Representational formalisms: what they are and why we haven't had any

Lev Goldfarb

ETS group
Faculty of Computer Science, UNB
Fredericton, Canada

Abstract. Currently, the *only* discipline that has dealt with *scientific* representations—albeit non-structural ones—is mathematics (as distinct from logic). I suggest that it is this discipline, only vastly expanded based on a new, structural, foundation, that will also deal with structural representations.

Logic (including computability theory) is not concerned with the issues of various representations useful in natural sciences. Artificial intelligence was supposed to address these issues but has, in fact, hardly advanced them at all.

How do we, then, approach the development of representational formalisms? It appears that the only reasonable starting point is the primordial point at which *all* of mathematics began, i.e. we should start with the generalization of the process of construction of natural numbers, replacing the identical *structureless* units, out of which numbers are built, by structural ones, each signifying an atomic "transforming" event.

This paper is conceived as a companion to [1], and is a revised version of [2].

> *Mathematics is the science of the infinite*, its goal is the symbolic comprehension of the infinite with human, that is finite, means.
>
> Hermann Weyl

> The definition of mathematics accepted in the 20th century, as a science of the infinite, should be replaced by another one, which more accurately captures its nature: it is the science of the relationship between the finite and the infinite.
>
> N. N. Yanenko

1 Introduction: the situation in artificial intelligence

Why is the problem of representation so strategically critical? Mainly because, once we discover the "right" representational formalism, we will be able "to know the mind of god", to use a popular phrase coined by modern physicist Stephen Hawking, i.e. we will truly understand what the nature of intelligence in the Universe is. "The problem [of representation] persists because it is extraordinarily difficult, perhaps the most difficult one in all of science. It is essentially the question of knowledge and of thought, and all that these imply ..." [3, p. 3].

The scientific area that was supposed to focus on various representational formalisms is artificial intelligence (AI), and almost from the very beginning of AI a very paradoxical situation arose: representation was indeed considered to be a key issue but practically no work *at a basic scientific level* has been directed toward it. In other words, this area has evolved along the lines of least resistance: after some minor explorations, the researchers have chosen not to embark on the development of radically new representational formalisms.

The name "artificial intelligence" was proposed by John McCarthy at the 1956 Dartmouth workshop on "thinking machines", and later widely accepted as the name of the area dealing with the (computer) modeling of various "intelligent" processes [4, pp. 39–40, 48–50]. Unfortunately for the development of the field, "Dartmouth indeed defined the AI establishment: for almost two decades afterwards all 'significant AI advances' were made by the original group members or their students" [4, p. 49; the single quotation marks are added]. In particular, two historical facts have exacerbated this state of affairs: the dominance of logical formalisms in AI[1] and the exclusion of pattern recognition from AI[2].

The issue of why logical formalisms are not relevant to the development of representational formalisms, and are therefore not central to the development of AI, will be addressed briefly in Section 7. Here it suffices to quote Bertrand Russell, one of the prominent logicians of the last century: "Nature herself cannot err, because *she makes no statements*. It is men who may fall into error, when they formulate propositions" [9, p. 311; emphasis is added]. At the same time, it is not difficult to see why the choice of the logical formalism by McCarthy is, to some extent, not really surprising: at that time it was the only widely known "qualitative", as opposed to quantitative/mathematical, formalism. (In fact, this characterization is superficial.) Although many AI conferences as well as topics in AI conferences have the popular but nebulous name of "knowledge representation", they have not addressed representational formalisms in the sense that all natural sciences would find useful: *for the most part* they have been dealing with logical formalisms.

The separation of pattern recognition and AI had a very negative effect on the development of both fields but particularly on AI, since it put on the back burner the modeling of inductive learning, the process I believe to be the *central* intelligent process, i.e. the process around which all other relevant processes have evolved. Nevertheless, as stated in [10, p. 18], "in the late '80s, ... once political events intervened (in the form of the recent rise of connectionism ...), the situation has begun to change as can be seen from the content of

[1] See some of the popular AI texts [5], [6], [7]. The dominating role of logical formalisms in AI is mainly due to the influence of McCarthy and his Stanford AI school.

[2] "In 1973 ... pattern recognition was explicitly excluded causing much sound and fury, and wailing and gnashing of teeth. ..." [8, p. 8].

recent leading AI textbooks" (such as, for example, [11]). But this had no substantive effect on the investigation of representational formalisms, except by introducing various hybrids of classical formalisms. Thus, for example inductive logic programming is still one of the popular machine learning paradigms. What has recently changed is that the relative role of the vector-space-based statistical learning methods has increased in AI, i.e. only a very small part of the damage caused by the above exclusion of pattern recognition from AI during the last 30 years has been patched up. It is quite clear that this exclusion was motivated by the incompatibility of logical and vector-spaced-based formalisms, which instead should have motivated the development of new representational formalisms.

Primarily, because AI researchers avoided dealing with the representational issues at a more fundamental level, I believe that, when all said and done, no truly lasting ideas have emerged within this field so far, an opinion echoed in the following quote from Hilary Putnam, one of the leading philosophers of mind of the second half of the 20th century [12, pp. 182–83]:

> I am disturbed by the following, which is an undeniable sociological fact of Artificial Intelligence: no branch or sub-branch of science in the twentieth century has engaged in the kind of salesmanship that Artificial Intelligence has engaged in. A scientist in any other field who tried to sell his accomplishments in the way Artificial Intelligence people have consistently done since the sixties would be thrown out. This is something very disturbing. There are valuable engineering achievements, but the claim to have made something like a breakthrough or even a really new approach in thinking about the mind and psychology, seems to me fraudulent.

I want to emphasize one general point about the practice of AI that is largely responsible for its current regrettable state of affairs: the reliance on the language of existing (especially logical) formalisms to formalize the relevant concepts, instead of relying on the "logic" of the relevant intuitive concepts to drive the development of the corresponding formalisms.

This reliance on basic logical and mathematical formalisms also suggests that the failure of AI can be explained by the lack of relevant representational formalisms in present-day mathematics and logic, which will be discussed in Sections 3, 7, and 8.

The situation in pattern recognition (and machine learning) will be addressed in several sections below. As just mentioned in regards to AI, the basic methodological point to keep in mind here is, again, an understandable but inappropriate reliance on the *existing* basic formalisms to lead the way.

In light of the above, the main lesson from the development of AI and the basic point of the paper is this: we need to focus our efforts on the development of radically new representational formalisms that are "structural" generalizations of the classical numeric representational formalism and *allow them to lead the way.* I believe that, at present, we have *no other reasonable option.* I am also convinced that these new representational formalisms must shed new light on the nature of mathematics, as it is clarified in the epigraphs. Regrettably, as far as I know, there is only one such formalism being developed—the evolving transformation system (ETS) [1]—which is, for convenience, briefly summarized in the next section.

