Pattern-Recognition: a Foundational Approach

Agnese Augello[3], Salvatore Gaglio[1,3], Gianluigi Oliveri[2,3], and Giovanni Pilato[3]

[1] DICGIM - Università di Palermo
Viale delle Scienze, Edificio 6 - 90128, Palermo - ITALY
[2] Dipartimento di Scienze Umanistiche - Università di Palermo
Viale delle Scienze, Edificio 12 - 90128, Palermo - ITALY
[3] ICAR - Italian National Research Council
Viale delle Scienze - Edificio 11 - 90128 Palermo, Italy
{gaglio,oliveri}@unipa.it
{agnese.augello,giovanni.pilato}@cnr.it

Abstract. This paper aims at giving a contribution to the ongoing attempt to turn the theory of pattern-recognition into a rigorous science. In this article we address two problems which lie at the foundations of pattern-recognition theory: (i) What is a pattern? and (ii) How do we come to know patterns? In so doing much attention will be paid to tracing a non-arbitrary connection between (i) and (ii), a connection which will be ultimately based on considerations relating to Darwin's theory of evolution.

1 Introduction

As is well known, the main aim of pattern-recognition theory is to determine whether, and to what extent, what we call 'pattern-recognition' can be accounted for in terms of automatic processes. From this it follows that two of its central problems are how to: (i) describe and explain the way humans, and other biological systems, produce/discover and characterize patterns; and how to (ii) develop automatic systems capable of performing pattern recognition behaviour.

Having stated these important facts, we need to point out that at the foundations of pattern-recognition theory there are two more basic questions which we can formulate in the following way: (a) what is a pattern? (b) how do we come to know patterns? And it is clear that, if we intend to develop a science of pattern recognition able to provide a rigorous way of achieving its main aim, and of pursuing its central objects of study, it is very important to answer questions (a) and (b).

After having addressed the problem of providing a definition of the concept of pattern in §2, a case-study of a particular type of finite geometry is discussed in §3 in the hope that by so doing we might obtain a rigorous characterization of the concept of mathematical pattern.

Section 4 is then dedicated to the examination of some of the interesting lessons that can be learned from the case-study in §3. In particular, one of these has to do with the characterization of the concept of mathematical pattern in terms of mathematical structure; and another concerns the possibility of generalizing the view of mathematical patterns as structures to patterns belonging to fields different from mathematics.

Finally, sections 5 and 6, armed with the notion of pattern developed so far, bring the paper to a close by addressing question (b) above: how do we come to know patterns?

2 Searching for a definition

A potentially fruitful approach to the problem 'What is a pattern?' is that of Daniel Dennett. For Dennett, who in his discussion of the concept of pattern is concerned with issues belonging to the philosophy of mind and action,

> [W]e are to understand the pattern to be what Anscombe called the "order which is there" in the rational coherence of a person's set of beliefs, desires, and intentions. [[6], §IV, p. 47.]

However, although taking into account final causes, beliefs and intentions often can both reveal an order existing among a certain individual's actions and explain his behaviour in terms of giving an account not only of how, but also of why he did what he did, it must be admitted that talking about 'the order which is there in the rational coherence of a person's set of beliefs, desires, and intentions' is too vague to shed light on the notion of pattern. This is, in particular, the case when the accounts of the order which is there . . . etc. are several, radically differ from one another, and all seem to agree with the facts.

Moreover, since patterns do not occur only within the context of human actions and beliefs, what happens when we are dealing with patterns displayed by crystals of snowflakes? Of course, also in the case of crystals of snowflakes (see Fig. 1)

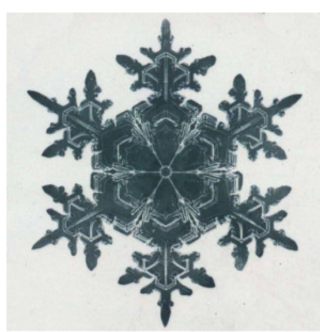

Fig. 1. A crystal of a snowflake

the patterns they display are related to the order in which the components of the crystals of snowflakes are to one another. But, whereas in the case of the crystals of snowflakes, if we use a microscope, we can actually see them, when we turn to actions the verb 'seeing' appears to let us down. For an action, in contrast to the crystal of a snowflake, is not just a brute physical fact and, therefore, the order manifested by a sequence of actions 'which is there in the rational coherence of a person's set of beliefs, desires, and intentions' is not something we can perceive by simply keeping our eyes wide open, and using instruments of observation.

This is an important point, because, if there is some truth in Dennett's definition of pattern, it means that what we might call 'brute seeing', that is, the mere act of representing within visual perceptual space a given input — like what happens with a photo-camera when we take a picture — cannot provide a satisfactory account of what happens when we perceive a pattern.

Therefore, if we intend to give an account of perceiving a pattern which is in accord with Dennett's definition, we should appeal to a concept of seeing which is much richer than brute seeing. A good candidate for such a concept of seeing is the concept that in the *Philosophical Investigations* Wittgenstein famously called 'seeing something as' or 'aspect seeing'.[4]

Notice, for example, that in seeing something as a square the perception of the square-pattern is not brute, because it presupposes, among other things, that the observer has a grasp of the concept of square.

