



A glimpse at the metaphysics of Bongard problems

Alexandre Linhares¹

*National Space Research Institute of the Brazilian Ministry of Science and Technology, LAC-INPE,
Av. Astronautas 1758, S.J. Campos, SP 12227-010, Brazil*

Received 4 November 1999; received in revised form 17 July 2000

Abstract

Bongard problems present an outstanding challenge to artificial intelligence. They consist of visual pattern understanding problems on which the task of the pattern perceiver is to find an abstract aspect of distinction between two classes of figures. This paper examines the philosophical question of whether objects in Bongard problems can be ascribed an *a priori*, metaphysical, existence—the ontological question of whether objects, and their boundaries, come pre-defined, independently of any understanding or context. This is an essential issue, because it determines whether a priori symbolic representations can be of use for solving Bongard problems. The resulting conclusion of this analysis is that in the case of Bongard problems there can be no units ascribed an *a priori* existence—and thus the objects dealt with in any specific problem must be found by solution methods (rather than given to them). This view ultimately leads to the emerging alternatives to the philosophical doctrine of metaphysical realism. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Philosophy; Pattern understanding; Bongard problems; Metaphysical realism; Multiperception

1. Introduction to Bongard problems

Three decades ago the intelligence theorist Mikhail Moiseevich Bongard posed an outstanding challenge to artificial intelligence [1]. His book *Pattern Recognition* (a translation of the Russian *Problema Uznvaniya*) brought a remarkable set of 100 visual pattern understanding problems where two classes of figures are presented and the pattern recognizer (either a human or a machine) is asked to identify the conceptual distinction between them. Sometimes the classes are opposite in terms of this conceptual distinction,

E-mail address: linhares@crt.umontreal.ca (A. Linhares).

¹ Current address: Centre de Recherche sur les Transports, Université de Montréal, Pavillon André-Aisenstadt, CP6128, succ. Centre-ville, Montreal, Quebec, Canada H3C 3J7.

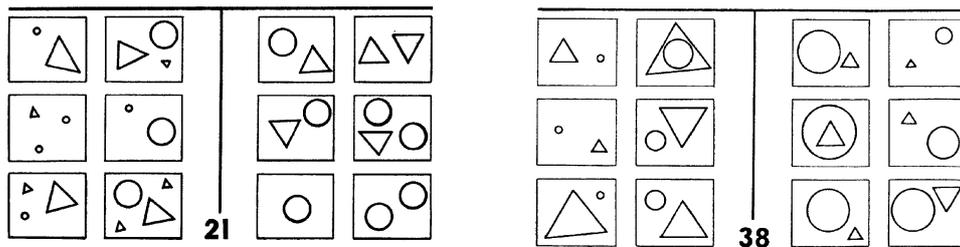


Fig. 1. Bongard problems BP#21 and BP#38. What abstract aspect distinguishes the boxes on the right (class 1) from the boxes on the left (class 2)? [From M.M. Bongard, *Pattern Recognition*, Spartan Books, 1970.]

such as large figures versus small figures, and other times there may be properties or relations holding between boxes in one class, but not in the other, such that there is always some aspect to distinguish the classes.

Fig. 1 displays two simple Bongard problems dealing with triangles and circles. One of the most important characteristics of such problems is that, although humans can generally solve them intuitively, their automation is simply daunting: there is always much relevant information to be perceived and much irrelevant information to be discarded. There is also a need for an interplay of underlying mechanisms that includes (low- and high-level) vision with high-level cognition, and such mechanisms are pervaded with ambiguity at all levels. This *functional integration* is one of the reasons that makes the automation of Bongard problems such a formidable task.

A skeptic might be led to argue that Bongard problems are “of no more interest than any other toy-world”. That would be a great underestimation of the information-processing tasks required to solve them; there is no cannibals and missionaries problem here [19]. Bongard was convinced that these problems captured the core of the problem of perception, and demonstrated that classical approaches to pattern recognition, such as classification by hyperplanes or by Rosenblatt perceptrons, were bound to fail. For example, should one attempt to project the input in the “receptor” to an n -dimensional space and draw a hyperplane in order to separate (and thus classify) the inputs, this would lead to the water and sponge situation: both classes are intensely interconnected and there is no hyperplane separating “water” from “sponge”. A similar inadequacy arises with the Rosenblatt perceptron [1].

Though involving an integration of pattern recognition and artificial intelligence, Bongard problems have been virtually ignored by the research community, and there are hardly any published studies on the subject. Nearly ten years after the appearance of Bongard’s book, intelligence theorist Douglas Hofstadter sketched out an elegant parallel processing architecture based on the HEARSAY II speech-understanding project [11]. However, given the technological standards, and the magnitude of the task, he withdrew from implementing the project (Hofstadter currently advises a Ph.D. project on the subject at the university of Indiana, Bloomington). Bongard problems then remained untouched for more 15 years, until computer scientists Kazumi Saito and Ryohei Nakano of NTT Communication Science Labs developed the RF4 project [26,27] briefly described below.

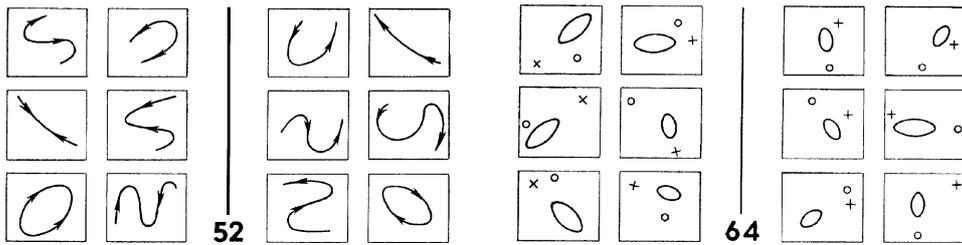


Fig. 2. Bongard problems BP#52 and BP#64 demonstrate the influence of cultural factors. [From M.M. Bongard, *Pattern Recognition*, Spartan Books, 1970.]

1.1. Influence of cultural factors

An interesting but generally ignored aspect of Bongard problems is that their difficulty for a given subject is directly associated his or hers (or the system's) previous experience. Since the problems consist of geometric figures, one may be led to believe that cultural factors do not influence the performance of a person attempting to solve them. This is not the case. Just as the Turing Test for intelligence is permeated with cultural issues [6], Bongard problems are prone to the specific—and subjective—way subjects perceive things. For instance, BP#52 is easy for humans because we learn at infancy to follow arrows; it would in fact be a very hard problem for an alien intelligence that had never worked with this specific symbolism. Consider, on the other hand, problem BP#64, which is not an easy problem for humans.