Finally, embarking on the development of a representational formalism, we must be prepared, to an even greater extent than Einstein was suggesting to physicists, that "In order further to approach this goal, we must resign to the fact that the logical basis departs more and more from the facts of experience, and that the path of our thought from the fundamental basis to those derived propositions, which correlate with sense experiences, becomes continually harder and longer" [13, p. 322]. I would like to add that while this "path" had quite some time to become "harder and longer" in physics, in AI, due to the more abstract nature of the field, it *must* be "harder and longer" *from the very beginning.*

2 Class-oriented representational formalisms: a one-page outline of a key topic

I now very briefly outline (see also [1, Section 1.4]) the concept of a representational formalism as that of a class-oriented representational formalism, which, I believe, the only reasonable orientation in approaching the former. The reason for this view is related to *our basic hypothesis about the informational nature of the Universe: we view the Universe as a system of evolving and interacting classes of processes/objects*. Consistent with this hypothesis, I assume that any representational formalism must represent real entities in such a manner as to *preserve their original class identity*, i.e. the corresponding mapping from the set of physical objects to the set of abstract objects in the representational formalism should preserve all original class delineations (see also Fig. 1 in [1]).

Classical, numeric, formalisms do not satisfy this requirement[3], hence their need to introduce extrinsic "decision surfaces" to deal with classes, which cannot and do not really clarify *anything* about the structure or nature of the classes. The (algebraic) *structure* of such decision surfaces has *practically nothing*—and can have hardly anything—to do with the training examples from the class: these surfaces are not sensitive at all to the data inside the decision regions, where most of the data is, and, moreover, in the case of non-linear surfaces, they have nothing to do with the underlying representational, i.e. linear, structure. So what is the *class structure* that is being captured by such surfaces?

Of course, it is understood that the above representational mapping must be realizable via some sensors (see also [14]). The main conceptual question, however, is this: What is the structure of the (abstract) underlying operations that the above representational formalism should be based on? Three points should be relatively clear. First, the operations must be chosen so as to be structurally isomorphic to the corresponding actual object operations, and hence should be structurally sufficiently diverse. Second, the same chosen operations must be adequate for supporting—in an "algebraic" sense—a meaningful concept of **class representation**, or less formally, class description (see Sections 3 and 6). In particular, class description/representation should be specified by means of the chosen operations, otherwise the representational mapping might not be able to preserve the integrity of a class. The third point concerns the preservation (by the mapping) of the modular/hierarchical object and class structure: the formal underlying structure of the formalism should allow for a *natural* transition from one "level" of representation to the next one, i.e. this transition should be possible without any modification of the underlying formal language.

3 Representational formalism through the eyes of ETS: a two-page summary

[3]This is a consequence of the fact that in such formalisms *an object is represented as a point in an abstract space, rather than as a structural entity with its own "combinatorial" structure.*

In this section, I briefly summarize the lessons learned from the development of the ETS formalism (for details see [1]). Let me start by characterizing this representational formalism as one designed to support the view of the **environment** as being formed by a hierarchically evolved—and therefore **hierarchically organized**—interacting system of **classes** (of entities, where each entity should be properly *viewed and represented as a process*).

From a representational stand-point, the environment is viewed as being multi-staged[4]. An object in an environment—approached at a particular representational **stage**—is represented by what we call a **struct**, which is a *sequence* of interconnected structural **primitives**, or **primitive transformations** constituting this struct (see Figs. 1,3). Each such primitive (see Fig. 1)—except those from the initial stage—stands for a temporal **event** signifying a compressed previous stage macro-event, which represents the interaction of one or several previous stage processes (see the top two processes in Fig. 2, left) resulting in one or several news processes (the bottom process in Fig. 2, left).

In general, the *interaction* of several real processes representationally manifests itself in their structs overlapping, i.e. sharing some their events (primitives).

The novel nature of ETS representation is its assumption that the *"structure" should be understood as a temporal recording of structured events.* In particular, this allows the introduction of structural primitives whose *syntax and semantics are inseparable*, and it also makes such primitives fundamentally different from similar concepts in previous formalisms. Moreover, the classes, stages, and the number of stages are not fixed, i.e. they can be modified by the **inductive learning process**, *which is the basic process supporting the "reconstruction" of the multistage environment by an intelligent agent.* An agent—relying on its own system of classes—tries to learn inductively the structure of other classes. It is understood that the formalism must provide the same formal language for dealing with each stage, i.e. *the formal language must be stage independent.*

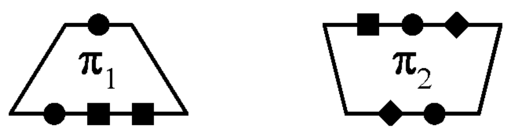

Figure 1: Pictorial illustration of two (abstract) primitives. The atomic/primitive *event* designated by primitive π_1 captures a particular split of a single initial process, depicted by the top circle, into three terminal processes, depicted by the bottom circle and two squares (see Fig. 9 in [1]).

To take an example, "water molecule" and "mammal's ear" are names (but not representations) of corresponding classes of entities separated by several stages. Obviously, during the evolution of the universe, the former class has participated, at a much later time, in the formation of the latter. A good, suggestive example for modeling a *single stage* representational structure is provided by basic chemical *structures*, i.e. molecules. Of course, this is not to suggest that chemists or physicists already knew how to represent these observable structures in a satisfactory manner.

Within a single representational stage, each **class** is also viewed as possibly **multileveled** (see Part III in [1]). A **single-level class representation** is specified by means of a (single-level) class generating system. *Each step* of the corresponding generative process is specified

[4]Each transition to the next "stage" corresponds to the compaction of representation—associated with climbing to the next conceptual level— expressed in *the same formal language* (see Fig. 2).

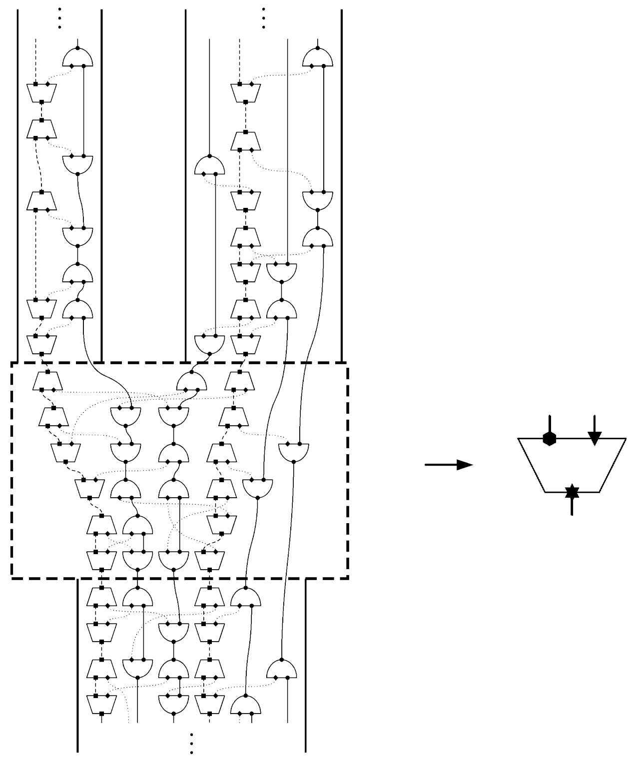

Figure 2: Pictorial illustration of a transformation (a macro-event) with two initial and one terminal processes (left) and the corresponding next-stage primitive (right) (see Fig. 5 in [1]).

by a set of structural **constraints**, one of which must be satisfied, i.e. when generating a class element out of structural pieces—each composed of the above primitives—this piece must satisfy at least one of the step's constraints. In the multilevel version of class representation—still within a single representational stage—each step is also specified via a set of structural constraints, but the structural pieces admissible at the step are composed out of elements from the previous level classes, i.e. out of **constituent elements**.

