A MODEL OF KNOWLEDGE BY RESEMBLANCE

Suppose we find ourselves in the thorny situation of having to explain the logical **AND** operator to a person new to logic—let's call her Francesca. As a first attempt, we propose to her an explanation about how to use this concept:

> "The logical AND operator connects two statements and results in a true statement if both statements are true and in a false statement otherwise."

However, Francesca then gaslights us, requesting an explanation that does not rely on logic or algebra in the first place. It seems that our limits in explaining the meaning of the AND operator to her coincide with the limits of her prior knowledge, as we realize that our explanation implicitly required some knowledge of logic itself. Therefore, it seems that our best shot could be to try to explain the meaning of AND by example, and thus we write on the blackboard the usual truth table.

CASES	**x1**	**x2**	**x1 AND x2 - the result**
case 1	T	T	T
case 2	T	F	F
case 3	F	T	F
case 4	F	F	F

Yet, as we begin writing, Francesca warns us that a mere enumeration of facts does not constitute an explanation and insists that something in common must be found between the examples to account for its knowledge.

If Francesca understands the idea of 'something in common', it must be the case that, like us, she is entitled to tell whether two material objects or symbols are the same or not. Getting smart with this knowledge, we then craft our explanation by relying solely on such comparisons.

> Cases 1 and 4 are similar because, in them, **x1**, **x2**, and the **result** all resemble one another. This means:
>
> **i)** Whenever x1 and x2 resemble each other, the result is whatever the value of either x is.
>
> Cases 2 and 3 are also similar:

ii) Whenever x1 and x2 do not resemble each other, the result is the constant symbol **false**.

At this point, Francesca is satisfied with the explanation and she demonstrates to us her understanding by showing the following cases:

CASES	x1	x2	x1 AND x2 - the result
case 5	Luca	Francesca	F
case 6	Francesca	Francesca	Francesca

These cases seem odd. We did not expect to apply the **AND** operator to natural language words—what could that even mean? Francesca replies that, in fact, the logical operator **AND** does not seem so different from some uses of the connector and in her own language. She explains that the statement

"Luca and Francesca ARE friends"

makes intuitive sense because both 'Luca' and 'Francesca' are persons and thus resemble each other. Furthermore, because we understand that persons can be friends, the connector and places 'Luca' and 'Francesca' within the same relationship of friendship. By contrast, the statement

"Luca and a stone ARE friends"

does not seem intuitively right. A stone is an object that cannot reciprocate friendship, so in the absence of additional information, the natural conclusion is that the statement lacks reasonable meaning.

Finally, she concludes that our explanation seems to suggest that the implicit meaning of 'F' in our explanation denotes the impossibility of composition.

The objective of this article is to propose an account of knowledge derived, as in the motivating example, from the perception of whether things are similar or not, and from the inferences drawn on that basis.

The problem with knowledge derived in this way is that we have no guarantee future observations will confirm our beliefs. At most, we can ensure our knowledge is consistent with what has been observed so far. This, however, fails to explain what happens when new information contradicts what was previously believed. If we allow knowledge to be simply contradicted by observation, the very notion of understanding collapses. Therefore, a model of knowledge grounded in observation must explain the nature of contradictory information in terms of what can be learned from such exceptional cases.

Consider a decision-maker who has learned from experience that in situation X her best choice is alternative A. When she later encounters a problem Y resembling X, she will then prefer alternative A, expecting good outcomes.

Now suppose that in Y, despite the resemblance, alternative A leads to a bad outcome. What should she conclude? Because her past experience has proven valuable so far, she may reasonably conclude that there must be some exceptional circumstance that makes Y different from X, validating the decision rule in all other cases. We call this the principle of the exception, the old legal principle that an exception confirms the rule as valid in all other cases. Hence, by excluding the exceptional circumstance from the belief responsible for the decision, she preserves her previous knowledge while assuming that this exceptional circumstance could be the cause of the bad outcome itself. This revision mechanism keeps her beliefs rational while proposing a solution to the problem advanced above.

In this work, I propose a model of knowledge in which beliefs are bold conjectures about the unintended consequences of resemblance and contradictions are maximally informative events that can be understood using the principle of the exception. I propose a philosophical experiment: an epistemic machine whose goal is to learn the meaning of the unknown program generating her experience. By formulating bold conjectures and revising her thinking in light of contradictory information, the machine exemplifies the problem of the use of knowledge in society—how to make sense of dispersed, potentially contradictory bits of information and create meaning from them.