Problem BP#64 would indeed be very easy for an experienced sonar operator used to seeing ovals as submarines, crosses as mines, and small circles as whales. This operator would conclude instantly that on class 1 the submarines head towards the mines, while on class 2 the submarines head towards the whales. Note that, though there are no submarines, mines or whales involved, for this person it is impulsive to perceive the ovals as submarines, and unnatural to perceive them as ovals, simple geometric figures. The previous experience demands that from such a sonar operator, and, in this case, as soon as he or she tracks the “path of the submarine”, the answer to the problem emerges. An ordinary person would have no special reason for tracking such path, and could thus face a harder problem.

The reader should note that this does not mean that cultural factors are *essential* for solving Bongard problems. This is not the case. In fact, one may assume that, in order to be a valid problem, only geometric information should be needed to solve it. Thus, the claim here is that cultural factors, though not essential, do indeed *influence* the processing of geometrical information when conducted by humans.

We may now initiate our exploration of the metaphysics underlying Bongard problems.

1.2. The philosophical problem

This paper addresses the ontological problem of the *a priori* existence of objects (i.e., bounded entities) on the domain created by Bongard. This is a fundamental question, for the way one stands in relation to it largely determines one's approach towards automating their solution. As we may see, some prized intuitions do not hold, at least in Bongard's

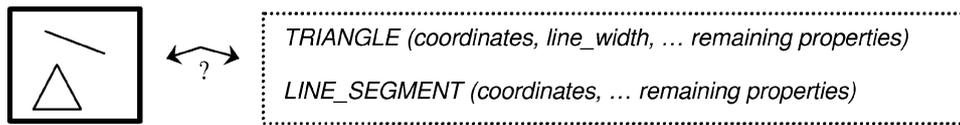


Fig. 3. Raw geometrical information versus symbolic descriptions.

domain; as we start with an intuitive notion of perception, a notion that could be traced back to Greek philosophers, and then gradually come to see that perception is unlike that simple notion at all.

It is better to start our discussion with the *independence assumption*:

“existence and fact are independent of human cognitive capacities. No true fact can depend upon people’s believing on it, on their knowledge of it, or on any other aspect of cognition.” [15, p. 164]

The independence assumption defines what is a mind-independent fact. Consider, for instance, the one-hundred dollar bill. Physics tells us that it has mass, and that under the right circumstances it irradiates a variety of wavelengths of light. These are metaphysically external facts because they happen to be true without depending on any understanding whatsoever. For all we know, these facts remain true independently of any cognition or observation. However, the *value* of the bill does not, as it is determined by human institutions and standards. Another example would be the *colors* perceived in the bill; there are no colors “out there”. Outside of human cognition there are only wavelengths of light, and colors exist only in the eye of the beholder.

The reader may be wondering whether this might be a venture into a territory of purely abstract questions of no significance to those interested in automating the solution of Bongard problems. This is not so, and to make absolutely clear how this study relates to solution methods, we may focus our effort on this particular issue: suppose that a specific Bongard problem includes the following box shown in Fig. 3. Would it be appropriate to discard the raw geometrical information in favor of a simple symbolic description, such as that presented?

This is the core question of our investigation. What is in question here is not only the appropriateness of this type of representation, but also the underlying ontological question concerning the very existence of the objects TRIANGLE and LINE_SEGMENT in the first place—an *a priori* existence which is external to any understanding. A possible answer to this question is “Yes, there is clearly a triangle and a line segment, this fact is independent of cognition, and nothing will be lost by describing them symbolically instead of using the raw data”. This is the classical view: there exist objects (i.e., bounded entities) in a realm outside of any understanding, and, for each one of those objects, the process of perception is responsible for finding its one fundamentally correct description. We can put this idea as the “one object, one view” hypothesis.

We will now argue that these are faltering philosophical grounds on which to approach Bongard problems.

2. One object out there, one view in here

Once again, we are considering the notion that perception is the process by which objects (existing outside of any understanding) come to be mapped into their (single correct) corresponding descriptions. This is the view that we “just happen to see” the things out there, we see those things *as they are*, a widely held and rather intuitive notion.

In order to make an accurate picture of this “one object, one view” hypothesis, let us first provide a brief introduction to the philosophical doctrine known as *metaphysical realism*. Metaphysical realism is the view that there is a real world external to humans (realism), and, furthermore, that there can be an external (metaphysical) perspective of reality—a *God’s eye view*—that is not only correct, but unique, and also independent of human cognition. Its division of the world into objects is particularly interesting. To borrow the words from Lakoff,

“An ‘object’ is a single bounded entity. According to metaphysical realism, there is a correct and unique division of reality into objects, with properties and relations holding among them. Each ‘object’ is a single bounded entity, and that is *the only correct description* of that object. It cannot also be correctly described as a plurality of objects or a mass of waves. That is what metaphysical realism says: there is only one correct way in which reality is divided up into objects.” [15, p. 262]

He then goes on arguing that a chair may be seen as a single object, or as some distinct parts, *which are objects in themselves*, or maybe as a huge collection of molecules, or even as wave forms (taking the extreme point of view of wave equations in physics). The point is that *all views are correct*, and that there is no single correct God’s eye view division of the world into objects (and into categories, his main thesis).

In fact, perception is seen under this view as a simple one-to-one mapping, a mere function from objects existing “out there in reality” to descriptions existing “in here”, holding the intrinsic assumption that any object X needs only one representation (or perception-mapping) $P(X)$. A representation that could, in principle, be defined *a priori* and given to the system. This leads us to our first claim.

Proposition 1. *The use of fixed, a priori, representations presupposes metaphysical realism.*

Representations fixed *a priori* have actually been adopted in a general-purpose search program, known as RF4.² RF4 is referred to as a “concept learning algorithm” that uses formulas to discriminate between examples, as it is stated that “*a concept that discriminates examples is assumed to be describable by a formula*” [26,27]. For instance, on BP#38, where class 1 contains triangles that are larger than the circles, while class 2 contains triangles smaller than the circles, the reported formula for the solution is [26]:

² It should be noted that RF4 is not supposed to be defending any theory in cognitive science. It is an advanced AI general-purpose search algorithm, one specific application of which is in the domain of Bongard problems (in fact, it is the only published attempt at Bongard problems to date). As we will be considering, in the case of Bongard problems, symbolic *a priori* representations impose serious limitations for search algorithms, but these limitations may not arise in the other domains for which RF4 is applicable.

$$\begin{aligned} \text{BP\#38: } \forall B \in \text{boxes}, \exists x_1 x_2 \in B, \text{ shape}(x_1) = \text{polygon} \wedge \\ \text{shape}(x_2) = \text{oval} \wedge \\ \text{size}(x_1) > \text{size}(x_2) \Rightarrow \text{class } 1 \end{aligned}$$

where B is an instance of each figure of the class.

These formulas can be easily applied because the representation of objects is given by the users as “compact input information”: line segments, circles, triangles, and polygons in general are given their coordinates, outline, internal texture, and other properties. The papers do not state how complex curves or other objects of higher entropy are dealt with. We refer the reader to [26,27] for more detailed descriptions of the project.