Thus, even within a single representational stage, classes may have a hierarchical structure, which provides sufficient room for the evolution of an object structure within a single stage. This structure, however, is qualitatively distinct from the *large scale, stage-related hierarchical structure*, in which the larger structural patterns—reflecting the interaction of (structural) processes—are contracted into new next stage primitives (Fig. 2). The former refers to the constituent class organization, while the latter refers to the more compact representation obtained by replacing a macro-event corresponding to the interaction of several classes, i.e. a **transformation**, with the corresponding **next stage primitive** (Fig. 2). The introduction of stages allows for an exponential reduction of representational complexity: as was just mentioned, each higher stage primitive stands for a *complex event* encapsulating an interaction of several previous stage processes.

Having fixed the nature of the generating process, any class must be adequately reconstructable from its representation. At the same time, the class representation must be efficiently learnable from a small training sample, which of course cannot be assumed to be noise free. The last two statements point to a powerful and most general *connection between* *the finite (class representation) and the infinite (set of class elements).*

In a broader context of the philosophy of science, the ETS formalism points to the basic validity of the Francis Bacon's theoretico-inductive approach to science. Moreover, it becomes quite clear why David Hume's and especially the last century's criticisms leveled at such a program went wrong: without a structural representational formalism (plus the corresponding hypothesis about the inductive structure of all classes in nature), induction may, indeed, appear to be much less reliable than it really is. In particular, looking at several "similar" ETS structs one can easily intuit why, during the inductive learning process, the enormous structural information present in the structs representing training examples allows for a quite reliable recovery of the class representation.

4 The present-day mathematics: numbers as the sole foundation of, and directing force for, the entire current scientific edifice

The story of numbers goes back in history *at least* many tens of millennia, predating the development of civilization itself. In [15, Section 3], I briefly outlined the inductive origin of natural numbers, including four stages in their emergence, and tried to justify why "the decisive and irreplaceable role of the induction axiom in the well-known Peano axiomatization of natural numbers is a reflection of the inductive cognitive origin of numbers".

A lesser-known fact about the much more recent role of numbers in shaping Western civilization is researched by the noted historian Alfred Crosby.

> Europeans of the late middle ages inherited a profound change in *mentalité* that had been fermenting for centuries, a change from the ancient qualitative way of comprehending the world—what Crosby calls the "venerable model"—to a quantitative model that would soon dominate Western society and provide Europe with the power to dominate the world.... In the space of less than a century just before and after 1300, Europe produced its first mechanical clock (which quantized time), marine charts and perspective painting (which quantized space), and double-entry bookkeeping (which quantized financial accounts).
>
> ... In 1300 everyone thought of nature as heterogeneous, each quality with its own measure. Yet 250 years later Pieter Bruegel (in his 1560 painting *Temperance*) portrayed people engaged in visualizing reality as aggregates of uniform units (or quanta): leagues, miles, degrees, letters, guilders, hours, minutes, musical notes. [16, p. 1] (based on [17])

> In practical terms, the new approach was simply this: reduce what you are trying to think about to the minimum required by its definition; visualize it on paper, or at least in your mind, be it the fluctuation of wool prices at the Champagne fairs or the course of Mars through the heavens, and divide it, either in fact or in imagination, into equal quanta. Then you can measure it, that is, count the quanta.
>
> Then you possess a *quantitative representation* of your subject that is, however simplified, even in its errors and omissions, precise. You can think about it rigorously. [17, pp. 228–29; emphasis added]

In fact, in the preface of his book [17], Alfred Crosby attributes to the above events the later amazing success of European expansion around the beginning of the 14^{th} century.

Why is there, then, a relatively *recent* perception, particularly among some mathematicians and theoreticians, that "mathematics is not about numbers"? There is no doubt that, *partly*, the latter is true, but *fundamentally* it is only an illusion. The main reason for this recent perception, and misconception, is the nature of the axiomatic form of mathematics and especially of one of its main areas—algebra—as it gradually emerged over the last 100–150 years. During the last century, algebra became a general "repository" for many useful abstract mathematical concepts/structures, e.g. group, ring, field, vector space, category, morphism. Simultaneously a tradition[5] emerged in mathematics for organizing these various abstract (axiomatic) structures in a hierarchical manner. As a result, many (but not all) texts, including some authoritative ones, have propagated the illusion of the primacy of abstract structures themselves, without explaining where these structures came from, i.e. what motivated them in the first place or by tracing their historical roots. For example, one should know that the concept of group has emerged gradually from the research on the solutions of higher degree (numeric) polynomials by Lagrange, Ruffini, Abel, and Galois around the start of the 19^{th} century.

Indeed, the concept of group is one of the most general concepts in algebra. In science, particularly in physics and chemistry, once introduced, it gradually began to play the role of a concept that is supposed to algebraically characterize the nature of a particular "invariant" of a phenomenon, i.e. the nature of the set of all transformations under which the corresponding "feature", or "structure", is preserved (see, for example, [18], [19]). In fact, the group can be used to characterize the nature, or "degree", of symmetry of a given geometric figure. I would like to emphasize here that, although the group concept appears to be independent of numeric forms of representation, this is a misapprehension: currently, the formal understanding of geometric as well as non-geometric objects is based on their *numeric representations*.

It is interesting to note that the group concept can be viewed as a concept of "*class description*" when the set of admissible classes is substantially restricted, i.e. when dealing with much more "regular" classes of phenomena. With the introduction of a structural (non-numeric) representational formalism, e.g. the ETS formalism, the description of much more general classes of phenomena becomes possible.

Another example is furnished by one of the most central mathematical concepts, the concept of a space of (numeric) functions. Although the underlying set could potentially be endowed with many abstract (and non-linear) structures, what we see in mathematics is a particular dominating structure—the linear, or vector space, structure—inspired by corresponding "numeric" considerations. And, of course, it couldn't be any other way: generalizations are always inspired by some concrete structure.

Thus, to be sure, abstractions are present in mathematics, but because they emerged as *tools* for dealing with *numeric structures*, they *must* carry at the very least *some* limitations associated with this numeric origin. We have it on the authority of no less a mathematician than Nicolas Bourbaki that "the axiomatic research of the 19th and 20th centuries has there, too, substituted little by little a unitary conception progressively *leading all mathematical notions back, first to the notion of number, then, as a second step, to that of set*." [20, p. 28, footnote 6; emphasis added]

I would like to make one final point regarding a future, "structural", mathematics, to which I will return in the next section. Although the concept of set is an auxiliary one in mathematics[6], it still occupies a very prominent, foundational, place. It appears that in

[5] The major protagonist of this tradition was a French group of mathematicians under the pseudonym of Nicolas Bourbaki.

[6] As Bourbaki correctly insists, the basic mathematical concepts are those related to the main mathematical structures, e.g. in algebra, the ones mentioned above.

a "structural" mathematics its place will be occupied by the concept of a *class* (of related objects). Interestingly, Cantor's conception of the set as "the multitude that might be thought as oneness" *may* be viewed as already pointing in that direction.

