

Article

# Information and Inference

Paul Walton

Capgemini UK, Forge End, Woking, Surrey GU21 6DB, UK; paulnicholaswalton@gmail.com;  
Tel.: +44-13-0688-3140

Academic Editor: Willy Susilo

Received: 16 April 2017; Accepted: 22 May 2017; Published: 27 May 2017

**Abstract:** Inference is expressed using information and is therefore subject to the limitations of information. The conventions that determine the reliability of inference have developed in information ecosystems under the influence of a range of selection pressures. These conventions embed limitations in information measures like quality, pace and friction caused by selection trade-offs. Some selection pressures improve the reliability of inference; others diminish it by reinforcing the limitations of the conventions. This paper shows how to apply these ideas to inference in order to analyse the limitations; the analysis is applied to various theories of inference including examples from the philosophies of science and mathematics as well as machine learning. The analysis highlights the limitations of these theories and how different, seemingly competing, ideas about inference can relate to each other.

**Keywords:** information; philosophy of science; inference; induction; information quality; information friction; machine learning

## 1. Introduction

Inference pervades society. Examples include legal trials, science, computer modelling and those inferences that routinely form the basis of individual or organisational decisions. Increasingly, new forms of inference, based on business analytics and machine learning are becoming embedded in digital services [1]. However, inference uses information and is therefore subject to the difficulties inherent in information. This paper examines inference in this context building on the techniques developed in [2–4], summarised in Section 2 and showing how they apply to inference.

Section 3 explores the relationship between information and inference and, in particular, the impact of information ecosystems (just “ecosystem” where the context is clear) that have developed under the influence of a range of selection processes. In each ecosystem, conventions have developed that define the ways in which information is abstracted, processed and exchanged by the various actors in the ecosystem, which we call Interacting Entities (IEs). In different ecosystems, people, animals, organisations, parts of organisations, living cells [5,6] and computer systems are all IEs. The conventions incorporate the trade-offs made to information measures like pace and friction and, in particular, the impact on information quality [1,3]. Inside the envelope of environmental circumstances in which selection pressures apply, IEs need to be able to rely on inference to achieve favourable enough outcomes; this is how the term “reliable” is used in this paper. However, the conventions that drive these inferences may not provide sufficiently reliable inferences for other purposes or under different selection pressures [7].

These trade-offs and the constraints of individual circumstances (shortages of time or resources, for example) affect information in two ways. Individual interactions may not achieve the quality standards of the ecosystem embedded in ecosystem conventions. Secondly, ecosystem conventions may impose limitations on the effectiveness of inference. Inference uses a number of different mechanisms

like induction, deduction and probability and the constraints that apply to information may be different in each case. Section 4 examines the reliability of inference.

In addition to selection trade-offs, there are pragmatic issues with inference—chaotic effects [8] and the impossibility of completely accurate measurement mean that it may not be possible to discriminate different environment states reliably. At a deeper level, reliable models of complicated systems do not exist (think of economics); and even for the simplest system models may not exist (think of the three-body problem). In [7], the authors express this in the following way:

*“What makes relevant inferences possible [ . . . ] is the existence in the world of dependable regularities. Some, like the laws of physics, are quite general. Others, like the bell-food regularity in Pavlov’s lab, are quite transient and local. [ . . . ] No regularities, no inference. No inference, no action.”*

The corollary of this is that, generally, an approach to inference is reliable only to the extent that selection pressures have challenged it thoroughly. Moreover, it is reliable enough only within the envelope in which the selection pressures have applied.

In Section 5, we use the ideas of Sections 3 and 4 to analyse some examples in the philosophy of science, mathematics and machine learning. By considering how the types of inference are subject to selection processes and how closely they match the generic inference pattern we can identify their limitations. We can also show how different, seemingly competing, ideas about inference in the philosophy of science relate to each other.

## 2. The Model for Information

This section sets out the approach to information contained in [1–4]. In this approach, information corresponds to a set of physical properties that conform to conventions that evolve in information ecosystems. Consider the elements of this statement in turn.

Information processing entities interact with their environment and so we call them Interacting Entities (IEs). Through interaction, IEs gain access to resources like money, food, drink or votes for themselves or related IEs, and develop new capabilities including channels to other IEs and sources of information (see Figure 1).

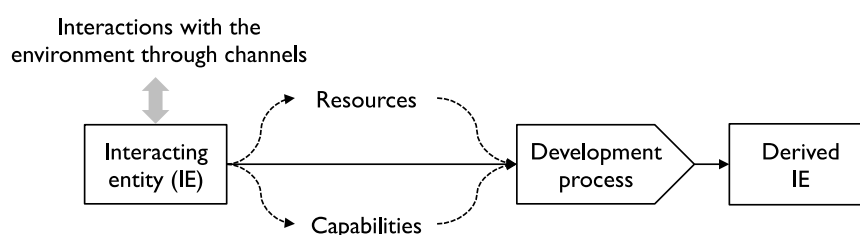


Figure 1. Selection processes.

These ideas apply very broadly. In [5,6], the authors discuss information in relation to living cells and say that “living existence is informational” and that “the signaling system provides the cell with a generic ability to ‘grab’ or ‘abduct’ information from the environment”. With this starting point they discuss meaning, value and adaptability in the context of information.

Through a range of processes and feedback mechanisms, derived IEs (e.g., children, new product versions, changed organisations) are created from IEs. The health of the IE—its ability to continue to interact and achieve favourable outcomes—and the nature of any derived IE depend on the resources and channels the IE has access to and the outcomes it achieves. The nature of interactions and the outcomes available create selection pressures that affect the characteristics of derived IEs. Examples of selection processes include the market, natural selection, elections, personal choice and sexual selection but for any IE different combinations of selection process may apply.

As selection pressures change, there is an incentive to combine the conventions of different ecosystems where this will achieve more favourable outcomes. In [5,6], the authors discuss knowledge recombination and its relationship with “separate, specialized disciplines”, relating to information ecosystems, and the “disciplinary reliance on accepted paradigms”, relating to ecosystem conventions. They relate the dynamics of science to the evolutionary nature of technological change, quoting [9]:

*“Technologies therefore share common ancestries and combine, morph, and combine again to create further technologies. Technology evolves much as a coral reef builds itself from activities of small organisms—it creates itself from itself.”*

The ability of an IE to achieve favourable outcomes from an environment state requires information processing—information enables connections to be made between the environment state, potential outcomes and the actions to help achieve them. So, we can talk of descriptive, predictive and prescriptive information, respectively.

Selection processes lead to the formation of information ecosystems. Note that the use of the term “ecosystem” here is consistent with [1–4] and is based on an evolutionary perspective like that of ecology in the natural world but differs from that used in [10]. Examples include English speakers, computer systems that exchange specific types of banking information, Mathematicians and many others. Each ecosystem has its own conventions for exchanging and processing information. Within different ecosystems, modelling tools such as languages, Mathematics and computer protocols have evolved to structure information within the ecosystem. The conventions shape how potentially arbitrary physical properties are interpreted. The existence of ecosystem conventions confirms that information is relative (see, for example, [11]—“the information in a message is a function of the recipient as well as the message” or [12]—“information is not an absolute but depends on the context in which it is being used”). An IE outside the ecosystem may not be able to interpret the information—think of a classical languages scholar trying to understand quantum mechanics.