So, in the perspective of RF4, the answer to our previously posed question is “YES”. RF4 does not deal with raw information from binary images, but only with pre-defined symbolic descriptions, such as those presented. It follows implicitly that this is such a correct, unique, and external (independent of cognition) description of the pattern. The project regards a triangle, for example, as a single bounded entity, which cannot be regarded in any distinct manner—this is why it implicitly supports the doctrine of metaphysical realism.

The algorithm is, according to the references [26,27], capable of solving 41 Bongard problems. This means that a satisfactory formula could be found (however, the papers do not provide the specific problems which were solved, neither the formulas obtained by RF4).

An idea that comes to mind is that of a “perception module” P that could eventually provide the representations that search algorithms need. This should bring “perception” capabilities to the system, by exchanging the information in image form to proper symbolic descriptions. It also should be a separable module that does not interfere with the task of searching for the solution—for which the search algorithms are devised.

However, on a closer look, could P really work? In order to be an attachable module, it would have to be absolutely context-free, in the sense that it could look into one box only, and create a single correct description of what it finds in that box. It is not appropriate to consider otherwise (the possibility that it might look into the other boxes before describing each one of them), because in this case it would not be a separable module in the first place, as it would be absolutely integrated with the system, and would itself in fact be looking for the solution. After all, if the module has to look into two or more boxes, it is because it is looking for the similarities and differences among them (and thus this perception “module” is really inseparable from higher-level processing).

In fact, this module would have to be a one-to-one mapping. If it is a separable module, then it starts from the raw data, perceives what is in there, creates a description, and stops at this point. Given an object X , it finds *the corresponding* representation $P(X)$. It could never, from the same raw data, create potentially different descriptions.³ It would need to create a single description and pass it on for the further processing of the system (i.e., the search algorithms). Thus, we have:

³ Well, in fact, it could, for example, if one poses the description problem as an optimization problem, and resorts to an approximate solution, instead of taking an optimal one. However, this would not alter anything, because the real claim here is that the information needed to create a description can reside out of the corresponding box; it may be distributed over all the other boxes, over the whole context of the problem.

Proposition 2. *A separable perception module would necessarily have to provide a single description for a single box, in all the cases that the box might appear.*

Proposition 2 brings up the question: Could this single-description module work? Not for solving Bongard problems in all their richness, because, as we will see, these problems require multiple distinct descriptions for each single box of raw data.

To demonstrate this, we may use some of Bongard’s ambiguous examples, even for simple objects such as a triangle, or a line segment. A triangle has a simple mathematical definition. In a God’s eye view there can be no ambiguity for such thing as a triangle, because either a geometrical arrangement is a triangle or it isn’t. However, *there can be no single correct description* of a geometric arrangement in Bongard problems, for such a fixed description would inevitably break down in many cases. Let us state this explicitly:

Proposition 3. *The very same geometrical arrangement may need fundamentally different representations for each context on which it arises.*

Bongard provided a great number of examples where the same geometrical arrangement should be described in *fundamentally distinct* ways. A case in point: what is to be seen in BP#06 as a single object (i.e., a triangle) later appears in BP#91 (and in BP#85) as three distinct objects (i.e., three line segments that just happen to meet by coincidence).

Another example: in BP#85, straight line segments must be seen as continuous, even when they intersect with another line segment. This does not happen, for instance, in BP#87, where the intersections happen to *split the line segments*. There is a persistent problem on the definition of what is to be regarded as a bounded entity in a problem, of where those boundaries reside, and of what should be the right level of description of the objects.

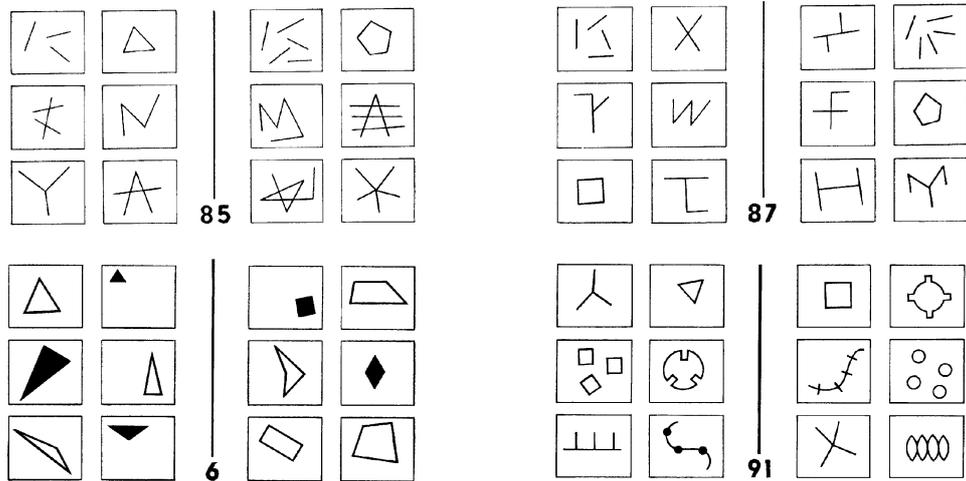


Fig. 4. Bongard problems require multiple potential descriptions—even for arrangements as simple as line segments and triangles. [From M.M. Bongard, *Pattern Recognition*, Spartan Books, 1970.]

As one examines these problems, it becomes apparent that a perception module could not, and this is *in principle*, create a description of their content without examining the overall context of the problem (i.e., looking at the other boxes), without interfering with the task of finding out what the solution might be like. It is impossible to separate the process of representation of one box from the consideration of the whole group.

Proposition 4. *It follows from Propositions 2 and 3 that there could be no separable perception module—and hence the process of describing the content of the boxes must thus be tightly integrated with the process of perceiving similarities and distinguishing differences among them.*

Even such clear—and mathematically definable—concepts such as triangles and line segments can have a great number of distinct descriptions, all of them correct. In short, following Lakoff [15], there is no God’s eye view of a triangle, or even of a line segment, because *there is no God’s eye view of any geometric arrangement whatsoever*. Hence, there can be no rigid representation given *a priori* for any arrangement whatsoever.

Since this is an extremely important point, as it shows that Bongard problems are beyond the capabilities of those computational approaches founded on metaphysical realism, it is worthwhile to restate the significant objections against the “one object, one view” hypothesis:

First, *there cannot be a separate perception module “feeding” context-free descriptions into an inference engine* (or any other higher-level cognition module conceivable). The system must be tightly integrated. The reader should note that this is hardly an original claim [2,3,7–9,12].

Second, *there is more than one way to represent the “very same thing”* [12,14].

So, as it turns out, when properly tested, *the “one object, one view” hypothesis is proven false*. There exist problems (such as those shown) that cannot be solved by any approach founded on this notion.