5 Our starting point: generalization of Peano axioms by replacing the successor operation by more general structural operations

What should guide us in the construction of a formalism for structural representation? Here we are entering *absolutely* unfamiliar territory: we don't have any examples of formalisms on which our (inductively trained) esthetical selection principles can be based [21]. So what should our starting point be?

We do have *one example* of a "representational formalism", albeit a non-structural one: the set of natural numbers, which are the abstract entities on top of which the edifices of mathematics and natural sciences are built. So, since this is all we have, we should try to learn *as much as we can* from this example. To this end, we should look very carefully at the *process* of construction of natural numbers. The well-known Peano axioms that define natural numbers also perfectly capture their process of construction via *one simple operation*, which is applied successively to "objects", starting from some initial "object" (see [22] or [23]). I will not discuss here the overriding importance of this construction process to the development of mathematics (see [24, chapter 1], [25]).

The temporal nature of this construction process strongly suggests that, when generalizing it, our construction must also proceed in a *temporal manner* (with a possible allowance for atemporal/parallel application of a few operations at a time, see Fig. 3)[7]. The critical point to note, however, is this: if one of the *identical* simple, i.e. almost *non-structural*, operations in the construction of natural numbers is now replaced by a few parallel structural operations, one can now "see", in contrast to the former case, *which structural operation was applied and when it was applied* (Fig. 3). In other words, for the first time, the "structure" is emerging as the corresponding temporal structure.

Such representational power, mainly due to the transparency/explicitness of the representation, has *never* before been available in any mathematical formalism, and it is this newly discovered power that we will have to learn to garner and exploit, starting literally from scratch. The full practical implications of such a representation are overwhelming since they include, in particular, the concept of a "structural measurement process" [14]. But even outside the considerations of structural measurement, we are still faced with the unprecedented challenge of gradually developing the mathematics of such temporal entities, which no doubt will require the development of completely new "tools" as compared to those developed within the classical numeric-based mathematics. (By perusing [1] it is not difficult to get a feeling why this is so.) However, for incremental progress, it is important to keep in mind that the development of any formalism proceeds much more effectively and efficiently if, at least initially, when necessary, we separate the issues related to formalism construction from the purely implementational ones. Of course, one of the more *immediate promises*

[7] We are still dealing with temporal order but extended now to small *sets of operations*, which are atemporal within each such set but whose (i.e. sets) relation to each other is temporal: the linear order of operations becomes a partial order.

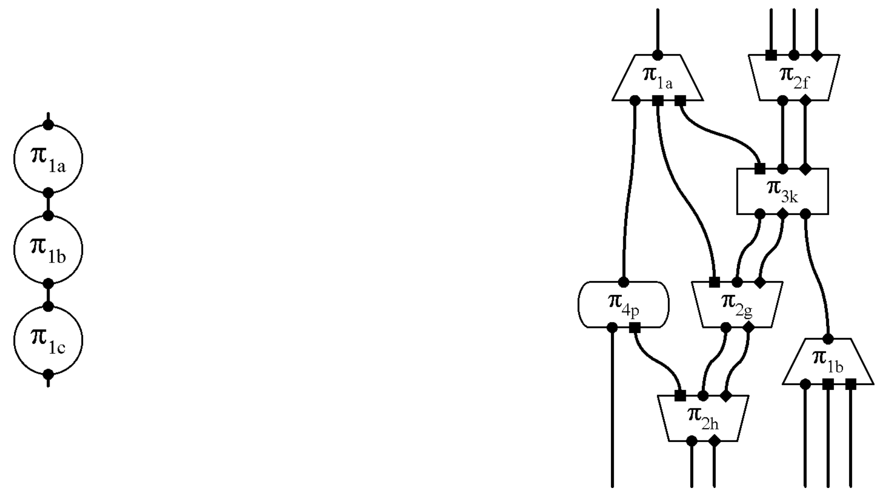

Figure 3: ETS Representation of the natural number three (left) and of a more "general" object (right), where each π_{ij} denotes an ETS primitive.

of such a powerful representation is the reliability it affords in the solution of problems in pattern recognition and other information retrieval areas.

Returning to the above structural operations—called primitive transformations (Section 2)—I note again that these (syntactic) primitives can also be considered as semantic primitives, which is not the case with the prototype they generalized, i.e. with the single numeric, or Peano, primitive. Indeed, a primitive transformation *is supposed to mimic/model the corresponding actual object's transformation*, which appears to be the key feature of structural, as compared to numeric, representation (see Section 8).

The question of whether the information processes are irreversible is also resolved. Obviously, in view of the temporal nature of such processes, the *appearance* of a primitive transformation in the representation cannot be undone, simply because all future primitives appear after it.

One of the most interesting questions raised by such an "irreversible" representation is this: Is the formative/generative history that is captured by this representation a necessary feature? From biology—more specifically, from developmental biology—we already have a clear affirmative answer: in order to "produce" an organism Nature apparently relies on some kind of formative history of the species, which the organism must recapitulate during the early stages of its development. But is this wisdom applicable to pattern recognition? I believe yes, it is, and it could not be otherwise. For example, if, in computer vision, when representing a face, along with conventional face features, we were able to incorporate some contour lines and expressions *captured temporally* as the expression changes, this rich structural information would dramatically reduce the size of the class of "similar" faces. Moreover, when then trying to recognize a face, the recognition process would abort earlier if it does not encounter some structural piece expected in the representation of the given face.

In any case, to answer the above question, one should proceed in the usual scientific manner, which was succinctly summarized by well-known physicist Richard Feynman [26, p. 19]: "How do we *know* that there are atoms? By one of the tricks mentioned earlier: we make the hypothesis that there are atoms, and one after the other results come out the way we predict, as they ought to if things *are* made of atoms." So, how will we know that an object's generative history is important? We adopt the above representation and then see if the predictions of new class elements based on this representation are valid.

I should also mention one quite obvious but critical advantage of the above temporal structural representation: given a small training set from a class, in contrast to the numeric representation, we can "directly" observe/learn the larger structural operations, macro-operations, that delineate this class. This correlates, for example, with our visual experience in a familiar environment, when we can fairly quickly see all "interesting" class features.

Finally, I draw your attention to an important philosophical observation. Although a number of past philosophers, starting with Heraclitus, insisted on "change" as the main "reality", as compared to the "visible object" reality, these philosophical observations have never been realized—except in ETS, which postulates transformations as the basis of representation—as the foundation of a representational formalism.

6 Finite class representation of a possibly infinite class

Assuming that the concept of class is primary, the concept of a naturally evolving *class representation* must be the central concept in a representational formalism. I should note that, from the very beginning, it was this vision that inspired and directed the development of the ETS formalism.