Call a slice a contiguous subset of space–time. A slice can correspond to an entity at a point in time (or more properly within a very short interval of time), a fixed piece of space over a fixed period of time or, much more generally, an event that moves through space and time. This definition allows great flexibility in discussing information. For example, slices are sufficiently general to support a common discussion of nouns and verbs, the past and the future.

Information corresponding to ecosystem conventions we call content with respect to the ecosystem. Content is structured in terms of chunks and assertions. A chunk specifies a constraint on sets of slices (e.g., “John”, “lives in Rome”, “four-colouring”). An assertion hypothesises a relationship between constraints (e.g., “John lives in Rome”) that may or may not correspond to the actual relationship. Pieces of information are connected in an associative model (for example, Quine’s “field of force whose boundary conditions are experience” [13], the World Wide Web, or Kahneman’s “associative memory” [14] for people) with the nature of the connections determined by ecosystem conventions and constraints.

In [10], the author defines epistemological information for a human subject in terms of its ontological information: “the state in which the thing is and the way by which the state varies from others”. In this paper, “state” corresponds to properties of slices where, in any ecosystem, a property is measured using an ecosystem process. Also, an ecosystem may include non-human actors. Assertions are the mechanism for relating the properties of different slices and how they vary from each other. The relationship with other definitions of information is discussed in [2] but, as discussed above, the approach in this paper and [2] is based on the idea that information is relative and that it is the ecosystem conventions that determine how content is processed and linked to actions and outcomes.

The effect of selection pressures over time is to improve the ability of IEs and ecosystems to process information in a number of different ways. The quality of information may improve, in the sense that it is better able to support the achievement of favourable outcomes; it may be produced with

lower friction [15] or it may be produced faster [5]. These three measures, pace, friction and quality, are limited by ecosystem conventions.

Selection pressures ensure that information is generally reliable enough for the purposes of the ecosystem and within the envelope in which the selection pressures apply. However, quality issues and ecosystem limitations mean that outside the envelope we should not expect ecosystem conventions to be reliable. This is particularly important in an era of rapid change, such as the current digital revolution, in which ecosystem conventions cannot keep pace with the change.

### 3. Information and Inference

IEs make inferences to guide their actions and improve their chances of favourable outcomes. Inferences are based on information received as well as that stored in IE or ecosystem memory. Depending on ecosystem conventions, inference can use a variety of mechanisms including, for example, association, correlation, induction, deduction, machine learning, probability and the scientific method. This section analyses the relationship between information and inference using the model described in Section 2.

#### 3.1. Selection and Inference

The relationship between information and selection processes is examined in [1,2]. The ability of an IE to achieve favourable outcomes depends on how well it can respond to interactions with the environment. The question that IEs address, often implicitly, is: what is the action to take, if any, that has the best chance of a favourable outcome? Better quality information and relevant connections between different pieces and types of information improve the chances of favourable outcomes. So, it is clear that there is a selection advantage in understanding what can be inferred from descriptive information.

However, inference is subject to huge combinatorial problems and rigorous inference incurs very high friction (think, for example, of the cost of science). So we would expect that selection pressures would drive trade-offs between the measures and, indeed, this is what happens. For example, in [7], the authors make the case that many specialised inference mechanisms have evolved in people; they say: “inference, and cognition more generally, are achieved by a coalition of relatively autonomous modules that have evolved [ . . . ] to solve problems and exploit opportunities.”

In a similar way, organisations implement computer systems that automate many different routine activities in different ways but richer, more complex information processing requires people. In the field of machine learning, there are many different approaches [16] and one of the key tasks is to understand where to hand off decisions to a human being [17]. Selection processes select for conventions that embed these trade-offs in ecosystem processing and communication including, as described in [7,14], into the fundamental design of the IE.

#### 3.2. Reliability

For an IE, an inference is reliable if it helps achieve a favourable-enough outcome—the IE can rely on the inference under the ecosystem selection pressures.

So, how do ecosystems improve the reliability of inference? An element that many ecosystems have in common is that of challenge. Table 1 shows some examples.

**Table 1.** Challenge.

Ecosystem	Inference	Challenge
English criminal law (prosecution)	The defendant is guilty	The defence (plus, potentially, the appeals process)
Science	A prediction made by a hypothesis is true	Experiments to refute or confirm the prediction
Mathematics	A theorem is proved	Peer review
Computer systems	The system will perform as required	Tests that the system meets its requirements

The objective of each challenge is to identify weaknesses in the inference either in terms of its output (e.g., refutation in scientific experiments), the input assertions it is based on (e.g., the evidence in a trial) or the steps of the inference (e.g., peer review of Mathematics).

The generic mechanism is similar in each case. A related ecosystem has selection processes in which favourable outcomes correspond to successful challenges. The degree to which the challenge is rigorous depends on the selection pressures that apply to it and, in some cases, the degree to which a different IE from the one making the inference conducts it.

However, selection pressures can also diminish quality. “Conflict of interest” is a term used widely to capture the difficulties of such selection pressures. In [18], the author gives examples of the impact on science.

### 3.3. Inference Patterns

Different types of inference have a structure in common. In each, there is a set of input assertions and from different combinations of assertions further assertions can be inferred. As Quine says in [13] the set of input assertions to consider may be much larger than those under immediate consideration and include IE or ecosystem memory. In making any inference, there may be intermediate steps, so an overall inference will have the form of an acyclic, connected, directed graph, as shown in Figure 2 (although in many cases inference may have just a single step).

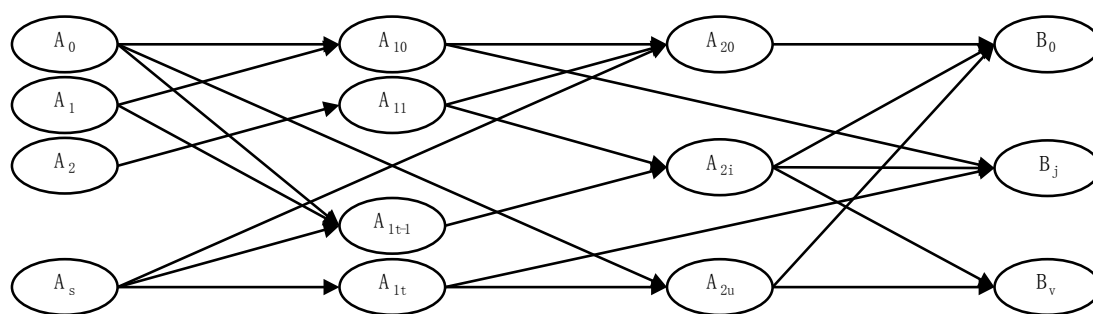


Figure 2. Inference instance.

There will also be a measure of quality associated with the derived assertions. When an IE makes an inference the result may add, delete, confirm or change those assertions and quality measures in the IE’s memory (and potentially in the ecosystem memory) not directly associated with the inputs.

What does each connection (based on a single inference) look like? Inference patterns and inference instances can be modelled using linnets [2,3]—a generalisation of graphs in which connectors (generalising edges) can connect other connectors as well as vertices. Figure 3 shows the general case.