The machine will think in the Infant Language, a minimal language consisting of three means, Resemblance, Diversity and Conclusion, mimicking the abilities of an Infant 3-12 months old who is capable of little inferences and distinguishing things. We will then impersonate an epistemic machine and challenge ourselves with the quest of learning an unknown program from radical ignorance. Thus, we will show that thinking in the infant language while using the principle of the exception to solve contradictory thinking will lead us to learn a partition of states, family of resemblances, characterized by the outcomes we expect to follow.

The main contribution of this work is to propose a path through the related ideas of information, knowledge, and the practical consequences of empiricism. This path could not be drawn without a proper understanding of information itself and its means of combination. For this reason, we adopt an algebraic approach in which the fundamental operation is an associative, commutative, and idempotent merge—an operation that adds information while preserving the core intuition that information cannot decrease.

From a philosophical perspective, this model proposes a computational synthesis of the ideas of natural kinds and family resemblances. The evolution of unreason into science is analogous to the epistemic machine's quest: learning the unknown program generating her observational data. Unreason becomes reason through the principle of the exception, which continually refines the scope and application of family resemblances derived from observations and their attendant expectations.

The remaining of this article is organized as follows. In Part 1 we introduce the building blocks of the epistemological model, in Part 2 we familiarize with the Infant Language, bold conjectures and the principle of the exception. In Part 3 we will impersonate the epistemic machine simulating the learning path of an unknown program. Discussions and conclusions will then follow.

1. Part 1: Information, Knowledge, Beliefs.

Information is knowledge of circumstances and facts upon which thinking is based, and we represent it by means of information structures. An information structure can either be complete or partial. Complete information structures are complete in the sense that their content cannot be incremented, whereas partial information structures can be incremented. The absence of information is denoted with the symbol nothing, which is a partial information structure that can always be incremented by adding some information to it. The idea of adding information is given by a binary merge operation that is **ASSOCIATIVE**, **COMMUTATIVE**, and **IDEMPOTENT**, defined as:

$$\oplus : I^2 \to I$$

where I is the set of possible information structures. The merge operation is generic in the sense that the meaning of its application depends on the kind of structures that are being added together. In the baseline case in which the information structures are complete, like numbers and symbols, the operation is given as

$$a \oplus b = \begin{cases} a & \text{if } b = \text{nothing or } a = b \\ \text{contradiction} & \text{otherwise.} \end{cases}$$

A contradiction is a primitive information structure whose meaning has to be interpreted as the impossibility of adding or composing different bits of information together. Thus, merge defines a bounded join-semilattice in which **nothing** is the bottom element whereas **contradiction** is maximal. This structure is similar to information lattices in domain theory, where ⊥ represents absence of information, with the addition of contradictions intended as maximally informative information structures.

Example

The number 1 is a complete information structure and, from the property of merge we have:

$$1 \oplus 1 = 1;$$
$$1 \oplus \text{nothing} = \text{nothing} \oplus 1 = 1;$$
$$1 \oplus 2 = 2 \oplus 1 = \text{contradiction}.$$

The reason why we cannot add the information structure 2 to the information structure 1 is because they are complete. Hence, if we know that "1" we cannot possibly know that "2", that would be a contradiction.

Consider now a closed Interval like $[1,3]$ and let us define define merge as the intuitive interval intersection operation so that

$$[1,3] \oplus [2,5] = [2,5] \oplus [1,3] = [2,3];$$
$$[1,3] \oplus 2 = 2 \oplus [1,3] = 2;$$
$$[1,3] \oplus [5,10] = [5,10] \oplus [1,3] = \text{contradiction.}$$

The ACI properties of merge reify the idea that the addition or composition of two bits of information cannot result in a less informative structure.

Because of the idempotency of the merge operation, we can define the idea of resemblance and diversity between information structures using merge as a primitive operator.

We define the **Resemblance** operator as:

$$R(s_i, s_j) = \begin{cases} y, & \text{if merge}(s_i, s_j) \neq \text{contradiction}, \\ n, & \text{otherwise.} \end{cases}$$

The meaning of resemblance is that two things resemble each other if they can be composed without contradiction. We use the symbols y, "yes" and n, "no" instead of "true" and "false" to remark the pre-logical nature of resemblance.