The term *class representation* was introduced in [27], and although it refers to a particular form of *class description*, it may also (more loosely) be interpreted as associated with a particular structure of the underlying formalism, within which this form of class description is a *natural* one. By introducing this term, I wanted to draw greater attention to the *form* of class description used in *any* representational formalism, given that very little attention has been paid to it: the inductive adequacy of the class representation should play the central role in assessing the adequacy of a formalism for pattern recognition. I believe that the lack of attention to the concept of class representation is not accidental, but is a consequence of an important aspect of the structure of basic (mathematical) formalisms: they lack a satisfactory concept of class description/representation.[8] In other words, it turns out that their (underlying) *formal structure* cannot support a satisfactory concept of class, and therefore *the situation simply cannot be remedied within existing formalisms by developing "more powerful" classification algorithms.*

In connection with this, one of the most remarkable telltale signs is the lack of both the concept of class *as well as* the concept of class representation in AI—including machine learning—and the avoidance of both via the introduction of "a boolean-valued function from training examples" [28, p. 21] (see also [11, p. 651]). I find this remarkable because a boolean-valued mapping is a very unnatural/roundabout way of dealing with the concept of class and highlights the intrinsic limitations of the conventional formalisms.

Moreover, the substitution of the concept of class representation by the class indicator, or characteristic, mapping has a number of serious consequences, one of which is the concealment/removal from scientific attention of the question of the relationship between a (finite)

[8] As was mentioned above, this situation can be explained by the fact that classical mathematical formalisms—which have served mainly the needs of physics—represent an object as a point in some space.

class representation and the (possibly infinite) set of its elements. I believe that the resolution of this *most central scientific question* is of vital importance to the development of AI (including pattern recognition), and that the *quality* of the answer to this question, within a particular representational formalism, is the decisive factor in judging the adequacy of that formalism. In particular, if a representational formalism offers a scientifically satisfactory and experimentally supported solution to the nature of this relationship, it can be taken seriously. Not surprisingly, the great riddle of induction is closely associated with this scientific question. Also, in light of the above epigraphs addressing the nature of mathematics, I suggest that the concept of class representation, rather than the concept of set (with the subsequent development of *non-mathematical* logical languages to support it, see Section 8), should be in the center of attention for mathematicians.

What is the relationship between a class and its representation in the ETS formalism? Very briefly (see Part III in [1]): since class representation is specified by a finite sequence of sets of constraints, the objects of the class can be generated in a systematic generative manner. In other words, a possibly infinite set of class objects is specified by a finite sequence of sets of constraints. As the class evolves, its sets of constraints also evolve in a *very natural* manner, i.e. incrementally. This hypothesis about the nature of class representation, as was mentioned in Section 4, can and should be experimentally verified. Another important point regarding this form of class representation is that we expect it to be efficiently learnable from a small training set and also to be stable with respect to various kinds of "noise" present in the object representation.

7 A new phase in applications: construction of structural representation

In this section, I want to briefly address a major shift in the relative importance of the representation phase as compared to other implementation phases.

First, following conventional scientific experience, it is quite natural to hypothesize the existence of a *structural measurement* process, which is a far-reaching generalization of the classical, numeric, measurement process. While the numeric process involves a systematic procedure (built into the measurement device) for "dismantling" the structural information present in the original object and encoding it into identical non-structural units, the structural measurement process involves the transduction of one kind of structural information into another via much more complex and dynamic interactions associated with the sensing of primitives and their connections. Although we touched on this issue in [14], it should be quite clear that this paper is not the place to address this vast topic.

Next, I want to address a radical shift in the relative importance of various constructive phases associated with the application of the ETS formalism as compared to the application of the vector-space-based formalism. In short, within the latter framework, it appears that representation is "easy" but learning is "difficult", while within the ETS framework, the expectation is that *representation is difficult but learning is easy.*

Indeed, in the implementation of a *representational* formalism, especially in the initial stages of its development, it is only natural to expect the development of the basic representation to be of central importance. (Of course, once it is satisfactorily developed, this representation can be used in a routine manner.) At the same time, once this implementation phase is complete, the learning phase is relatively "simple", in the sense that it can now, for the first time, rely on the sufficiently rich structural information present in the constructed representation. I would also like to draw your attention to the observation that human experience seems to support this "division of labor": when we find ourselves in a new environment, we spend most of our time building necessary representations.

On the other hand, within the vector-space-based formalism, which I consider separately in the next section, the main theoretical and practical efforts are directed towards the development of "good" classification algorithms (without a meaningful concept of class representation), while the construction of "good" representations, of necessity, is accomplished in an ad hoc manner. As a result, we have no reliable object and class representations, and further, at the end of the learning process, we are also no better off with respect to our ability to utilize the results of learning for various other information processing needs, e.g. discovering the connections between classes and gaining insight into the nature of data, including its main "features" (as are the needs of data mining and information retrieval). In light of this *basic* deficiency, I am also suggesting that *it is very misleading to call such statistical algorithms "learning" algorithms.*

8 The vector space as a formalism and the important applied role of its underlying formal operations

I will focus here on two key points, both of which concern the role of the formal operations by means of which a formal model is axiomatically defined: their conceptual and applied roles.

The conceptual role of such operations in mathematics has been clarified by the work of the Bourbaki group (see Section 3) but, apparently, is not widely understood. Basically, this point can be summarized as follows: the formal operations involved in the *axiomatic definition* of a particular mathematical structure (and the compositions of those operations) are the only *legitimate/admissible* operations within that structure. In case of the vector space, the two basic operations are vector addition and multiplication of a vector by a scalar. Their compositions specify what are called "linear" operations, which are the only legitimate operations in a vector space, also called a *linear space*. *No non-linear operation* is a part of this mathematical structure. Incidentally, there are uncountably many such operations, and so *even if* we wanted to choose one of them over the other, we have no *legitimate* way of doing so. Thus, if we decide to ignore this point, we are no longer dealing with a bona fide mathematical structure, and the many benefits that come from working within a mathematical structure are lost. Regrettably, many statisticians and applied mathematicians are not concerned about this.

The second point is related to the applied role of the defining operations, and has not, to my knowledge, been addressed (except briefly in several of my papers), but it is of even greater importance for us. Let us consider an "intelligent" agent operating in some environment. In order for this agent to be able to interact effectively with the environment, the agent must possess a finite set of *basic representational operations* whose compositions should be able to capture all critical "features" in the environment. Moreover, since many of these features are *compositionally* related to one another in a very particular way, the corresponding compositions of the agent's operations must be related in a *similar way* (see

Section 2)[9]. Otherwise the agent will not be able to capture the relationships within a class as well as between classes of objects (and between the features). For example, the agent will not be able to understand what a dog's ear is—which should be thought of as a class at a lower "level" and at the same time a feature at a higher "level"—and that ears of dogs of various breeds play the same compositional "role", independent of their shape, size, texture, etc. The common role of the ears can be hypothesized on the basis of the fact that they form a class with a fixed compositional role in the structure of the head. Without such a capability, no autonomous functioning in a non-trivial environment is possible. This observation implies that if we ignore the structure of basic representational operations and their interrelations, we will not be able to build a truly autonomous/intelligent system or agent, since, in this case, such a system would completely depend on the implementer to hardcode the necessary features in the environment. Furthermore, in many areas, such as AI, data mining, and information retrieval, we would also need to hardcode the meanings of all the features. However, both of these tasks are practically impossible. And even in the case that we were to hardcode most of the relevant features, the learning and classification processes will be quite brittle, as has been experienced so far.