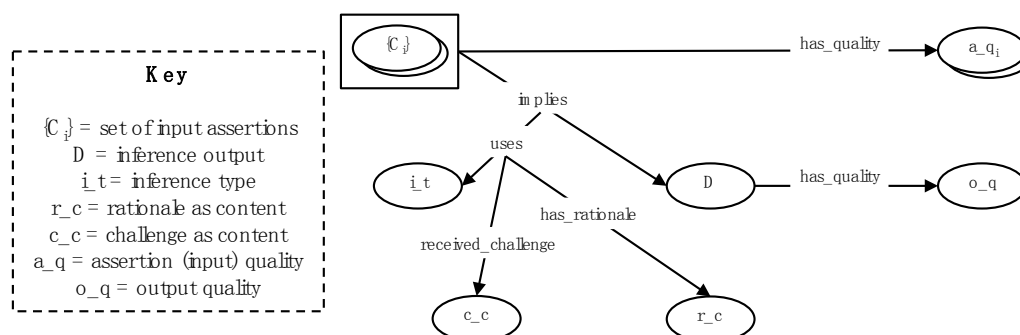


Figure 3. Single inference patterns.

As shown in Figure 3, apart from the initial assertions  $\{A_i\}$  any assertion (D) is derived from other assertions (for clarity, the figure shows one output—there may be more than one). Each input assertion has a measure of quality (that may or may not be captured). The inference uses some type of inference that determines how this step in the inference is accomplished. The overall rationale may be expressed as content and may be challenged independently of the inference process itself or embedded as part of the inference process. The following sections expand on these ideas.

### 3.4. Content and Event Inference

In [13], Quine discussed the prevailing “fundamental cleavage” between analytic and synthetic truth and, in the context of science discussed the “field of force whose boundary conditions are experience”. In [4], the “cleavage” is discussed in term of ecosystem conventions and their limitations. Some limitations are associated with the distinction between content interpretation and event interpretation that depends on the ecosystem-specific difference between content slices and event slices. Content slices are those that are interpreted as content within the ecosystem—all others are event slices. So, content slices include speech, animal calls, computer messages corresponding to particular protocols, documents and many others (each within its own ecosystem). IEs interpret information by connecting it with sets of other slices and so the question of which slices are considered in the search is highly pertinent. As it is expressed in [4]:

*“does the search include content only [ . . . ] or also events [ . . . ]? Call the former ‘content interpretation’ and the latter ‘event interpretation’. In both cases, the interpretation connects the [information] with sets of slices, but in the first case, the slices will be content slices only. For example:*

- *a Logician or Mathematician may treat the content of [information] as a set of symbols with rules (content interpretation);*
- *a computer query may just search for the string “John Smith” (content interpretation);*
- *a person may also consider memories of John Smith as a person (event interpretation);*
- *more generally, people may relate content to events that they remember (for example, relating people to holidays, parties, or other social events)—this is a major focus of modern social media (event interpretation).”*

We can apply the same approach to inference. Inferences using inference models that apply to content only we can call content inference. Event inference includes the rest.

### 3.5. Types of Inference

This section considers different types of inference and how they relate to the inference pattern shown in Figures 2 and 3. In each case, the inference can be expressed in terms of properties of slices where a property is measured by an ecosystem process. In terms of the discussion so far there are the following types:

- Content inference;
- Causation—in this case, the inference is based on one or more causation processes;
- Similarity—in this case, the inference is based on the similarity between sets of slices and the assumption that the similarity will extend.

Before we look at these, some terminology:

- $P(\sigma, t) = V$  means that the set of properties  $P$  of slice  $\sigma$  at time  $t$  has value  $V$  ( $P$  and  $V$  are tuples);
- $P(\sigma) = V$  means that  $P(\sigma, t) = V$  is true for all  $t$  under consideration;
- $P(\Sigma) = V$  means that  $P(\sigma) = V$  for all  $\sigma \in \Sigma$ .



V can be any content allowed by the ecosystem (so it can be a single value, a range or a more complex function, depending on the ecosystem).

Each of these statements may be considered either to be defining a chunk or an assertion. For example,  $P(\Sigma) = V$  defines a chunk where  $\Sigma$  is the set of all slices that meet this constraint. However, if  $\Sigma$  is defined elsewhere then it is an assertion.

Consider the following syllogism (E/C means that event or content interpretation are possible):

(E/C) Socrates is a man.  
 (E/C) All men are mortal.  
 Therefore, Socrates is mortal.

As discussed below, the difference between event and content interpretation is important. If we use event interpretation, this captures simple set containment based on the slices defined by the sets {Socrates}, {men} and {mortal entities}. In this case there is a direct relationship between the assertions and the sets.

In the more general case we have in which (C) indicates that this is content interpretation:

(C)  $P \rightarrow Q$ .  
 (C)  $Q \rightarrow R$ .  
 Therefore,  $P \rightarrow R$ .

This syllogism is an example of a general case that applies to modelling tools like Logic and Mathematics as well as any reasoning that only considers rules about symbols and content rather than the event slices they refer to. We can describe the general case using properties of slices. Suppose that:

- $C_i$  and  $D_j$  are assertions in an ecosystem for various  $i, j$ ;
- $\Delta_i$  is the set of slices representing  $C_i$  in the ecosystem;
- $\Phi_j$  is the set of slices representing  $D_j$  in the ecosystem;
- $R$  is ecosystem content that represents application of ecosystem rules to demonstrate that  $\{C_i\}$  implies  $\{D_j\}$ ;
- $\Psi$  is the set of slices that represent the content  $R$  in the ecosystem;
- $T$  is the ecosystem representation of truth;
- $E(\Omega) = T$  represents the application of an ecosystem process to measure the ecosystem truthfulness of  $\Omega$  and the process returns  $T$ .

With these definitions we can express the general case as:

(E/C)  $E(\Delta_i) = T$  for all  $i$ .  
 (C)  $E(\Psi) = T$ .  
 Therefore,  $E(\Phi_j) = T$  for all  $j$ .

This case includes deduction in both formal and less formal senses, depending on the ecosystem in which it is applied.

Consider causation and the following example from [19]:  $A =$  "The rain caused the wheat to grow". There are many assumed assertions here. In particular, it is assumed that there is a slice  $\sigma$  in which wheat is growing over a period of time; also that  $t'$  is a time period that is sufficient to demonstrate growth. So,  $A$  means the following: "It rains on  $\sigma$  at time  $t$  implies that the wheat will have grown by time  $t + t'$ " (in the absence of other factors).

In any example of causation there is an implied slice (like  $\sigma$ ), a starting time period,  $t$ , an elapsed time period,  $t'$ , and a physical process that operates from  $t$  to  $t + t'$ . As in the example above, none of these may be immediately apparent in the content and the slice of interest at time  $t + t'$  may be different from  $\sigma$  (call it  $\sigma'$ ) and we may be interested in a different set of properties of it. So, any example of causation takes the following form:

(E)  $P(\sigma, t) = V$ . (e.g., the height of the wheat)

(E) There is a slice  $\tau$  containing  $\sigma$  from time  $t$  that instantiates an ecosystem process pattern. (e.g., it rains)

Therefore,  $\exists \sigma' \subseteq \tau$  and time  $t'$  for which  $Q(\sigma', t + t') = U$ . (e.g., the wheat has grown).

The property values are established through ecosystem measurement processes and values are provided in ecosystem content. The property values  $\{(p_i, v_i)\}$  match the input assertions of the inference (but also see Section 4 for a discussion of the reliability of the match).

Inference associated with causation works forward (as in this case), but also backwards:

(E)  $P(\sigma, t) = V$ .