But what if an object could have multiple descriptions available? Would that provide a genuine way to attack Bongard problems? We may thus move on to a new hypothesis, based on the idea of multiple potential descriptions arising from a single object, an idea that we may refer to as the “one object, many views” hypothesis.

3. One object out there, multiple views in here

Let us examine this hypothesis closely. What is the idea like? In this conception, we reject the idea of perception as a simple function from objects to descriptions, and now conceive it as a one-to-many relation, or at least a function taking additional parameters (instead of just external objects). In fact, to mark this conceptual distinction, we shall be using the term *multiperception*. Computer science is pervaded with the idea of multiple instantiations of objects, such as multiprocessing, multitasking, multiprogramming, multithreading, etc. What we are claiming here is that there is a need for an analogous concept of multiperception, by which an object *X* may have multiple ways of being perceived (i.e., represented). This is not the trivial capability

of multiple instantiations of representations for X , but instead of multiple modes of representing X .

This capability entails yet another problem: if a system has multiple ways of describing the objects, what is to distinguish the representation given to a specific one? If there is no intrinsic (and objective) representation $P(X)$ always applicable to the perception-mapping of a particular X , how is it to be decided? This leads us to the *context factor*. In Bongard problems, the perception mapping $P(X)$ of an object X is also a function of the context on which X is embedded—that is, $P(X)$ by itself does not make sense, for there is no God’s eye view of X . If we must think in terms of a function, the appropriate function would be $P(X, S)$, where S is the context (scene) to which X belongs. It does not make sense to ask for $P(X)$ from X alone, there is a need for contextual information that lies outside of X . At this point, this is the view of perception as a one-to-many relation, from one object to its many potential descriptions.

But there is more to the story. As Hofstadter reminds us, “we must see a thing *as another thing*” [13]. In cognitive psychology, the notion that seeing includes “seeing as” is more broadly disseminated (see, for instance, [10]). A very basic example: two lights are flashed in quick succession. What does the subject see? A single moving point. Subjects perceive as a single moving light what in fact consists of two distinct lights. We all know what that is like, for we can not see the still pictures of video *as they really are*.

In Bongard problems, this is a notorious effect to be found ubiquitously.

Let us look at some of them. A simple object such as a line segment can be represented as something different from a line segment. Consider, for instance, BP#96, where there is not, mathematically speaking, a single triangle in the problem. There are only sets of lines, and yet the right answer, by any human standard, is “triangles versus quadrangles”. Bongard used this kind of arrangement many times; in BP#97 there is a triangle made out of circles, and a circle made out of triangles (among other “false” triangles and “false” circles). Other interesting arrangements are the “zigzag triangles” and “zigzag rectangles” of BP#10. The computational capability to flexibly represent these figures as triangles, even when they are not (in mathematically definable objective terms), is imperative for the solution of Bongard problems, and is another facet of multiperception.

It is important to understand precisely what is involved in the process of “seeing as”. In fact, this may be one of the most important problems in high-level perception, as it shows that the process of perception includes a kind of “analogy-making”—that representations can “slip” (this thesis has been fully developed elsewhere [12]). When we make an analogy, we interpret a situation in terms of another. And this is what “slipping” means: we may interpret a zigzag as a line segment, or a set of triangles as a circle, or just about anything as another thing. So, when the zigzag slips into a line segment, we not only interpret it as a line segment, but also project onto it all the properties of a line segment, such as straightness. Though zigzags can never be straight, one could easily imagine a “straight zigzags versus curved zigzags” Bongard problem—and this shows how deep slippages go.

And this is not just an issue of noise, of imperfect information, for a very simple reason: when we *see as*, we trade meaningful structures for other meaningful structures. We trade a set of triangles for a circle (BP#97), a set of line segments for a triangle (BP#96, BP#97), or maybe zigzags or smooth curves for a triangle (BP#10). When a circle is perceived in an arrangement of triangles, *the triangles hold per-*

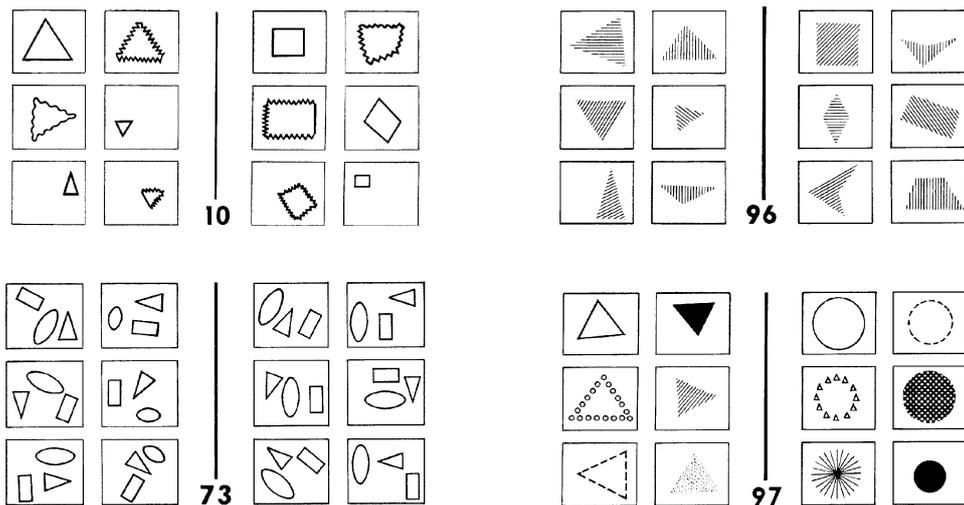


Fig. 5. Bongard problems require “seeing as”. [From M.M. Bongard, *Pattern Recognition*, Spartan Books, 1970.]

fect information—they do not contain any noise. The triangles are meaningful structures perceived with perfect information, and their collective arrangement suggests another meaningful structure, a circle. Thus, we trade a (set of) meaningful structure(s) for other meaningful structure(s), and this is different from just “an issue of noise”.

Sometimes, an arrangement should not even be characterized at all. The problem of discriminating what is relevant from what is irrelevant appears in BP#73, where the triangle distracts from the solution and should in fact be ignored, for it consists of irrelevant data to make the problem harder.

3.1. Objects cracking under pressure

A connected arrangement should not always be taken as an individual object. A glimpse at Fig. 6 will demonstrate that there is a multitude of ways of segmenting a triangle into bounded objects. The figure makes it clear that it is not always correct to perceive such a triangular structure as a single object. And since triangles are not special in any sense here, this will also apply to rectangles, stars, crosses, zigzags, and almost any other geometrical arrangement: since it is not always correct to perceive connected geometrical arrangements as bounded objects without regard to their overall context, we may conclude that the *objects of a specific problem are found during the process of segmentation*. And as soon as we take this to be correct, it follows that

Proposition 5. *In Bongard problems, objects do not exist outside of understanding.*

This is probably the most important reason for discarding the use of *a priori* representations as a genuine approach towards their solution.