Thus, again as in Section 2, I want to draw attention to the applied role of the basic operations (in an axiomatic specification) of a mathematical formalism: **they must be structurally isomorphic to the basic object operations in the modeled environment**. In other words, one should choose that representational formalism whose basic operations "mimic" most accurately those of the environment. The latter can be verified by observing how well the agent can *autonomously* learn various classes in the environment.

In this respect, the representational capability of the vector space model is incredibly weak, were it not sanctioned by the millennia-old tradition supported/entrenched by numeric measurement devices (see Section 4): there are only two basic operations and both are modeled on numeric operations rather than on structural ones. Moreover, their (algebraic) compositions do not offer sufficient richness, and this is precisely the main reason why various non-linear operations, including kernels, are being (inevitably) introduced.[10] But, as explained above, this will not save the situation: non-linear operations may somewhat improve classification, but since *the representational question is not being addressed at all*, the whole construction becomes very brittle, with no benefits to various postclassification processings, which are increasingly in demand, for example, in data mining and information retrieval.

One final but important point relates to a number of recent developments in pattern recognition and machine learning, in particular to various *kernel*- and *dissimilarity*-based approaches. The main observation I want to make is, in a way, quite obvious: to obtain good kernels or dissimilarity measures one needs to rely on some good "representation", since this is how they are computed. In other words, without such an (implicit, at least) representation, no kernel or dissimilarity can be introduced. Thus, without the underlying operations associated with a transformation of one object into another, such concepts cannot be implemented. In fact, this is how I was led to the development of transformation-based,

[9] During the last century, this point was also made a number of times by philosophers of mind in the context of "symbolic computation", tracing back to [29] at least.

[10] In particular, additional layers in a neural network increase the power of discriminating functions, but are not meaningfully related at all to the *class representation*.

or structural, representations, when I realized that, in order to be successful in a sufficiently rich environment, the agent must be able to evolve a *wide variety* of new dissimilarities—defined via such transformations—without the help of a human designer [30]: even for the string environment, how does one chose the "right" (sub)string operations for computing the appropriate dissimilarity measure? I firmly believe that we need fundamentally new kinds of formalisms offering a sufficiently rich and *evolving* repertoire of (structural) operations.

9 Logical formalisms

As was suggested above, the key to the success of a representational formalism is the *quality* of the class representation it provides. In the case of a logical formalism, although the situation may initially appear to be more promising than in the case of the vector space formalism[11], in reality, this "promise" turns out to be illusory.

Logical formalisms emerged from relative obscurity early in the 20th century, when researchers working on the foundations of mathematics—mainly trying to clean up the "language" of set theory introduced by Georg Cantor in the last quarter of the previous century—discovered a number of serious antinomies, e.g. Russell's, Burali-Forti's, Richard's, and Berry's. The rise of logical formalisms to prominence in the '20s and '30s was fueled by David Hilbert, the most influential mathematician of the time (after Poincare's death in 1912), who became involved in the program to secure the foundations of mathematics[12]. At that time, the work of George Boole, Gottlob Frege, Bertrand Russell, and several other researchers on the *foundations of mathematics*—and definitely not on "modeling the mind"—suddenly became of interest to some other *mathematicians*.

But what was this logical formalism (whose main ideas are attributed to Frege) motivated by? Here are some relevant points regarding Frege's motivation, which shaped, and is fully reflected in, the formalism [31, pp. 14–18; emphasis added]:

> Although ... some ... have been tempted to claim that mathematical truths in general are ultimately justified on the basis of our experience, Frege *rejects the view that the contents of our experience are at all relevant* to the truth of arithmetical claims.
>
> ....
>
> Logic, according to Frege, limns the laws of thought. *These must not be confused with the laws of thinking*, the psychological processes that might take place as one reasons. Rather, they are laws that constrain what can be rationally thought. *Logic does not aim to describe how humans think*, but instead to characterize how they must think if their thought is to remain within the bounds of reason.
>
> ....
>
> The laws of logic govern all rational thought, regardless of its particular content: *logic '[disregards] the particular characteristics of objects*, [and] depends solely on those laws upon which all knowledge rests.'

Thus, the basic logical formalism (developed for needs of the foundations of mathematics) is deliberately structured in such a way that the structure of inductive experience, as well as

[11] This is simply because logical formalisms appear to be more "qualitative" as compared to quantitative mathematical formalisms.

[12] It was Hilbert who insisted "No one shall be able to drive us from the paradise that Cantor created for us."

the structure of the corresponding biological processes, are ignored, while these are precisely the processes that should be of central importance to AI.

In connection with the development of logic at the beginning of the century, it is *instructive* to read the scathing reaction of the great Henri Poincare to those developments, which he gave in chapters 3, 5 of [32]. Poincare also criticized the introduction in mathematics of *actual* infinity, as opposed to potential infinity. Most importantly, he was convinced of the central role of inductive constructions in mathematics and that

> All the efforts that have been made to upset this order, and to reduce mathematical induction to the rules of logic, have ended in failure, [and this failure is] but poorly disguised by the use of a language inaccessible to the uninitiated [32, last section titled "General Conclusions"].

It is also interesting to note that, as his quotes in Section 1 and below clearly indicate, later on in his life Bertrand Russell changed his mind about the role of logical formalisms:

> Inference is supposed to be a mark of intelligence and to show the superiority of men to machines. At the same time, the treatment of inference in traditional logic is so stupid as to throw doubt on this claim .... I have never come across any ... case of new knowledge obtained by means of a syllogism. It must be admitted that, for a method which dominated logic for two thousand years, this contribution to the world's stock of information cannot be considered very weighty. [33, p. 63]

> I had become increasingly aware of the very limited scope of deductive inference as practiced in logic and pure mathematics. I realized that all the inferences used both in common sense and in science are of a different sort from those in deductive logic. [34, p. 141]

However, most interestingly, one should look at how the original father of logic, Aristotle, perceived the place of logical mechanisms. I quote one of the most known philosophers of science of the last century, Karl Popper, who, in his words "didn't like Aristotle" precisely for his epistemology:

> [Aristotle] tries to give a theory of *epistēmē*, of demonstrable knowledge; and being a clever man, and a good logician, he finds that his assumption that there is demonstrable knowledge involves him in an infinite regress, because this knowledge, if demonstrated, must be logically deduced from something else, which in turn must also be demonstrated knowledge, and therefore in its turn deduced from something else, and so on.
>
> So he gets to the problem: how can this infinite regress be stopped? Or: what are the real original premises, and how do we make sure of their truth? He solves this fundamental problem of knowledge by the doctrine that the real original premises are statements of definitions. ...Definitions, on the one hand, give to words a meaning by convention and are therefore certain (analytic, tautological). But if they are only conventional, and therefore certain, then all *epistēmē* is truth by convention and therefore certain. In other words, all *epistēmē* is tautological, deduced from our definitions. This conclusion Aristotle does not want, and he therefore proposes that there exists, on the other hand, also definitions that are not conventional and not certain. Yet he does not stress that they are not certain, only that they are the result of 'seeing the essence of a thing', and so synthetic; they are the result of induction.
>
> This seems to have been the way in which induction entered into the theory of scientific method, of epistemology. [35, p. 2]

To the above, I want to add a few, more formal, reasons why logical formalisms cannot be used as representational ones. Basically, all main features in the characterization of a representational formalism, as proposed in Sections 2 and 3, are missing: logical primitives cannot at all model structural object operations (and *they were not intended* for this purpose)[13], no

[13] That is why, in logical formalisms, one has to deal with syntax, semantics, and interpretations.

multilevel structural representation can be supported[14], no inductive class representation is possible, and thus *there is no representational mapping* as outlined in Section 2. In particular, predicates cannot be *sensed by a sensor* as they are *language-like entities* created by the human mind, *after* it has sensed and interpreted the environment. (For additional discussion see also [30, Section 5], in which I outlined a possible role of logical mechanisms that turns out to be reasonably consistent with that advocated by Aristotle.)