(E) There is a slice  $\tau$  containing  $\sigma$  from time  $t - t'$  to time  $t$  that instantiates an ecosystem process pattern.

Therefore,  $\exists \sigma' \subseteq \tau$  for which  $Q(\sigma', t - t') = U$ .

In this way, inference based on causation can be indirect—it can work forwards and backwards, through slices  $\sigma, \sigma', \dots, \sigma' \cdots'$ , and corresponding times to draw conclusions seemingly remote from the starting point.

In all of these cases, there is an underlying ecosystem pattern (perhaps based on science at one extreme or folklore at the other) that forms the basis for establishing inference. For example, causation is the basis for computer simulations for subjects as diverse as economics and climate change.

The degree to which causation-based inference makes sense in people-based ecosystems depends, in part, on the degree to which the type of causation accords with human psychology (like the example A above). Kahneman [14] points out the “innate biases” of System 1; these biases, and limitations in human senses, form the basis, for example, of magic tricks. In [20], the authors show how the tricks involved in magic are linked to neuroscience. A more extreme example is quantum mechanics and Einstein’s quote [21]:

*“Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the “old one.” I, at any rate, am convinced that He does not throw dice.”*

Correspondingly, causation that accords with human psychology is more likely to be adopted in human ecosystem conventions than causation, which does not. Or, expressed differently, human psychology can apply selection pressures.

Now for similarity. In this case, inference is based on the similarity between sets of slices with the assumption that the similarity will extend, assuming the “dependable regularities” discussed in [7]. We can express a pattern for similarity-based inference as follows (for sets of slices  $\Sigma_1, \Sigma_2$ ):

(E)  $P(\Sigma_1 \cup \Sigma_2) = V$ .

(E)  $Q(\Sigma_1) = U$ .

Therefore,  $Q(\Sigma_2 \cup \Sigma_1) = U$ .

A stronger form has the following pattern:

(E)  $P(\Sigma_1 \cup \Sigma_2) = V$ .

(E)  $Q(\Sigma_1) = U$ .

(E) For all  $\sigma$  for which  $Q$  has been measured, if  $Q(\sigma) \neq U$ , then  $P(\sigma) \neq V$ . (i.e., there are no known counter-examples)

Therefore,  $Q(\Sigma_2 \cup \Sigma_1) = U$ .

We can compare this pattern with the taxonomy of inductive inference created by Carnap [22]. Table 2 relates them in terms of various specialisations of the “similarity” pattern.



**Table 2.** Carnap’s taxonomy of inductive inference.

Variety of Inductive Inference	Specialisation of the “Similarity” Inference Pattern
Direct inference	$\Sigma_1 \subset \Sigma_2$ , Q measures the frequency of $P = V$ . $Q(\Sigma_2) = U$ . Therefore, $Q(\Sigma_1) = U$ .
Predictive inference	$\Sigma_1 \cap \Sigma_2 = \emptyset$ . $P(\Sigma_1) = P(\Sigma_2) = V$ . $Q(\Sigma_1) = U$ . Therefore, $Q(\Sigma_2) = U$ .
Singular predictive inference	As for predictive inference but, in addition, $ \Sigma_2  = 1$ .
Inference by analogy	$ \Sigma_1  =  \Sigma_2  = 1$ . $\Sigma_1 \neq \Sigma_2$ . The rest follows as per the general similarity case.
Inverse inference	$\Sigma_1 \subset \Sigma_2$ . $Q(\Sigma_1) = U$ . Therefore, $Q(\Sigma_2) = U$ .
Universal inference	As for inverse inference but with the stronger version of similarity.

Each of the examples above is based on properties of slices including the existence of slices with particular properties. The input assertions correspond to relationships between properties of slices and so do output assertions. So, each form of inference has the following form:

$P_i(\Sigma_i) = V_i$ , for sets of slices  $\Sigma_i$ .

Therefore  $Q_j(\Phi_j) = U_j$ , for sets of slices  $\Phi_j$ .

The reliability of inference with respect to the physical world is based on the level of confidence that the model applies correctly to the given inputs (and that the known regularity described in [7] can be extended). This is the traditional problem of induction [23]. The uncertainties apply at two levels: those that result from the particular inference instance not meeting the ecosystem conventions and those that result from the ecosystem conventions themselves. In the former case there is a reference point; the ecosystem determines what quality means in terms of inference is like and we can use an approach to analyse the quality like the approach for descriptive information described in [2,4]. In the latter case we can use an equivalent of viewpoint analysis, as described in [4], to highlight the limitations of the ecosystem conventions.

#### 4. Inference Reliability

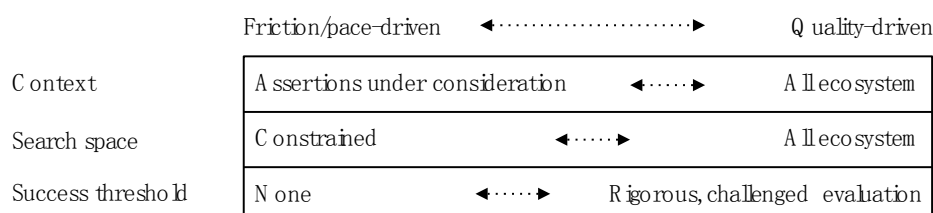
In this section we use the ideas developed in Section 3 to analyse the reliability of inference using a similar approach to that followed in [4]. First, we need to understand more about the processes that IEs and ecosystems undertake to draw inferences, since these processes are the source of the limitations. Subsequent sections consider the quality of inference as measured against ecosystem standards and then the limitations of the ecosystem conventions themselves.

##### 4.1. Inference Strategy

Ecosystem conventions define norms for making inferences. However, to make an inference an IE has to (implicitly or explicitly) decide how to instantiate the pattern in Figures 2 and 3. For example, as the authors put it in [7], “humans are like other animals: instead of one general inferential ability, they use a wide variety of specialized mechanisms”.

This shows the importance of inference strategy—the approach taken to inference. Inference strategy addresses the factors shown in Figure 4 based on the corresponding figure in [3]. Inference strategy is constrained at two levels: the ecosystem conventions (some of which are embedded in the

basic capabilities of the IE) and the circumstances in which the IE finds itself (including, for example, time or resource constraints).



**Figure 4.** Inference strategy factors.

The context factor is driven by the set of inputs ( $\{A_i\}$  in Figure 3) that are taken into account—how much should be included from memory as well as the specific assertions under consideration? At one extreme, only the assertions in question would be included; at the other, all of ecosystem memory as well.

The search space factor determines the rigour with which the inference pattern is sought—how much IE and ecosystem memory is searched to find the inference pattern(s). The success threshold option considers how the search for a model completes: is there a rigorous evaluation incorporating multiple options?

#### 4.2. Inference Quality

The quality of descriptive information is discussed in [2,3]. With respect to inference, the measures relate to the components shown in Figures 2 and 3: the input assertions; the structure of the inference instance; the quality of the challenge; interim and output assertions. The input assertions can be distinguished from the inference in the same way as unconditional and conditional probability.

Consider the reference point provided by the ecosystem. This will be an ecosystem inference instance that conforms to an ecosystem inference pattern. An example of such a pattern is the UK government 5-case model for making business cases [24]—business cases make an inference that with a particular starting point specified actions undertaken by an organisation will produce specified outcomes. Other examples are the rules that apply to mathematical or logical deductions. The ecosystem inference pattern (and there may be more than one for an ecosystem) expresses the convention in the ecosystem for making inferences of the level of quality expected.