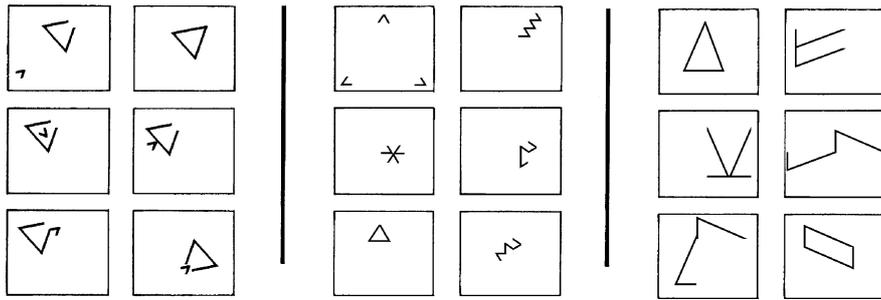


Fig. 6. Some potential segmentations of a triangular arrangement.

Should boundaries exist metaphysically in an arrangement, then these boundaries could not crack under pressure, subject to contextual influences, and there could be only one correct description of each arrangement as an object. However, as we have seen, there are multiple correct descriptions, which emerge from the context of the problem—multiple correct segmentations of a single connected arrangement into separate objects. This is what makes a symbolic representation such as `TRIANGLE(properties...)` inadequate: a segmentation as a single object does not break smoothly into the multiple potential segmentations required by Bongard problems.

And it does not make sense to say, for instance, that “threeness” is a real property of a triangle, to preserve the view that it still can be seen as a triangle, a single object holding such a property, a property which could also be found on the other boxes and might then lead to the solution: an extreme resort to threeness might, it could be argued, ultimately solve the problems BP#85 and BP#91, but, then again, such a resort would fail to solve problems holding the segmentations into other-than-three objects, such as the leftmost segmentation of Fig. 6. And, for instance, concerning the leftmost segmentation of Fig. 6, the corresponding argument that “a triangle holds the property of twoness” would be meaningless, as there is no *a priori* basis for such argument. In the leftmost segmentation of Fig. 6, there is not one triangle, but really two objects that happen to meet in a way such that a triangular arrangement emerges from their conjunction.

Let us not confuse the issues here. There is obviously a structure that “affords to be described as a triangle”. But *from this fact only* it does not follow that *it is a triangle in all cases*. Even if our first (and best) impression of it is as a triangle, that object does not exist outside of a general context, outside of the understanding of its being a triangle. In fact, the definition of what it is for something to be a triangle does not exist in a realm outside of understanding [17]. That structure “affords to be described as a triangle”, because there are people in the world that “afford to actually describe it as a triangle”. Were it not for our presence, the triangle would continue to exist *as a particular pattern*, but not *as a single bounded entity*.

Fig. 7 makes it clear. How many objects are there in each box? Since the boxes are taken absolutely out of context, there is no basis whatsoever to find an answer (and, for some, no basis on which to distinguish figure from ground). There are ambiguous figures, such as the

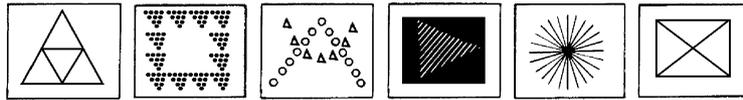


Fig. 7. How many objects are present in the boxes? Is there a single correct solution?

two triangles/three triangles/four triangles; and figures on which objects emerge from other arrangements, such as the rectangle/12 triangles/120 circles. Another is the scratched black rectangle/many line segments/one triangle. Under the proper context, any of these views could be correct. Hence, in order to understand these figures as made up of individual objects, we must have a general context on which to describe them.

What are objects, then, if not single bounded entities of an external reality? Objects are, in this alternative view, *describable units of representation*. Products of cognition, such that, when we actually perceive a triangle, even if nothing slips and we see a *perfect* triangle, we never happen to “just see it” (i.e., to find its correct description), we instead perceive the geometrical arrangement “as a triangle” (i.e., we look for a segmentation which affords that description). Even when nothing slips, we are always “seeing as”.

Thus it seems that we are required to deny, in the case of Bongard problems, the metaphysical, *a priori*, existence of objects in the first place. The raw binary image must be segmented into clear bounded objects, and the “one object, multiple views” hypothesis does not stand up to this critique: While we certainly have multiple potential views of a single geometrical arrangement, we do not, however, have clearly demarcated objects existing outside of understanding, for objects only arise when embedded on a larger context. This leads to the hypothesis that perception may be best understood when considered from the viewpoint of the idea of “no objects, multiple views”.

But before we analyze this idea further, there is one more problem arising in metaphysical-realism-based approaches to Bongard problems.

3.2. A further problem with metaphysical realism

Besides the problem of object identification and description, there is yet one more insurmountable problem with a metaphysical-realism-based approach to Bongard problems. Metaphysical realism states that reality comes complete with a fixed set of objects, properties, and relations. In this view, objects can easily be placed in their corresponding categories; a view that implies that *there is such a fixed set of categories*. Thus, for metaphysical-realism-based projects, and in the perspective of the representation problem, there should be a fixed set of object-types to describe each and every arrangement (token) contained in a Bongard problem.

But as soon as one fixes a set of categories, it is always possible to invent yet one more category *which is not contained in the set*. Though this may seem a minor problem *for the philosophical view*, which could in fact contain an infinite set of categories, *it is insurmountable for any technology implicitly based on that view*, because of their restriction to a finite set of object-types.

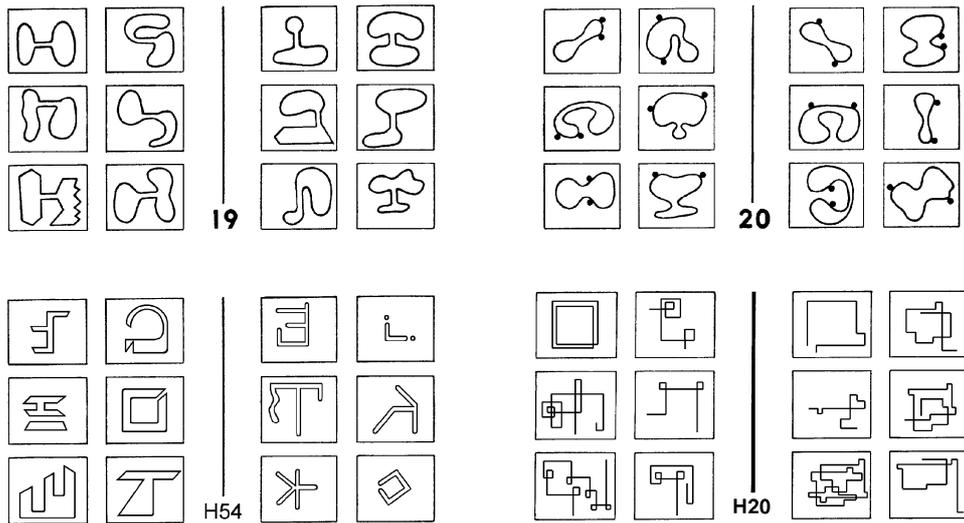


Fig. 8. How could one describe the contents of these problems over a phone line? [From M.M. Bongard, *Pattern Recognition*, Spartan Books, 1970.]