Similarly, we define the **Diversity** operator as:

$$D(s_i, s_j) = \begin{cases} y, & \text{if merge}(s_i, s_j) = \text{contradiction}, \\ n, & \text{otherwise.} \end{cases}$$

Lemma 1 (Properties of Resemblance and Diversity). The resemblance and diversity operators satisfy the following properties:

1. **Complementarity:** $R(s_i, s_j) = y \iff D(s_i, s_j) = n$
2. **Reflexivity:** $R(s, s) = y$ for any non-contradictory information structure s
3. **Symmetry:** $R(s_i, s_j) = R(s_j, s_i)$ and $D(s_i, s_j) = D(s_j, s_i)$
4. **Non-transitivity:** Neither R nor D is transitive

Proof: See Appendix A. □

Example (Failure of Transitivity). Consider three interval information structures:

$$s_1 = [1,3], \quad s_2 = [2,5], \quad s_3 = [4,7]$$

For resemblance:

- $\text{merge}(s_1, s_2) = [2,3] \neq \text{contradiction}$, thus $R(s_1, s_2) = y$

- $\text{merge}(s_2, s_3) = [4,5] \neq \text{contradiction}$, thus $R(s_2, s_3) = y$
- $\text{merge}(s_1, s_3) = \text{contradiction}$ (disjoint intervals), thus $R(s_1, s_3) = n$

This demonstrates that $R(s_1, s_2) = y$ and $R(s_2, s_3) = y$ does not imply $R(s_1, s_3) = y$.

For diversity, consider:

$$s_1 = [1,3], \quad s_2 = [5,7], \quad s_3 = [2,4]$$

Then $D(s_1, s_2) = y$ and $D(s_2, s_3) = y$, but $D(s_1, s_3) = n$ since $\text{merge}(s_1, s_3) = [2,3]$.

Lemma 1 states that Resemblance and Diversity are complementary, reflexive and symmetric, like $=$ and $\neq$, however, the non-transitivity of both operators reflects the context-dependent nature of similarity and difference in information structures, aligning with Wittgenstein's notion of family resemblance where objects may share overlapping but non-identical properties.

1.1. Experience, Signals, Beliefs

We assume that Nature is a machine $N(\cdot)$ generating the experience we observe. Experience is consequential in the sense that we distinguish between before and after and therefore between inputs and output. A signal is a particular information structure accumulating knowledge about something specific. We denote signals with s and a profile of signals $\mathbf{s} = (s_1, s_2, \ldots)$ with ω and we call it a state. A generic process $\pi(\omega) = o$ is a function that maps states to outcomes whose generic element is denoted with o. We say that an observation is a pair (ω, o) and we use Ω and O to denote the set of observed states and outcomes respectively. An Event E is a subset of Ω and beliefs, map events to outcomes

$$B_E \subseteq O.$$

We use the notation B_E to denote the outcomes believed to be possible when the event E is observed and we denote with $\pi(E) = \{\pi(\omega) \in O : w \in E \subseteq \Omega\}$ the set of possible outcomes that has been observed with E.

Given some Ω and O, two observations (ω, o) and (ω, o') are said to be ambiguous if the observation of the same state is associated to different outcomes. A belief B_E is ambiguous whenever it contains more than one element and it is contradictory whenever there exists an observation (ω', o') such that $\omega' \in E$ and $o' \notin B_E$.

The following assumption restrict the attention to deterministic programs suggesting instead that ambiguity in observations has to be attributed to the lack of complete information about the signals determining the output. Then, we will illustrate the meaning of the assumption in this context and the model by means of an example.

Assumption 1: Nature is deterministic.

Nature processes are **deterministic** which means that, for any state $\omega \in \Omega$, there exists one and only one outcome $o \in O$.

Assumption 1 means that processes cannot possibly be ambiguous. However, from the perspective of a subject experiencing the world, information may not come as whole that is already interpreted, on the contrary, what we may call knowledge may be dispersed through different processes and we may not be able to gather or perceive it its entirety as the following example demonstrates.

Example

In the remainder of this article we will often represent observations by means of tables as follows.

ω	s_1	s_2	$\pi(\omega = (s_1, s_2))$
ω_0	a	a	x
ω_1	b	d	x
ω_2	a	a	z

$$\Omega = \{\omega_0, \omega_1, \omega_2\} \quad O = \{x, z\}$$

The states ω_0 and ω_2 are ambiguous for the observation of the same signal is associated with more than one outcome. Hence, we can conjecture the existence of another signal s_3 which could explain why observations ω_0 and ω_1 differ.

ω	s_1	s_2	s_3	$\pi(\omega = (s_1, s_2, s_3))$
ω_0	a	a	h	x
ω_1	b	d	nothing	x
ω_2	a	a	l	z

The model developed so far provides the building blocks to discuss a notion of knwoledge and beliefs based on family resemblance. Hence, we need to understand what could be considered knowledge when all available information is limited to our experience.