The structure of both the ecosystem inference instance and the inference instance in question can be expressed as linnets, extensions of graphs described in Section 3.3. This allows a comparison of:

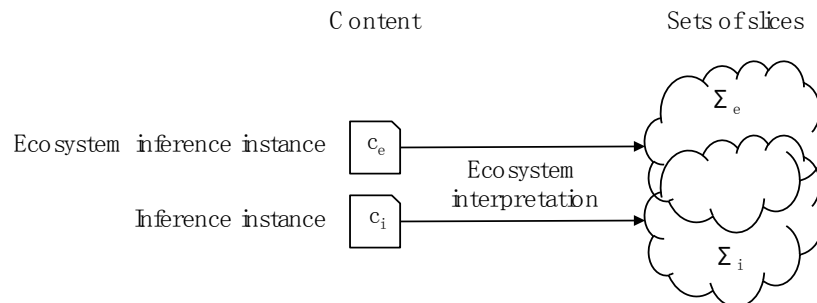
1. the completeness of the assertion vertices (i.e., including input assertions, interim assertions and output assertions) and the absence of extraneous vertices (Figure 2);
2. the completeness of the connections and absence of extraneous connections (Figure 2);
3. the inference model used for each connection (Figure 3).

Items 1 and 2 correspond to calculations about sub-linnets and so represent a partial order under set inclusion. Item 3 may represent just a comparison or there may be a richer relationship between inference models in the ecosystem. Again, in general, this is a partial order.

Input assertions are covered by the quality measures discussed in [2,3]. The various quality measures discussed are summarised in a single measure called trustworthiness. Trustworthiness combines the quality of the chunks forming an assertion and the plausibility of the assertion itself—the degree to which the actual relationship between the chunks is the same as that hypothesised by the assertion.

Does the same apply to interim and output assertions? The idea of plausibility transfers directly to inference so consider the quality of chunks. Interim and output assertions are created by the inference

and so we are interested in the following question: to what extent does the inference instance match the ecosystem inference instance? The content of the two instances could be expressed very differently, so a direct analysis of the content may not be helpful. Instead, using ideas from [2,3], we can measure the differences between the interpretations of the content and, to do this in a standard way, we can use the ecosystem interpretation. Figure 5 shows the comparison.



**Figure 5.** The quality of interim and output assertions.

Consider accuracy first. The accuracy of descriptive content addresses the relationship between ecosystem and IE interpretations of the same content. Here, we are concerned with the ecosystem interpretation of different pieces of content. However, the calculation, in both cases, is based on set inclusion (of sets of slices) and we can apply it to  $\Sigma_e$  and  $\Sigma_i$  (see Figure 5). In this sense, we can measure the accuracy of the inference instance.

The precision of descriptive content applies to interpretations of the same content at different times or in different assertions (or parts of the same assertion). We can use the same concept.

The other measures described in [2,3] apply either to the content itself or to the nature of the IE's interpretation so they do not apply to the question we are addressing here.

So, accuracy and precision can be used to measure the quality of chunks in interim and output assertions. As before, we cannot guarantee that these form a total order. So we cannot (in general) assign numbers to the overall quality of an inference; nor can we guarantee to assess the relative quality of different inferences. The best we can say is that we can assess the relative quality of some pairs of inferences.

To consider the quality of the challenge we can look at the characterisation of inference discussed in Section 3.5. Each form of inference has the following form:

1.  $P_1(\Sigma_i) = V_i$ , for sets of slices  $\Sigma_i$ .
2. Therefore  $Q_j(\Phi_j) = U_j$ , for sets of slices  $\Phi_j$ .

The challenge takes the following form: do there exist the following sets of slices for which the input assertions are trustworthy but not the output assertions:

1.  $\{\sigma_i: \sigma_i \in \Sigma_i\}$
2.  $\{\phi_j: \phi_j \in \Phi_j\}$
3.  $P_1(\sigma_i) = V_i$
4.  $Q_j(\phi_j) \neq U_j$ .

The purpose of the challenge is to identify what the software industry calls test cases [25]—a set of inputs and outputs designed to cover the range of possibilities thoroughly enough to provide confidence in the context of the ecosystem. The quality of the challenge determines how effectively the ecosystem model has been applied.

When there is high friction we should expect selection processes to find a trade-off. In practice, it is too difficult for IEs to assess quality so they often use proxies and the provenance of an inference

(linked to authority, brand, reputation, conformance to a data model or other characteristics) provides a shortcut. In [26], the authors express this idea elegantly with respect to documents:

*“For information has trouble, as we all do, testifying on its own behalf... Piling up information from the same source doesn’t increase reliability. In general, people look beyond information to triangulate reliability.”*

#### 4.3. Inference Limitations

Ecosystem conventions apply to each element in Figures 2 and 3 and the conventions are derived from trade-offs between quality, friction and pace and so we should expect there to be limitations in the conventions. In this section, we examine the following limitations:

- Ecosystem inertia: this applies when an ecosystem inference pattern has not kept up to date with the changes in selection pressures;
- Selection conflict: this applies when there are selection pressures that diminish the quality of inference (and ecosystem inertia may be a source of selection conflict);
- Inference tangling: this applies when different ecosystem conventions apply to different parts of the inference instance; an inference that is assumed to have been derived using a set of ecosystem conventions has not been;
- Output collapse: this applies when only one inference output is produced rather than a range;
- Inference strategy: this applies when the inference strategy is limited in its scope as described in Section 4.1.

These are discussed in turn below.

Ecosystem inertia is described in [4] and, with respect to digital change, in [1]. Ecosystem conventions take time to develop and may not keep up with the rate of change in the environment. In this case, the conventions that apply may not suit the changed environment and reduce the IE’s chances of favourable outcomes. Kuhn’s discussion of paradigm shifts in science [27] is an example.

In organisations, “change resistance” is a well-known concept. For example [28]:

*“One of the most baffling and recalcitrant of the problems which business executives face is employee resistance to change.”*

Or, in the social arena, some people find it difficult to keep up with changing digital technology, causing a “digital divide” [29] as they lose access to digital services.

One of the drivers for ecosystem inertia in human disciplines is the proliferation of knowledge. In [5,6], the authors discuss this and suggest that “all major research areas have to be surrounded by a ‘cloud’ of disciplines in order to convey the necessary [ . . . ] knowledge”. With respect to information science itself, the authors suggest a move from “self-centred inward looking”, driven by ecosystem conventions, towards “outward seeking”.

Selection conflict occurs when there are selection pressures that reduce the quality of inference. Goldacre [18] provides examples in science. The term “conflict of interest” applies more widely and covers many such selection conflicts including, for example, applications in the law and politics. The extent to which an ecosystem has mechanisms to detect the presence and influence of selection conflict has a corresponding impact on reliability.

Inference tangling occurs when different ecosystem conventions apply to different parts of the inference. An example of inference tangling is mentioned in Section 3.5—there are examples of content inference (like the “Socrates” example mentioned) for which content or event inference is possible. Formal systems apply to symbols but whenever sentences are written in English (or another language) then event interpretations are also possible, potentially leading to difficulties (especially in the light of Kahneman’s “innate biases” [14]). Elements of an inference can be based on the informality and, often, poor information quality of language.