For example, consider the problems BP#19, BP#20, BP#H20, and BP#H54 below.⁴

These problems are not dealing with triangles, circles, stars, or other easily categorizable structures for which we happen to have clear concepts. Given our conceptual structures, it would be a formidable exercise to explain their content to someone over a phone line, in a context-free manner, *without having solved the problems*.

And these examples make the following remark seem obvious: *There can be no finite set of categories to name and describe the geometrical arrangements found in Bongard problems*. Even if we decide to ignore the arguments of the previous section, and accept that a triangular arrangement could be “traded” for an *a priori* symbolic representation, even then, there could be no finite set of symbolic representations to categorize the arrangements found over Bongard problems. As soon as one fixes a set of categories, there will be structures which could not fit into any of those predefined categories.

4. No objects out there, multiple views in here

We have thus discarded the idea of perception as a function from external objects to internal descriptions, and it now seems that we must conceive it as a many-to-many relation: A single segmentation (identifying an object) could map to multiple potential descriptions; a single description could arise from multiple potential segmentations. Also, as the problem of segmentation makes it clear, we must let the concept of external object

⁴ Problems BP#H20 and BP#H54 come from Douglas Hofstadter’s “56 new Bongard problems”, unpublished manuscript, 1977 (available through the Center for Research on Concepts and Cognition, Indiana University, Bloomington).

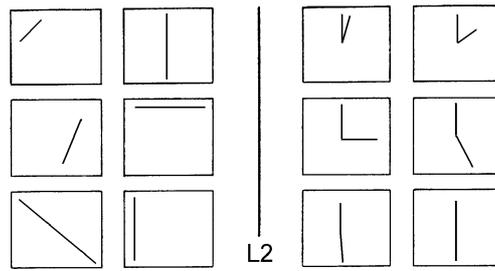


Fig. 9. Bongard problem BP#L2.

go. To favor a new term, in this paper we use “worldly arrangements”, or “geometrical arrangements”. (It should be noted that this term holds connotations of *discreteness*, which is not a problem when dealing with Bongard problems; however, such connotations cannot be assumed if one is investigating the true fabric of reality, as in [29]—and for which the rather fancy term “deictic flux” has been used.)

We may conclude that the only methodological approach in order to genuinely address Bongard problems is to provide the system with raw non-interpreted data in the first place, and that any symbolic information describing an *a priori* object existing in any given problem would consist of an implicit assumption of metaphysical realism, and thus lead to insurmountable problems. We have also seen that, since the problem of perception involves two interrelated and inseparable processes (segmentation and description), perception is thus a many-to-many relation, not a function with a clearly defined input consisting of metaphysically external objects.

Proposition 6. *There are no descriptions independent of segmentations, as if there were ready-made pre-segmented objects.*

A representation $P(X)$ is not simply a process of description of X , but also a segmentation of the world into an X in the first place, as there is no pre-segmented X out there. It is this idea, that a representation consists of a segmentation-description, intrinsically inseparable, that one must live by in order to obtain a satisfactory approach towards solving Bongard problems.

Proposition 7. *For each specific segmentation, there might be multiple potential descriptions.*

There are many examples of this. For instance, the single segmentation of a circle may be described, in the proper context, as a “closed curve”, or as “convex”, or as “large object”, etc. It has been noticed⁵ that the shape of a single circle in a box appears on BP#01, BP#02, BP#04, BP#05, BP#08, BP#09, BP#11, BP#12, BP#17, BP#21, BP#23, BP#31, BP#33, and BP#97 of Bongard’s original collection. This should be enough evidence to support Proposition 7.

⁵ Harry Foundalis (personal communication, April 2000).

Proposition 8. *For each specific description, there may be multiple segmentations.*

This should not occur with high frequency, but there is at least an existence proof, given by the following example: in BP#L2, the description *line segment* applies to both of the matching boxes, class 1 has *one* line segment, while class 2 has *two* line segments. In fact, this problem stretches the definition of Bongard problems a little, as *the same box appears on both sides*.

Bongard has never gone so far as to present the very same box on both sides of a problem, but that does not mean that it can not be done. After all, there are problems for which the solution is simply “one object versus three objects”, and the very same box might fit both interpretations (such as the “one triangle versus three line segments that just happen to meet by coincidence”, presented over problems BP#6, BP#85, BP#87, BP#91). As long as a meaningful distinction can be observed, the problem is still valid. This possibility is, however, quite limited, and the challenge provided by problems of this type is not nearly as great as that of a general Bongard problem.

When it comes to Bongard problems, the cherished intuitions of a world clear cut into *a priori* objects must go. Perception is not a one-to-one function, but a many-to-many relation. Objects, such as line segments, do not exist in an *a priori*, metaphysical, sense. Context not only plays a major role, but it absolutely dominates the search for answers. Perceptual processes interfere strongly with high-level cognition, and vice-versa. There are multiple potential descriptions. There are multiple potential segmentations. For each specific segmentation, there might be multiple descriptions. For each specific description, there might be multiple potential segmentations. These are some of the very reasons *why* Bongard problems are so hard in the first place, and why progress has been so slow in this arena.

Related analyses in the AI literature are, for instance, the incisive view presented in [28] of the much hyped project CYC [18], or the critical look provided in [3] of the *structure mapping engine* [5]. Some of the issues discussed here have been previously brought by [12,14]. On the side of philosophy, alternatives to metaphysical realism—which discard the idea of a metaphysically external object—have been proposed in [15,16,24,29].

4.1. *Objects as a fruit of cognition*

We should now put Bongard problems aside and face the broader philosophical issue. Is all of perception like this? Are there objects existing in reality outside of understanding? Or are we creating objects by imposing boundaries between regions of the world, and they are in fact our *units of description* of such a world?

It is clear that the arguments put forth here are restricted to Bongard problems, and should not apply directly to the very fabric of the world. However, a new philosophy has been emerging, a *philosophy of presence*, which considers an object not as an *a priori* external bounded entity, but really as a fruit of cognition, as emerging in the interaction with subjects, as a fruit of the *dance between s-regions (subjects) and o-regions (objects)* [29]. But maybe object is too much of a loaded word; and perhaps *unit* is a better term.