Thus, in view of the above, it was a *gross* misunderstanding to choose the logical formalism, even in an expanded form, as a basic *representational* formalism for AI.

10 Computational models

Turning our attention to computational models, I note that, as is generally accepted, computability theory is in fact part of logic[15], and since computational models do not have much claim on representational novelty, I want to touch on the following two issues only: Chomsky's generative grammar model and the inappropriate obsession with computational complexity in machine learning.

First of all, one should acknowledge the historical importance of Chomsky's formal grammar model: it was the first model to emphasize the critical role of generativity, and in this sense its importance should not be underestimated. However, even though Chomsky stressed its central role in his model, a generative object history simply cannot be represented as part of a string (over the basic alphabet), and therefore this critical feature of the model *could not be incorporated into an object's representation*, i.e. into a string (see also the next section). Moreover, since Chomsky has been consistently opposed to putting this model into an inductive, i.e. more dynamic, context, no important issues related to the inductive nature of languages and grammars or their inductive connections have been addressed by his school. In particular, a *grammar is not viewed as an (inductive) class representation*. When the corresponding issues arose in pattern recognition in the '70s and '80s, it gradually became reasonably clear that the framework in its current form was inadequate for the goals of pattern recognition [30], [37]. In essence, given a training sample—since no generative information is present in the training strings themselves—there is no reliable way to associate a class grammar with the training strings (even under reasonable restrictions on the class grammars). Of course, the presence in the model of the second, somewhat ad hoc, alphabet does not help, either.

The second point I would like to draw attention to is the absolutely inappropriate obsession with computational complexity *above all other considerations*, including representational issues. I am convinced that we should follow the wisdom of, for example, a physicist who would be appalled if she was approached about complexity issues before a satisfactory physical model of the phenomenon was available. Of course, complexity issues have their place, but *one should not put the cart before the horse*. Moreover, a satisfactory representational formalism *should* allow for efficient learning algorithms, otherwise, as was mentioned above, it simply cannot be "satisfactory". (Again, when have physicists worried about computational issues?) Thus, for example, the presence of a generative object history as a part of an

[14] At least one prominent logician has drawn attention to this: [36, part III].

[15] One should also not forget that the main tenets of computability theory were developed by logicians, i.e. Post, Gödel, Church, Turing, Kleene, and (more recently) Cook.

object's representation in the ETS formalism exponentially reduces the number of candidate classes from which a training set could have originated.

11 The inadequacy of popular discrete representations

First of all, it is important to note that all such "representations"—e.g. strings, trees, graphs—have emerged *outside the scope of any representational formalism* and became "representations" simply by default as the importance of "discrete" representations gradually grew with the advent of computers.

As a result, there is no (classical) mathematical framework in which, for example, the set of all strings over a fixed finite alphabet can be reliably considered as elements of some (formal/mathematical) "space", and there are actually some good reasons for this: in contrast to classical mathematical objects, there is no "universal space" that can *naturally* accommodate this set of strings. In my opinion, the latter situation is mainly due to the fact that the relationships between two strings are not defined well, and therefore there are no *well-defined* classes of strings. I believe that unless some temporal dimension of representation is introduced—as was done in the ETS formalism, for example—we are not dealing here with a reliable form of representation. A simple example might clarify this point: How do I compose/construct a new, non-trivial sentence? The process does not proceed from left to right as the final written form may suggest, simply because the generation of a sentence in my mind follows a more complex "non-linear" order, not (obviously) related to the linearly-written sequence of words. Of course, the same is true for DNA sequences. The representational situation with graphs is even more problematic, since a graph's formative history is more complex than that of a string.

Thus, the inadequacy of the popular discrete representations, e.g. strings, trees, and graphs, cannot be remedied in a simple manner: it requires radical rethinking. It is such rethinking that actually led to the ETS formalism. I am afraid that a "relational revolution", in which "statistical machine learning is [currently] in the midst of" [38], has nothing to do with the radical *representational* rethinking I am referring to (see Section 2).

12 Why conventional methodologies for the evaluation of performance of learning algorithms are inadequate for that purpose

First of all, it is very important to realize (particularly for funding agencies) that, today, in the age of computing, almost any model, including a very poor one, could be "made to work" in a number of applications, especially when given enough human resources. That does not at all mean that the model has a scientific merit. What makes a new model *scientifically* attractive is its important explanatory value, i.e. it must hypothesize a qualitatively new, non-trivial feature of *reality* that one should be able to verify experimentally (otherwise, at best, it is not a "new" model). If no interesting/novel feature of reality is being hypothesized, the model, accordingly, does not have much *scientific* value. In AI, one must insist on this criterion to an even greater extent, since there is a strong intuitive expectation that any natural environment is "meaningful", or "full of meaning". Thus, in pattern recognition (and machine learning), from a *scientific* point of view, a useful formalism is supposed to advance a “useful” hypothesis about the *structure of pattern classes* in the Universe[16]. Again, as was mentioned in Section 5, there is a longstanding scientific practice concerning the verification of such hypotheses. But, above all, one must *have* such a hypothesis. The ETS formalism advances such a hypothesis, while no other popular machine learning model does so.

On the other hand, from an engineering, or applied, perspective, the situation looks quite different. This is because an engineering practice, in contrast to a scientific one, revolves around implementations. Of course, this is not to say that implementations do not require creativity, but rather that this kind of creativity is typically confined to a different level/context: scientific creativity is about advancing a new hypothesis concerning the nature of reality, while engineering creativity is about finding original and efficient ways of implementing the experimental part of the enterprise, including the verification of scientific hypotheses.

So what is the situation in pattern recognition? Since pattern recognition (PR) has been perceived as an engineering field, engineering methodologies and practices apply, including the adoption of conventional statistical methodology. As a result, we have the following situation. Suppose you and I preprocess the data independently, and my PR algorithm performs at the level of 98% and yours performs at 88%. Does that mean that my algorithm is more satisfactory than yours? No, of course not. For example, no one knows which kinds of “talents” were put to work in the preprocessing stage. Moreover, if we use completely different pattern representations, the situation becomes even more confusing. For example, how do we know that my algorithm is as robust as yours when we substantially increase the number of classes or slightly change the original setting?

Thus my main point is this: the statistical methodology *could* be made more meaningful, *provided* the underlying representational *formalism* is fixed. What should be clear from the above is that, at this stage in the development of AI (including PR), it is the representational formalism that should be of main concern, and its development is much more a scientific, rather than an engineering, matter. Once a satisfactory representational formalism is developed, *only then* will it be “safe” to proceed with the development of learning algorithms, as well as the corresponding statistical methodology for the verification of their performance.