Inference tangling also occurs when content inference is associated with event interpretation—when a rigorous formal argument is applied, without a correspondingly rigorous event interpretation, to events in the physical world. Examples of this, with respect to well-known paradoxes, are discussed in [4].

Inference is about making connections between descriptive information and outcomes and in the most general case includes a range of different outputs (like, for example, a probability distribution). Although a particular inference may link to different outcomes with different levels of quality, in many cases the IE needs to act, so it needs to focus on one outcome. As a result, many ecosystem conventions collapse the range of outputs down to a single one resulting in output collapse.

The inference strategy captures the parameters of an inference instance. There are always constraints on the scope of the inference strategy (the combinatorial difficulties are formidable) and these constraints limit the quality of the inference. Any inference instance draws on the associative memory of the IE and ecosystem (for example, Quine’s “field of force whose boundary conditions are experience” [13], the World-wide Web, Kahneman’s “associative memory” [14]) to support the individual elements of the inference (see Figure 2). So, the impact of the limitations and quality may vary between parts of the inference and may also vary as the circumstances of the IE change. Outside the envelope of reliability, inference may be very sensitive to initial conditions and these may not be measured accurately enough to overcome the sensitivity. In addition, selection pressures are not static, as currently demonstrated by developments in digital technology [1]

Information suffers from extreme combinatorial issues that ecosystem conventions overcome at a cost. So, why should we expect reliability? Kahneman [14] indicates why: “System 1 is radically insensitive to both the quality and quantity of information that gives rise to impressions and intuitions”. Selection pressures have given us a level of confidence but that confidence can be misplaced.

Constraints on reliability are embedded in ecosystem conventions and may be further exacerbated by the individual circumstances of an IE (e.g., shortages of resources or time or the application of an inappropriate pattern to the inference).

Deductive logic gives a semblance of reliability but, again, at a cost. Apparent reliability is linked to content inference but only if it is divorced from physical events. Any attempt to apply content inference to physical events suffers from the limitations of event interpretation—the huge enterprise that is modern science testifies to the difficulty of this.

## 5. Examples of Inference

The approach presented in Sections 3 and 4 can be applied to different examples of inference. It can be used to analyse both individual examples of inference and general theories of inference and the extent to which they incorporate considerations of quality, ecosystem limitations and the need to consider all relevant selection processes. In this section, we consider some aspects of the philosophy of science, the philosophy of mathematics and machine learning.

### 5.1. Science

Different approaches to the philosophy of science examine different parts of the inference mechanism discussed in Section 3 and different subsets of the selection pressures that apply. Based on the approach in this paper we can ask the following questions:

- to what extent do theories consider the full selection environment and the limitations of ecosystem conventions?
- to what extent do theories consider the components of information quality discussed in Section 4.2?
- what does the selection-oriented view of information say about the structures of knowledge that persist over time?
- how can we interpret “truth” or “approximate truth” in relation to the theories?

The answers provide a mechanism to understand the scope of theories—what they take into account and what they leave out—and therefore understand their limitations. It also helps to demonstrate how the theories relate to each other. Let us consider selection pressures first.

Popper [30] focuses on the importance of refuting assertions through experiment, which forms part of the challenge shown in Figure 3. In his theory, experiments provide the selection processes and the number and type of experiments define the scope of the challenge. The focus on refutation highlights one of the selection pressures in science, but Popper does not consider other selection pressures and the ecosystem more widely (see [18] for examples in which additional selection pressures can subvert the challenge).

In [13], Quine touches on selection pressures and ecosystem inertia. As he puts it:

*“[ . . . ] it turns upon our vaguely pragmatic inclination to adjust one strand of the fabric of science rather than another in accommodating some particular recalcitrant experience. Conservatism figures in such choices, and so does the quest for simplicity.”*

The conservatism and quest for simplicity he discusses are selection pressures. Conservatism is an element of ecosystem inertia.

There are strong selection pressures in science to apply resources to refutation. However, Kuhn’s discussion of paradigm shifts in science [27] adds a wider range of selection pressures. As well as experimentation, Kuhn adds in wider societal pressures and it is the ecosystem inertia associated with these that leads to the nature of the paradigm shifts he describes. As Max Planck [31] (quoted in [27]) describes it:

*“A new scientific truth does not triumph by convincing its opponents . . . but rather because its opponents eventually die, and a new generation grows up that is familiar with it.”*

This is a more rounded view of the selection pressures than Popper provides but, again, it is limited. As Goldacre demonstrates in [18], there can be a further set of selection pressures associated with the nature of the funding, publication and review processes that lead to selection conflict and consequent reduction in quality. These three types of selection pressure, identified by Popper, Kuhn and Goldacre, show the nature of the selection conflict in science. However, they also demonstrate the complex relationships between different selection pressures of different types.

Now consider quality. Through the idea of degrees of corroboration, Popper touches on the trustworthiness of input assertions. However, he does not address the concept of quality measures in more detail, and nor does Kuhn. Quality measures are useful in determining the reliability of inference but also in determining the impact of what Quine calls “contrary experiences”. This is the problem of holistic underdetermination. In [32] this problem is expressed in the following way:

*“Holist underdetermination [ . . . ] arises whenever our inability to test hypotheses in isolation leaves us underdetermined in our response to a failed prediction or some other piece of disconfirming evidence. That is, because hypotheses have empirical implications or consequences only when conjoined with other hypotheses and/or background beliefs about the world, a failed prediction or falsified empirical consequence typically leaves open to us the possibility of blaming and abandoning one of these background beliefs and/or ‘auxiliary’ hypotheses rather than the hypothesis we set out to test in the first place.”*

Quine based his discussion of underdetermination on a connection-oriented view like that Kahneman [14]. Because of the interconnectedness, Quine argues that all input assertions may be challenged:

*“The totality of our so-called knowledge or beliefs [ . . . ] impinges on experience only along the edges. [ . . . ] the total field is so underdetermined by its boundary conditions, experience, that there is much latitude of choice as to what statements to reevaluate in the light of any single contrary experience.”*



Rudner [33] makes the following argument:

*“[S]ince no hypothesis is ever completely verified, in accepting a hypothesis the scientist must make the decision that the evidence is sufficiently strong or that the probability is sufficiently high to warrant the acceptance of the hypothesis.”*

The approach to inference in this paper is helpful in analysing this issue. Rudner’s use of the term “sufficiently” relates directly to quality (either of information or inference). We can therefore reframe the question of holistic underdetermination in terms of the relative quality of input and interim assertions and the inference itself (see Section 4.2 and [3]). Low quality will then become a guide to those assertions or elements of the inference to challenge more rigorously.

Other limitations may also play a role. Quality is often assessed by proxy within an ecosystem (see Section 4.3) and these proxy assessments may contribute to ecosystem inertia and the difficulty of Kuhn’s paradigm shifts. Relative assessments, as Rudner suggests, are the best we can expect, since quality measures do not generally support a total order.

We can also consider the relationship with structures of information and knowledge. When linked with an analysis of selection pressures, the connection-oriented approach to information has implications for the likelihood of structures to persist over time. The model for information presented in this paper embeds different types of structure in complicated sets of connections that we can model using linnets [2,3]. Selection pressures trade off friction against quality, so highly connected structures that are difficult to change (i.e., where the change incurs very high friction) are likely to persist—this is an aspect of ecosystem inertia. So, we should expect some structures to be preserved under selection pressures. Whether we can anticipate in advance which structures will be preserved is another question.