In this metaphysical overhaul, a new world is presented, a crucially continuous world (based on field theory of physics), which is referred to as a *deictic flux*. In that world, the boundaries between units (objects) are intrinsically artificial, and the argument does indeed go deep. Take, for instance, a discussion concerning the vacuum, the ultimate notion of empty space—pure nothingness—which, if taken to exist, could sustain a metaphysical notion of boundary. “Empty space, it is said, is not really empty if you look *very very hard*. Instead it everywhere and always (to say nothing of already) boils and bubbles, toils and troubles, with countless millions of subatomic particles and their antimatter opposites seething in a somewhat random but thoroughly intermixed pattern of activity” [29, p. 329]. When one seriously entertains this notion, that empty space is not really empty, that *even this case* does not constitute a clear metaphysical boundary between potential units in the world, one starts to perceive the alternative view, that objects, and their boundaries for that matter, arise from perceptual processes.⁶

But one does not have to go so deep to find such examples. Sometimes, units just fail to appear. Consider, for instance, the zebra stripes. Nature has come up with such outstanding patterns because it brings zebras a critical adaptive value. What exactly is that adaptive value? The answer, a very well-known fact of biology, is that it helps to make life miserable for predators such as lions, which cannot, over the reckless ride that characterizes predation in the Savannah, single an individual out of those confusing visual patterns emanating from the zebras. Lions cannot, or at least have a miserable time in trying to, in the terminology of [29], *register a zebra as an individual*, of correctly segmenting all that confusing information into a clear unit. And thus lions are destined to perceive a multitude of animals suddenly transform into a single one, which subsequently metamorphoses back into a bunch of them, and then back again, in the course of each and every hunt for this menu item. Thoughts like these bring to mind that famous remark: “*If lions could speak, we would not understand them*”—Ludwig Wittgenstein.

4.2. Particularity and individuality

A key notion of this new metaphysics is the distinction between *particularity* and *individuality* [29]. A sand dune has a particular structure, and so does a cloud in the sky, or a wave in the sea. There is *particularity*. However, there is no individuality: we cannot count the sand dunes in the desert, or the waves in the ocean, or the clouds in the sky. These are cases where there is particularity, *but no individuality*. (One could argue that basically, particularity concerns how the world is actually set up, while individuality concerns how we take it to be set up.)

Think of a blind man who has just undergone surgery, and can now “see”. We all know that at first he cannot see at all: show him an apple, and he will not perceive it as a bounded unit until he has a chance to touch it. Once again, units fail to appear. He has access to all the visual information and to the clear visual boundaries presented by the apple, but he cannot individuate it. And this is not just because he is still learning to focus and to use

⁶ The reader should by no means assume that this is all there is to the philosophy of [29]—see, for instance, the review in [20].

stereo vision: it is because he lacks the visual conception that defines what an apple as a unit looks like. He has access to particularity, but there is yet no individuality.

This distinction could, in fact, be used to draw a line between most of AI research [19].

There is individuality-only research, in what might be referred to as the *unity world* [19]. This is a world devoid of sounds, smells, and imagery; a world clear-cut into unambiguous bounded objects and nothing else. Some projects that are representative of this arena are chess playing programs, CYC [18], expert systems, and RF4.

On the other hand, there is particularity-only research, in what might be referred to as the *wave world* [19]. Some of the representative projects here are projects that take, for instance, an image as input and obtain other images as a result, as in stereo vision, or in shape from shading. This is a world rich in particularity, and the algorithms devised are generally taken to be independent of the units residing in the images.

And this is a major point about Bongard problems, for, as we have seen, they actually *demand* that the line from particularity (their two-dimensional square waves) to individuality (bounded units) be crossed. There is information in Bongard problems that can only be found in the unity world; but Bongard problems do not come pre-packaged in clear units—and one of the core computational requirements is that of finding the right description of the patterns into units. The input is a particular, and the output is individualized—the line between these worlds thus *must* be effectively crossed [19].

Not that this cannot be done; there are AI projects which are indeed capable of, given a particular structure, considering multiple potential descriptions and multiple potential segmentations, constantly influencing each other in a process that enables the systems to shift conceptual schemes, in order to finally decide on an individual description of that particular.

Consider, for instance, the pioneering speech-understanding program HEARSAY-II [4], which is rich in particularity and in individuality. Hearsay II starts its processing with the pure particularity of a sound wave, and gradually crosses the line to individuality, building, in a mix of bottom-up and top-down processing, the representation of an utterance [4]. It lives neither solely on the wave world nor on the unity world—and it plays back and forth between them. This remarkable project has lately been used as inspiration for innovative perception-based projects in intelligence theory (see [7,12,21–23,25]).

4.3. On the apparent stability of objects

We deal on a daily basis with what seems to be stable objects: a rock, an apple, a zebra, all of them are, to us, separable units. They behave as if they had clear boundaries, a definitive separation of what is rock from what is world-outside-rock. However, this clear individuality might be residing in our own cognition, and not on the world outside of understanding: if we look at the extreme perspectives of physics, we see no clear boundary whatsoever, existing outside of understanding, as atoms, planets and galaxies can also be seen as particular wave fields. Additional pressure on the metaphysical existence of individual bounded entities comes from applying a greater time scale: if we decide to take the perspective of millions of years, *that one* rock starts to look fluid, and it becomes harder and harder to delineate precisely at any given moment what exactly is *that one* rock from *not-that-one-rock*. Where are those stable boundaries, a million years from now? In fact,

is it really such a metaphysically stable object in the first place? Or is it an object *for us*, a part of the world which we have segmented as a single entity, *a unit of description* for our day-to-day use? Ontological questions such as these place an enormous pressure for the foundation of alternatives to metaphysical realism (for some attempts see [15,16,24, 29]). And as we have seen, Bongard problems do not succumb easily to this philosophical doctrine.

The Greek philosopher Heraclitus (ca 540–ca 480 B.C.) took the view that “everything flows and nothing abides; everything gives way and nothing stays fixed”. His words should strike us as truly outstanding, as he lived at a time so remote from our own that he could never imagine how profoundly modern science would ultimately confirm this vision. Here is a trivial sampling of what he did not know: for one, the earth moves. It was born 4.5 billions of years ago. It will certainly cease to exist in the far future. Species evolve, sometimes drifting apart towards specialization, and, other times, meshing into a series of subtle gradations. Islands emerge from underwater volcanoes, only to gradually submerge back into the ocean. Gigantic whales develop from a single, invisible, cell. DNA splits in two parts. Mountains rise. Atoms divide. Continents drift. Galaxies collapse. This list is endless—and reality is indeed in flux.

“You can never step into the same river twice”, he said. This vision—profoundly confirmed by modern science—puts an enormous pressure on the philosophical conception that reality comes complete with a fixed set of units. This widely held view is not necessarily true, and maybe it is time to properly consider the alternative: a view of the worldly arrangements existing as sand dunes, in a continuous flux, coming and going, being gradually formed, those forms existing temporarily, and then ceasing to exist, metamorphosing into other arrangements. Maybe there is absolutely no clear metaphysical demarcations between what we take to be as units, just as there are no clear boundaries between sand dunes. We can sit and watch a sand dune take form as the wind blows, but we can never delineate precisely the point where it starts to exist as an unit, and neither can we determine the very first second of its existence. Maybe this also applies to all those objects we see and talk about; perhaps these seemingly stable objects are not really single bounded entities of an external reality, but instead our own units of description of a reality undergoing such a severe transformation.