We say that Knowledge is a minimally coherent set of beliefs that are:

- rational
- complete
- mutually exclusive

Whenever the previously listed condition hold, then Knowledge can be said to be rational. The main problem of this work is therefore finding this set rational beliefs, in particular, how to derive them just from resemblance, diversity and the fundamental idea of adding information. The solution we propose here is a mechanism for adding knowledge that requires two distinct features

1. the ability to generate bold conjectures, that is to say, beliefs that are more general than each of the observations they explain
2. the ability review contradictory thinking by using the principle of the exception, a belief revision procedure inspired by the legal principle that exceptions prove the rule valid in all other cases.

Imagine then an epistemic machine whose objective is that of understanding the natural process generating the data she observes. We will simulate such a machine against an unknown program and we will show that, in finite time, the machine will be able to learn the unknown program in such a partitional form, and we interpret those partitions as family of resemblances, effectively grouping similar states to outcomes. The machine will think in the infant language and will resolve contradictions using the principle of the exception which are the subjects of the following parts.

2. Part 2: The Infant Language

3. Appendix

3.1. Appendix A

Proof of Lemma 1

We prove each property in turn.

(1) Complementarity. By definition:

$$R(s_i, s_j) = y \iff \text{merge}(s_i, s_j) \neq \text{contradiction}$$

$$D(s_i, s_j) = y \iff \text{merge}(s_i, s_j) = \text{contradiction}$$

Since for any merge operation, either $\text{merge}(s_i, s_j) = \text{contradiction}$ or $\text{merge}(s_i, s_j) \neq \text{contradiction}$ (but not both), we have:

$$R(s_i, s_j) = y \iff D(s_i, s_j) = n$$

as required. □

(2) Reflexivity.
For any non-contradictory information structure s, by the idempotence property of merge (Section 2.1), we have:

$$\text{merge}(s, s) = s$$

Since s is not itself a contradiction, $\text{merge}(s, s) \neq \text{contradiction}$. Therefore, by definition of R:

$$R(s, s) = y$$

as required. $\square$

(3) Symmetry.
By the commutativity property of merge, we have:

$$\text{merge}(s_i, s_j) = \text{merge}(s_j, s_i)$$

Therefore:

$$\text{merge}(s_i, s_j) = \text{contradiction} \iff \text{merge}(s_j, s_i) = \text{contradiction}$$

From the definitions of R and D, this immediately implies:

$$R(s_i, s_j) = R(s_j, s_i) \quad \text{and} \quad D(s_i, s_j) = D(s_j, s_i)$$

as required. $\square$

(4) Non-transitivity.
We prove non-transitivity by providing explicit counterexamples.
<u>Non-transitivity of Resemblance:</u> Let $s_1 = [1, 3]$, $s_2 = [2, 5]$, and $s_3 = [4, 7]$ be interval information structures where merge is defined as interval intersection.

Computing the merges:

- $\text{merge}(s_1, s_2) = [1, 3] \cap [2, 5] = [2, 3] \neq \text{contradiction}$, so $R(s_1, s_2) = y$
- $\text{merge}(s_2, s_3) = [2, 5] \cap [4, 7] = [4, 5] \neq \text{contradiction}$, so $R(s_2, s_3) = y$
- $\text{merge}(s_1, s_3) = [1, 3] \cap [4, 7] = \emptyset = \text{contradiction}$, so $R(s_1, s_3) = n$

Thus, $R(s_1, s_2) = y$ and $R(s_2, s_3) = y$, but $R(s_1, s_3) = n$, demonstrating that resemblance is not transitive.

<u>Non-transitivity of Diversity:</u> Let $s_1 = [1, 3]$, $s_2 = [5, 7]$, and $s_3 = [2, 4]$ be interval information structures.

Computing the merges:

- $\text{merge}(s_1, s_2) = [1, 3] \cap [5, 7] = \emptyset = \text{contradiction}$, so $D(s_1, s_2) = y$
- $\text{merge}(s_2, s_3) = [5, 7] \cap [2, 4] = \emptyset = \text{contradiction}$, so $D(s_2, s_3) = y$
- $\text{merge}(s_1, s_3) = [1, 3] \cap [2, 4] = [2, 3] \neq \text{contradiction}$, so $D(s_1, s_3) = n$

Thus, $D(s_1, s_2) = y$ and $D(s_2, s_3) = y$, but $D(s_1, s_3) = n$, demonstrating that diversity is not transitive. $\square$