Structural realism, as discussed by Worrall in [34], provides an example. Worrall says:

*“The rule in the history of physics seems to be that whenever a theory replaces a predecessor, which has however itself enjoyed genuine predictive success, the ‘correspondence principle’ applies. This requires the mathematical equations of the old theory to re-emerge as limiting cases of [ . . . ] the new. [ . . . ] The principle operates, not just as after-the-event requirement on a new theory [ . . . ] but often also as a heuristic tool in the actual development of the new theory.”*

That an existing structure will operate as a heuristic tool, which we can describe as a pattern with respect to connection-oriented information, is what we would expect from ecosystem conventions. We can also expect that successful theories will be very highly connected (to information about experiments as well as to other theories) and that there will be very high friction associated with dismantling these connections.

This idea is developed in [5,6] in which the authors discuss scientomics, which they describe as “a genuine informational approach to science” with an evolutionary approach to the development of knowledge in which knowledge recombination between disciplines plays a major role. In this sense, scientomics is explicitly about connecting different ecosystems, which naturally incurs high friction, the possibility of inference tangling and, implicitly, about overcoming the limitations of ecosystem conventions. As discussed in [5], “scientomics posits an inner structure of major recombination events along science history”.

Naturally, theories of the philosophy of science are concerned with truth. However, Worrall discusses one of the underlying difficulties when he addresses approximate truth, which is similar to the concept of truthlikeness discussed in [35]. He asks the question: “what exactly is involved in approximate truth?” but does not provide a detailed analysis. A rigorous analysis requires a more detailed understanding of this concept and its relationship with the quality of information and inference. This analysis is provided in [3] in which truth, truthlikeness and other concepts are analysed as measures of information. As discussed in [3], truth is not a rich enough measure of information to capture the comparisons suggested by Quine and Rudner. The idea of trustworthiness, which brings

together different measures of quality and plausibility (see Section 4.2), provides a more useful tool to analyse “approximate truth” and its relationship with quality. In this way, using the inference pattern in Figures 2 and 3 integrates the notion of “approximate truth”, via quality and truthlikeness, into the structure of any inference.

In summary, a limited view of selection pressures handicaps any theory. Without an understanding of the selection pressures a full understanding of the ecosystem conventions in play is not possible. In turn, this means both that the impact of the conventions on quality and the nature of the limitations will not be clear. Similarly, theories that rely on over-simplistic measures of information do not have the richness to handle concepts like quality, “degrees of corroboration” and “approximate truth”.

Each of the theories discussed focuses on a subset of the overall inference pattern. This highlights their limitations but also their position in the inference pattern highlights the relationships between them.

### 5.2. Mathematics

We can also apply the ideas above to mathematics. Mathematics straddles the worlds of content inference and event inference; the nature of inference is very different in these two cases.

Popper [36] discussed a related point when he suggested that a statement can be logically true but also falsifiable as a statement about the world. His example was “2 apples + 2 apples = 4 apples”. The first interpretation (“logically true”) relies on content interpretation and inference using mathematics. By contrast, the second interpretation relies on event interpretation and inference. The first uses mathematics as the modelling tool and the second uses both language and mathematics. For his example, it is easy to see that the interpretations coincide fairly readily. However if we consider the example “2 piles + 2 piles = 4 piles” (included in [4]) then it is more difficult. The second form of interpretation depends on what “+” means since it is also possible that “2 piles + 2 piles = 1 pile” (or other numbers).

Content ecosystems for mathematics deal with mathematics as a formal system. In this case, inference consists of proofs of conjectures or the demonstration of counter-examples. These can be long-term and ecosystem-wide as in the case of Fermat’s Last Theorem before it was proved by Wiles [37] or within a single IE for a short time (perhaps for individual elements of a proof as discussed by Villani in [38]). This is consistent with Curry’s formalist view [39].

Some ecosystems deal with applications of mathematics, for example in applied mathematics or science. In this case the inference to be challenged is an event inference of the form: there is a trustworthy relationship between a hypothesis expressed in mathematics and a set of slices. This form of inference is very different from content inference and the relationship between the two is complicated by the use of language in mathematics. Human language is used in mathematics to make it understandable. Wittgenstein referred to this idea in [40,41]:

*“[I]t is essential to mathematics that its signs are also employed in mufti”*

*“[i]t is the use outside mathematics, and so the meaning [‘Bedeutung’] of the signs, that makes the sign-game into mathematics”*

This means that there is considerable scope for inference tangling (see Section 4.3) and so we cannot necessarily be sure of the exact nature of the inference process.

### 5.3. Machine Learning

As the name implies, selection is built into concepts of machine learning. There are several different approaches to machine learning that Domingos [16] summarises as “symbolologists, connectionists, evolutionaries, Bayesians, and analogizers”. As he says:

*“Symbolists view learning as the inverse of deduction and take ideas from philosophy, psychology, and logic. Connectionists reverse engineer the brain and are inspired by neuroscience and physics.*

*Evolutionaries simulate evolution on the computer and draw on genetics and evolutionary biology. Bayesians believe learning is a form of probabilistic inference and have their roots in statistics. Analogizers learn by extrapolating from similarity judgments and are influenced by psychology and mathematical optimization."*

The tribes form ecosystems: "each tribe has a set of core beliefs, and a particular problem that it cares most about". Each has its own conventions.

Machine learning uses data. As Domingos says: "[machine] learners turn data into algorithms". They assume what he calls the unwritten rule of machine learning: "whatever is true of everything we've seen is true of everything in the universe" and they "induce the most widely applicable rules [ . . . ] and reduce their scope only when the data forces [them] to". This point of view pushes the "dependable regularities" discussed in [7] to their limits and emphasises the importance of challenge.

Machine learning quality has been analysed in depth. Machine learning finds patterns in data and the patterns depend both on the data and the type of pattern. It is prone to overfitting—"hallucinating patterns that aren't really there" [16]. At the extreme, machine learning will find patterns that fit the learning data but no others, so effective machine learning relies extensively on testing with data that has not been used for learning.

This demonstrates the main selection pressure they experience—what they infer is defined by the data used to define and test them. However, some types of machine learning, for example using neural networks, make the rationale for an inference opaque which exposes a different type of selection pressure. As Bostrom concludes [17]:

*"It will become increasingly important to develop AI algorithms that are not just powerful and scalable, but also transparent to inspection—to name one of many socially important properties."*

*"It is also important that AI algorithms taking over social functions be predictable to those they govern."*

These points demonstrate that, although machine learning is explicitly exposed to some selection pressures (by training and testing on selected inputs), these selection pressures only include a narrow assessment of function and quality. Wider issues, such as those presented by Bostrom are not necessarily yet part of the learning. As he says:

*" . . . it will require an AGI [Artificial General Intelligence] that thinks like a human engineer concerned about ethics, not just a simple product of ethical engineering."*

There is an analogy with Popper [30] and Kuhn [27]. Machine learning is currently addressing a narrow set of selection pressures (by analogy with Popper's description of refutation). Only when the machine learning ecosystem can also incorporate societal selection pressures (by analogy with Kuhn) will the wider questions raised by Bostrom be addressed.