This view should be distinguished from that of “reality being a subjective construction of the human mind”, as if the sun really revolved around the Earth before Galileo. This is not the case. There is a strong commitment to the existence of a reality, which will continue to exist even if we close our eyes to it. What this vision suggests, however, is that, instead of external single bounded entities, objects should be seen as *internal units of description*. The clear segmentation of the world into separable units, the imposition of boundaries between those units—the individuation of particularity—is a job for perceptual processes, and not just in the case of Bongard problems. This is in fact an understanding that when an astronomer talks about galaxies (rather than wave fields), or a biologist talks about zebras (rather than a multitude of cells, or even of their constituting parts), these scientists are adopting a conceptual scheme, and their objects (and corresponding boundaries), emerge *from the adoption of a conceptual scheme* [15,24].

This is the question of, given a complex particular pattern, selecting the optimum level of its description—which is, incidentally, the very challenge posed by Bongard.

We are still a long way towards understanding precisely the nature of perception and synthesizing it on a machine. We can only hope for the day when a machine will be able to perceive the deep meaning of all those patterns we are involved with, be them sand dunes, Gödelian theorems, musical styles, chess threats, or esthetical judgements. Such a machine will undoubtedly have the capability to segment and describe worldly arrangements in potentially distinct ways—and settle for one according to context—as we have referred to it here, *multiperception*. It must be able to go from the rich particularity of our world to the internal individuality of its description. It is simply inconceivable that it will not, as this is not a merely accidental obstacle towards the automation of the solution of Bongard problems, but a core cognitive capability that shapes human existence, and it clearly belongs to the fundamental nature of intelligence.

Acknowledgement

I would very much like to thank the people involved in reviewing all those previous versions of this manuscript: besides the four anonymous referees, José Ricardo Torreão, Douglas Hofstadter, and Harry Foundalis have contributed to this paper to a large extent. I also thank the encouragement provided by Melanie Mitchell of the Los Alamos National Laboratory; and I would like to recognize Dr. Horacio H. Yanasse and the whole research community of the National Space Research Institute of Brazil, for a much needed freedom, and for such a great scientific environment, respectively, both so hard to be found in an emerging economy. And finally, this work could not have been accomplished without the financial support provided by the FAPESP Foundation.

References

- [1] M.M. Bongard, Pattern Recognition, Spartan Books, New York, 1970.
- [2] R. Brooks, Intelligence without representation, Artificial Intelligence 47 (1991) 139–160.
- [3] D. Chalmers, R. French, D. Hofstadter, High-level perception, representation, and analogy: A critique of artificial intelligence methodology, J. Experiment. Theoret. Artificial Intelligence 4 (1992) 185–211.
- [4] L.D. Erman, F. Hayes-Roth, V.R. Lesser, D.R. Reddy, The HEARSAY-II speech understanding system: Integrating knowledge to resolve uncertainty, Computing Surveys 12 (1980) 213–253.
- [5] B. Falkenhainer, K.D. Forbus, D. Gentner, The structure-mapping engine: Algorithm and examples, Artificial Intelligence 41 (1990) 1–63.
- [6] R.M. French, Subcognition and the limits of the Turing test, Mind 99 (393) (1990) 53–65.
- [7] R.M. French, The Subtlety of Sameness, MIT Press, Cambridge, MA, 1995.
- [8] R.M. French, When coffee cups are like old elephants, or, why representation modules don't make sense, in: A. Riegler, M. Peschl (Eds.), Proc. 1997 International Conference on New Trends in Cognitive Science, 1997, pp. 158–163.
- [9] R.M. French, P. Anselme, Interactively converging on context sensitive representations: A solution to the frame problem, Revue Internat. Philosophie 53 (1999) 365–385.
- [10] A. Gilchrist, I. Rock, Rational processes in perception, in: Proc. 3rd Annual Conference of the Cognitive Science Society, Berkeley, CA, 1981, pp. 50–56.
- [11] D. Hofstadter, Gödel, Escher, Bach: An Eternal Golden Braid, Basic Books, New York, 1979.
- [12] D. Hofstadter, Fluid Concepts and Creative Analogies, Basic Books, New York, 1995.
- [13] D. Hofstadter, On seeing A's and seeing As, Stanford Humanities Review 4 (2) (1995) 109–121.
- [14] B. Indurkha, Metaphor and Cognition: An Interactionist Approach, Kluwer, Norwell, MA, 1992.

- [15] G. Lakoff, *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*, Chicago University Press, Chicago, IL, 1987.
- [16] G. Lakoff, M. Johnson, *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, Basic Books, New York, 1999.
- [17] G. Lakoff, R. Núñez, *Where Mathematics Comes from: How the Embodied Mind Creates Mathematics*, Book manuscript, forthcoming.
- [18] D.B. Lenat, E.A. Feigenbaum, On the thresholds of knowledge, *Artificial Intelligence* 47 (1991) 185–250.
- [19] A. Linhares, The topology of microworlds, Working Paper, National Space Research Institute of Brazil, in preparation.
- [20] A. Linhares, *Metaphysics incorporated*, Unpublished manuscript, National Space Research Institute of Brazil, 2000.
- [21] J.B. Marshall, *Metacat: A self-watching cognitive architecture for analogy-making and high-level perception*, Ph.D. Dissertation, Indiana University, Bloomington, IN, 1999.
- [22] G. McGraw, *Letter Spirit (part one): Emergent high level perception of letters using fluid concepts*, Ph.D. Dissertation, Indiana University, Bloomington, IN, 1995.
- [23] M. Mitchell, *Analogy-making as Perception*, MIT Press, Cambridge, MA, 1993.
- [24] H. Putnam, *Reason, Truth, and History*, Cambridge University Press, Cambridge, 1981.
- [25] J. Rehling, *Letter Spirit (part two): Automating creative design in a visual domain*, Ph.D. Dissertation, Indiana University, Bloomington, IN, forthcoming.
- [26] K. Saito, R. Nakano, A concept learning algorithm with adaptive search, in: K. Furukawa, D. Michie, S. Muggleton (Eds.), *Machine Intelligence 14—Applied Machine Intelligence*, Oxford University Press, Oxford, 1995, pp. 347–363.
- [27] K. Saito, R. Nakano, Adaptive concept learning algorithm, *IFIP Transactions A—Computer Science and Technology* 51 (1994) 294–299.
- [28] B.C. Smith, The owl and the electric encyclopedia, *Artificial Intelligence* 47 (1991) 251–288.
- [29] B.C. Smith, *On the Origin of Objects*, MIT Press, London, 1996.