13 Conclusion

We need a representational formalism that offers powerful class representation capabilities and that would *clarify the nature of the ubiquitous link between an object and its class*. Since the development of mathematics and logic (including computability theory) were not motivated by, and their existing models are not appropriate for, such a goal[17], a radically new formalism is called for. Of course, I am not the first to suggest the need for a fundamentally new formalism. In fact, a number of leading scientists during the last century have suggested this; see, for example, Schrödinger [39, pp. 143–162] and, more relevantly, von Neumann [40, especially pp. 80–82].

[16] I don’t believe that by restricting the environment we can expect the corresponding hypothesis to be much simpler.

[17] In fact, they have not been dealing with *any* (structural) representational formalisms.

Over the last 50 years, we have witnessed too many non-radical and failed attempts in AI and PR, including the appearance in the '80s of such "new" areas as machine learning and neural networks. I believe that, by now, a healthy scientific sense should strongly suggest that we can't squeeze out of the conventional formalisms—e.g. out of decision surfaces or internal nodes of ANNs—any representational miracles. Moreover, the same scientific sense suggests that we should not expect any other miracles from conventional formalisms, or that "one should not expect to find a lost key, at night, only under an existing light".

I am convinced that we are at a *very important* crossroads in the history of science: we will not make substantive progress unless we embark on the development of a fundamentally new (structural) representational formalisms. In fact, there appears to be no way around this undertaking. It is a very big—indeed *unprecedented*—scientific step, precipitating an attendant restructuring of science. To some of us, however, it is also a very exciting journey, one which actually began more than ten millennia ago with a relatively small collection of clay shapes (spheres, disks, cones, tetrahedrons, ovoids, cylinders, rectangles, and others) [41, p. 11] and which, after a numeric detour, is once again leading to new, much more powerful, structural entities.

Acknowledgments. My sincere thanks to Sean Falconer, David Gay, Oleg Golubitsky, Alexander Gutkin, Dmitry Korkin, and Ian Scrimger for their comments and help.

References

[1] Lev Goldfarb, David Gay, Oleg Golubitsky, Dmitry Korkin, What is a structural representation? A proposal for a representational formalism, this special issue, 2006.

[2] Lev Goldfarb, Representational formalisms: why we haven't had one, in Pattern Representation and the Future of Pattern Recognition: A program for action, *Proc. Satellite Workshop of the 17th International Conference on Pattern Recognition*, ed. Lev Goldfarb, Cambridge UK, August 22, 2004, pp. 5–22

[3] Allan Paivio, *Mental Representations*, Oxford University Press, 1990.

[4] Daniel Crevier, *AI*, Basic Books, New York, 1993.

[5] David Poole, Alan Mackworth, Randy Goebel, *Computational Intelligence*, Oxford University Press, 1998.

[6] Michael R. Genesereth and Nils J. Nilsson, *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, 1988.

[7] Eugene Charniak and Drew McDermott, *Introduction to Artificial Intelligence*, Addison-Wesley, 1985.

[8] Alan K. Mackworth, IJCAI, *Canadian AI Magazine*, Winter 1995.

[9] Bertrand Russell, *Wisdom of the West*, Crescent Books, 1989.

[10] Lev Goldfarb, The inductive learning model, *Canadian AI Magazine*, Winter 1996.

[11] Stuart Russell and Peter Norvig, *Artificial Intelligence*, Prentice Hall, 2003.

[12] Hilary Putnam, Against the New Associationism, in Peter Baumgartner and Sabine Payr, eds., *Speaking Minds*, Princeton University Press, 1995.

[13] Albert Einstein, *Ideas and Opinions*, Bonanza Books, 1954.

[14] Lev Goldfarb and Oleg Golubitsky, What is a structural measurement process?, Faculty of Computer Science, U.N.B., Technical Report TR01-147, October 2001.

[15] Lev Goldfarb, Inductive theory of semiotic systems, *Proc. 1997 Conf. on Intelligent Systems and Semiotics*, ed. A.M. Meystel, NIST, 1997, pp. 111–116.

[16] Lynn Steen, Notes on [16], http://www.stolaf.edu/other/ql/crosby.html, 1999.

[17] Alfred W. Crosby, *The Measure of Reality*, Cambridge University Press, 1997.

[18] Joe Rosen, *Symmetry in Science*, Springer, 1995.

[19] Istvn Hargittai and Magdolna Hargittai, *Symmetry through the Eyes of a Chemist*, Plenum Press, 1995.

[20] Nicolas Bourbaki, The architecture of mathematics, in *Great Currents of Mathematical Thought*, ed. F. Le Lionnais, vol. 1, Dover Publ., 1971.

[21] James W. McAllister, *Beauty and Revolution in Science*, Cornell University Press, 1996.

[22] Edmund Landau, *Foundations of Analysis*, Chelsea, 1960.

[23] Leon W. Cohen and Gertrude Ehrlich, *The Structure of the Real Number System*, Van Nostrand, 1963.

[24] Henri Poincare, *Science and Hypothesis*, Dover, 1952.

[25] G. Polya, Mathematics and Plausible Reasoning, Volume 1, Princeton University Press, 1990

[26] Richard P. Feynman, *Six Easy Pieces*, Addison-Wesley, 1995.

[27] Lev Goldfarb, Inductive class representation and its central role in pattern recognition, *Proc. 1996 Conf. On Intelligent Systems: A Semiotic Perspective*, eds. J. Albus, A. Meystel, R. Quintero, Vol. 1, pp. 53–58.

[28] Tom M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.

[29] Kenneth J. W. Craik, *The Nature of Explanation*, Cambridge University Press, 1943.

[30] Lev Goldfarb, On the foundations of intelligent processes I: An evolving model for pattern learning, *Pattern Recognition 23*, 1990, pp. 595–616.

[31] Alexander George and Daniel J. Velleman, *Philosophies of Mathematics*, Blackwell, 2002.

[32] Henri Poincare, *Science and Method*, Dover, 1952.

[33] Bertrand Russell, *An Outline of Philosophy*, Unwin Paperbacks, 1986.

[34] Bertrand Russell, *My Philosophical Development*, Unwin Paperbacks, 1985.

[35] Karl Popper, *The World of Parmenides: Essays on the Presocratic Enlightenment*, Routledge, 2002.

[36] S. Yu. Maslov, *Theory of Deductive Systems and Its Applications*, MIT Press, 1987.

[37] Eiichi Tanaka, Theoretical aspects of syntactic pattern recognition, *Pattern recognition 28*, 1995, pp. 1053–1061.

[38] *SRL2004: Statistical Relational Learning and its Connections to Other Fields*, ICML 2004 Workshop, http://www.cs.umd.edu/srl2004

[39] Erwin Schrödinger, *Nature and the Greeks & Science and Humanism*, Cambridge University Press, 1996.

[40] John von Neumann, *The Computer and the Brain*, Yale University Press, 1958.

[41] Keith Devlin, *Mathematics: The Science of Patterns*, Scientific American Library, 1994.