## 6. Conclusions

The model for information presented here provides a set of tools with which to analyse information. These tools (about selection, ecosystem conventions and measures of information) highlight the limitations of inference.

Under the influence of selection processes, ecosystem conventions develop that establish the norms for inference. These conventions will provide inference that is reliable enough (but possibly not more) within the envelope of environment states determined by the selection pressures. The corollary of this is that, generally, an approach to inference is reliable only to the extent that all relevant selection pressures have challenged it thoroughly. And it is reliable only within the envelope in which selection pressures have applied.

If we use the pattern for inference in Figures 2 and 3, we can analyse different forms of inference in different ecosystems. This highlights the degree to which ideas in the philosophy of science,

mathematics and machine learning, for example, match up to the generic pattern and, therefore, where their limitations may be found. It also shows how different, seemingly competing, ideas about inference in the philosophy of science relate to each other within the general model.

**Acknowledgments:** The author would like to thank the referees for their constructive comments.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Walton, P. Digital information and value. *Information* **2015**, *6*, 733–749. [CrossRef]
- Walton, P. A Model for Information. *Information* **2014**, *5*, 479–507. [CrossRef]
- Walton, P. Measures of information. *Information* **2015**, *6*, 23–48. [CrossRef]
- Walton, P. Information and Meaning. *Information* **2016**, *7*, 41. [CrossRef]
- Moral, R.; González, M.; Navarro, J.; Marijuán, P.C. From Genomics to Scientomics: Expanding the Bioinformation Paradigm. *Information* **2011**, *2*, 651–671. [CrossRef]
- Marijuán, P.C.; Moral, R.; Navarro, J. Scientomics: An emergent perspective in knowledge organization. *Knowl. Organ.* **2012**, *39*, 153–164.
- Mercier, H.; Sperber, D. *The Enigma of Reason*; Harvard University Press: Cambridge, MA, USA, 2017.
- Lorenz, E.N. Deterministic Nonperiodic Flow. *J. Atmos. Sci.* **1963**, *20*, 130–141. [CrossRef]
- Arthur, B.W. *The Nature of Technology: What It Is and How It Evolves*; The Free Press and Penguin Books: New York, NY, USA, 2009.
- Zhong, Y. Unity-Based Diversity: System Approach to Defining Information. *Information* **2011**, *2*, 406–416. [CrossRef]
- Burgin, M. *Theory of Information: Fundamentality, Diversity and Unification*; World Scientific Publishing: Singapore, 2010.
- Logan, R. What Is Information?: Why Is It Relativistic and What Is Its Relationship to Materiality, Meaning and Organization. *Information* **2012**, *3*, 68–91. [CrossRef]
- Quine, W.V.O. “Two Dogmas of Empiricism”, Reprinted in *from a Logical Point of View*, 2nd ed.; Harvard University Press: Cambridge, MA, USA, 1951; pp. 20–46.
- Kahneman, D. *Thinking, Fast and Slow*; Macmillan: London, UK, 2011.
- Gates, B.; Myhrvold, N.; Rinearson, P. *The Road Ahead*; Viking Penguin: New York, NY, USA, 1995.
- Domingos, P. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*; Penguin: London, UK, 2015.
- Bostrom, N. *The Ethics of Artificial Intelligence (PDF)*; Cambridge University Press: Cambridge, MA, USA, 2011.
- Goldacre, B. *Bad Science*; Harper Perennial: London, UK, 2009.
- Simon, H.A.; Rescher, N. Cause and Counterfactual. *Philos. Sci.* **1966**, *33*, 323–340. [CrossRef]
- Macknik, S.L.; Martinez-Conde, S. *Sleights of Mind*; Picador: Surrey, UK, 2011.
- Einstein, A. Letter to Max Born (4 December 1926); *The Born-Einstein Letters (Translated by Irene Born)*; Walker and Company: New York, NY, USA, 1971.
- Carnap, R. *Logical Foundations of Probability*, 2nd ed.; University of Chicago Press: Chicago, IL, USA, 1962.
- Vickers, J. *The Problem of Induction*; Zalta, E.N., Ed.; Spring 2016 Edition; The Stanford Encyclopedia of Philosophy: Stanford, CA, USA; Available online: <https://plato.stanford.edu/archives/spr2016/entries/induction-problem/> (accessed on 25 May 2017).
- The Green Book: Appraisal and Evaluation in Central Government. Available online: <https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government> (accessed on 25 May 2017).
- Sommerville, I. *Software Engineering*; Addison-Wesley: Harlow, UK, 2010.
- Brown, J.S.; Duguid, P. *The Social Life of Information*; Harvard Business Press: Boston, MA, USA, 2000.
- Kuhn, T.S. *The Structure of Scientific Revolutions*, Enlarged 2nd ed.; University of Chicago Press: Chicago, IL, USA, 1970.
- Lawrence, P.R. How to Deal with Resistance to Change. *Harv. Bus. Rev.* **1954**, *32*, 49–57. [CrossRef]
- Norris, P. *Digital Divide: Civic Engagement, Information Poverty and the Internet Worldwide*; Cambridge University Press: New York, NY, USA, 2001.

30. Popper, K.R. *The Logic of Scientific Discovery*; Hutchinson: London, UK, 1959.
31. Planck, M. *Scientific Autobiography and Other Papers*; Gaynor, F., Translator; Williams & Northgate: London, UK, 1950.
32. Stanford, K. *Underdetermination of Scientific Theory*; Zalta, E.N., Ed.; The Stanford Encyclopedia of Philosophy: Stanford, CA, USA, Spring 2016 Edition. Available online: <https://plato.stanford.edu/archives/spr2016/entries/scientific-underdetermination/> (accessed on 25 May 2017).
33. Rudner, R. The Scientist qua Scientist Makes Value Judgments. *Philos. Sci.* **1953**, *20*, 1–6. [[CrossRef](#)]
34. Worrall, J. Structural Realism: The Best of Both Worlds? *Dialectica* **1989**, *43*, 99–124. [[CrossRef](#)]
35. Floridi, L. *The Philosophy of Information*; Oxford University Press: Oxford, UK, 2011.
36. Ryle, G.; Lewy, C.; Popper, K.R. Symposium: Why are the calculuses of logic and arithmetic applicable to reality? In *Proceedings of the Logic and Reality, Symposia Read at the Joint Session of the Aristotelian Society and the Mind Association, Manchester, UK, 5–7 July 1946*; pp. 20–60.
37. Wiles, A. Modular elliptic curves and Fermat's Last Theorem. *Ann. Math.* **1995**, *142*, 443–551. [[CrossRef](#)]
38. Villani, C. *The Birth of a Theorem: A Mathematical Adventure*; The Bodley Head: London, UK, 2015.
39. Curry, H. *Outlines of a Formalist Philosophy of Mathematics*; North-Holland: Amsterdam, The Netherlands, 1958.
40. Wittgenstein, L. *Remarks on the Foundations of Mathematics*, Revised Edition; Anscombe, G.E.M., Translator; von Wright, G.H., Rhees, R., Anscombe, G.E.M., Eds.; Basil Blackwell: Oxford, UK, 1978.
41. Wittgenstein, L. *Wittgenstein's Lectures on the Foundations of Mathematics*; Diamond, C., Ed.; Cornell University Press: Ithaca, NY, USA, 1976.



© 2017 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) license (<http://creativecommons.org/licenses/by/4.0/>).