defined in a Closure Space. In *Group, Ring, Language, and Related Areas in Computer Science*; Nishinaka, Ed.; Research Institute for Mathematical Sciences (RIMS), Kyoto University: Kyoto, Japan, 2023; Kôkyûroku No. 2265; Chapter 13; p. 10. Available online: <https://www.kurims.kyoto-u.ac.jp/kyodo/kokyuroku/contents/pdf/2265-13.pdf> (accessed on 31 March 2026).
77. Schroeder, M.J. Search for syllogistic structure of semantic information. *J. Appl. Non-Class. Log.* **2012**, *22*, 83–103. [CrossRef]
78. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
79. Hartley, R.V.L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, *7*, 535–563. [CrossRef]
80. Bar-Hillel, Y.; Carnap, R. *An Outline of a Theory of Semantic Information*; Technical Report No. 247; Research Laboratory of Electronics, MIT: Cambridge, MA, USA, 1952; reprinted in *Language and Information: Selected Essays on Their Theory and Application*; Bar-Hillel, Y., Ed.; Addison-Wesley: Reading, MA, USA, 1964; pp. 221–274.
81. Schroeder, M.J. The Complexity of Complexity: Structural vs. Quantitative Approach. In *Proceedings of the International Conference on Complexity, Cybernetics, and Informing Science CCISE 2013, Porto, Portugal, 30 June–6 July 2013*; pp. 41–46. Available online: <https://www.iiis.org/Proceedings/2013Proc/ViewProc2013.asp?id=CC0113> (accessed on 31 March 2026).
82. Schroeder, M.J. Hierarchic Information Systems in a Search for Methods to Transcend Limitations of Complexity. *Philosophies* **2016**, *1*, 1–14. [CrossRef]

83. Schroeder, M.J. Structural and Quantitative Characteristics of Complexity in Terms of Information. In *Information and Complexity, World Scientific Series in Information Studies*; Burgin, M., Calude, C.S., Eds.; World Scientific: New Jersey, NJ, USA, 2017; Volume 6, pp. 117–175.
84. Baer, R. *Linear Algebra and Projective Geometry*; Academic Press: New York, NY, USA, 1952.
85. Finkelstein, D.; Jauch, J.; Schiminovich, S.; Speiser, D. Foundations of quaternionic quantum mechanics. *J. Math. Phys.* **1962**, *3*, 207–220. [[CrossRef](#)]
86. Piron, C. Axiomatique quantique. *Helv. Phys. Acta* **1964**, *37*, 439.
87. Piron, C. Survey of general quantum physics. *Found. Phys.* **1972**, *2*, 287–314. [[CrossRef](#)]
88. Piron, C. *On the Foundations of Quantum Physics*; W.A. Benjamin: New York, NY, USA, 1976.
89. Solér, P.M. Characterization of Hilbert Spaces by Orthomodular Spaces. *Commun. Algebra* **1995**, *23*, 219–243. [[CrossRef](#)]
90. Rédei, M. Facets of Quantum Logic. *Stud. Hist. Philos. Mod. Phys.* **2001**, *32*, 101–111. [[CrossRef](#)]
91. Cassinelli, G.; Lahti, P. Quantum Mechanics: Why Complex Hilbert Space? *Phil. Trans. R. Soc. A* **2017**, *375*, 20160393. [[CrossRef](#)] [[PubMed](#)]
92. Schroeder, M.J. Quantum Coherence without Quantum Mechanics in Modeling the Unity of Consciousness. In *QI 2009*; LNAI Volume 5494; Bruza, P., Sofge, D., Lawless, W., van Rijsbergen, K., Klusch, M., Eds.; Springer: Heidelberg, Germany, 2009; pp. 97–112.
93. Schroeder, M.J. Algebraic Model for the Dualism of Selective and Structural Manifestations of Information. In *Logics, Algebras, and Languages in Computer Science*; Kondo, M., Ed.; RIMS Kôkyûroku, No. 1915; Research Institute for Mathematical Sciences, Kyoto University: Kyoto, Japan, 2014; pp. 44–52. Available online: <http://www.kurims.kyotou.ac.jp/~kyodo/kokyuroku/contents/1915.html> (accessed on 31 March 2026).
94. Schroeder, M.J. Structural and Quantitative Characteristics of Information and Complexity. In *New Contact Points of Algebraic Systems, Logics, Languages, and Computer Sciences*; Kondo, M., Ed.; RIMS Kôkyûroku: Kyoto, Japan; Research Institute for Mathematical Sciences, Kyoto University: Kyoto, Japan, 2015; No. 1964; pp. 88–97. Available online: <http://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/1964.html> (accessed on 31 March 2026).
95. Schroeder, M.J. Concept of Symmetry in Closure Spaces as a Tool for Naturalization of Information. In *Algebraic System, Logic, Language and Computer Science*; Horiuchi, K., Ed.; RIMS Kôkyûroku: Kyoto, Japan; Research Institute for Mathematical Sciences, Kyoto University: Kyoto, Japan, 2016; No. 2008; pp. 29–36. Available online: <http://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/pdf/2008-04.pdf> (accessed on 31 March 2026).
96. Schroeder, M.J. Structural Abstraction and Study of Symmetry. In *Algebraic Systems, Logic, Language, and Related Areas in Computer Science II*; Adachi, T., Ed.; RIMS Kôkyûroku: Research Institute for Mathematical Sciences, Kyoto University: Kyoto, Japan, 2021; No. 2188-12; pp. 71–80.
97. Tarski, A. *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*, 2nd ed.; Corcoran, J., Ed.; Hackett: Indianapolis, IN, USA, 1983.
98. Thorburn, W.M. The Myth of Occam’s Razor. *Mind* **1918**, *27*, 345–353. [[CrossRef](#)]
99. Spade, P.V. (Ed.) *The Cambridge Companion to Ockham*; Cambridge University Press: Cambridge, UK, 1999.
100. Hoffmann, R.; Minkin, V.I.; Carpenter, B.K. Ockham’s Razor and Chemistry. *HYLE—Int. J. Philos. Chem.* **1997**, *3*, 3–28.
101. Kolmogorov, A.N. Three approaches to the definition of the quantity of information. *Probl. Inf. Transm.* **1965**, *1*, 3–11.
102. Chaitin, G.J. On the Length of Programs for Computing Finite Binary Sequences. *J. ACM* **1966**, *13*, 547–569. [[CrossRef](#)]
103. Kolmogorov, A.N. Combinatorial Foundations of Information Theory and the Calculus of Probabilities. *Russ. Math. Surv.* **1983**, *38*, 29–40. [[CrossRef](#)]
104. Shen, A.; Uspensky, V.A.; Vereshchagin, N.K. *Kolmogorov Complexity and Algorithmic Randomness*; Mathematical Surveys and Monographs; American Mathematical Society: Providence, RI, USA, 2017; Volume 220. [[CrossRef](#)]
105. Lazareva, O. Transitive inference and transitivity: Two sides of the same coin? *J. Exp. Anal. Behav.* **2025**, *124*, e70031. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.