However, independently of questions relating to the nature of the 'order which is there . . . ' in different contexts, and any consideration concerning what we must mean by 'seeing an aspect' or 'perceiving a pattern', Dennett proposes a very interesting general test for the existence of patterns. Basing himself on Chaitin's definition of randomness:

> A series of numbers is random if the smallest algorithm capable of specifying it to a computer has about the same number of bits of information as the series itself. [[2], p. 48.]

Dennett asserts that:

> A pattern exists in some data — is real — if *there is* a description of the data that is more efficient than the bit map, whether or not anyone can concoct it. [[6], §II, p. 34.]

Although that offered by Dennett is a very plausible criterion which, in some cases, reveals the presence of patterns in a data-set, it is not specific to them. To see this consider the definite description 'The satellite of the Earth'. Such a definite description certainly provides an enormous compression of data with respect to the bit map of a computer visual representation of the Moon. But, it is a description which uniquely identifies an object not a pattern/structure.

Lastly, the phenomenon of seeing something as a square appears to hint at a structural feature of perception, where the pattern perceived is that of a square. In fact if, by zooming in or out on the object we perceive as a square, we change

[4] See on this [17], Part II, §XI, pp. 213^e–214^e.

(within a certain range) the magnitude of the picture of the object, we would still see the object as a square.

The structural character of the pattern perceived is particularly evident in the case of the crystal of a snowflake. Indeed, when we observe a crystal of a snowflake through a microscope or when we look at a photograph or at an artist's accurate impression of that very crystal of a snowflake, etc. in spite of being presented in each single case with a different object — the actual crystal, the photograph of the crystal, and the artist's accurate impression of the crystal — we recognize the presence of the same pattern *in* all these objects. Of course, the next question is 'What is a structural feature of an object?' or, in more general terms, 'What is a structure?' The latter is, indeed, the problem which is going to be at the heart of the next section.

3 Mathematical Patterns. A case study

If we are presented with objects **a** and **b** (see Figures 2 and 3), it is very difficult to see what interesting mathematical feature they might have in common, if any, let alone that they exemplify the same mathematical pattern.

A	B	C	D	E	F	G
B	C	D	E	F	G	A
D	E	F	G	A	B	C

Fig. 2. Object **a**

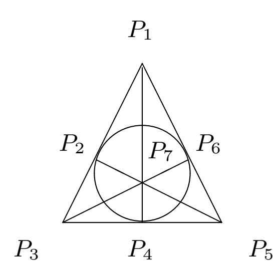

Fig. 3. Object **b**

Indeed, whereas object **a** is a 3×7 matrix whose elements are the first seven letters of the Italian alphabet, object **b** is a geometrical entity consisting of 7 lines and 7 points. The lines of object **b** are: the sides of the triangle drawn in Figure 3, the bisecting segments, and the inscribed circle. On the other hand, the 7 points are the points of intersection of three lines.

However, the situation radically changes if we introduce the following formal system T with the appropriate interpretations.

Let a formal system T be given such that the language of T contains a primitive binary relation 'x belongs to a set X' ($x \in X$), and its inverse 'X contains an element x' ($X \ni x$).

Furthermore, let us assume that D is a set of countably many undefined elements $a_1, a_2, \ldots$; call 'm-set' a subset X of D; and consider the following as the axioms of T:

Axiom 1 If x and y are distinct elements of D there is at least one m-set containing x and y;
Axiom 2 If x and y are distinct elements of D there is not more than one m-set containing x and y;
Axiom 3 Any two m-sets have at least one element of D in common;
Axiom 4 There exists at least one m-set.
Axiom 5 Every m-set contains at least three elements of D;
Axiom 6 All the elements of D do not belong to the same m-set;
Axiom 7 No m-set contains more than three elements of D.[5]

Now, the language of T contains two different sorts of variables: $x, y, \ldots$ and $X, Y, \ldots$ Let us assume that the variables $x, y, \ldots$ range over $\mathrm{D}_1 = \{\mathrm{A}, \ldots, \mathrm{G}\}$; and that the variables $X, Y, \ldots$ range over D_1^*, where D_1^* is a set whose elements are the subsets of D_1 the elements of which appear in the columns of the matrix in Figure 1, that is:

$$\mathrm{D}_1^* = \{\{\mathrm{A,B,D}\}, \{\mathrm{B,C,E}\}, \{\mathrm{C,D,F}\}, \{\mathrm{D,E,G}\}, \{\mathrm{E,F,A}\}, \{\mathrm{F,G,B}\}, \{\mathrm{G,A,C}\}\}.$$

(The elements of D_1^* are the m_1-sets.)

It turns out that $\mathrm{D}_1 \cup \mathrm{D}_1^*$ is the domain of the model of T represented in Figure 2. To see this, using the interpretation suggested above, it is sufficient to verify that **Axioms 1 − 7** are true of the matrix in Figure 2. We call such a model '$\mathcal{M}_1(\mathrm{T})$.'

On the other hand, if we change interpretation making: (a) the variables $x, y, \ldots$ range over $\mathrm{D}_2 = \{\mathrm{P}_1, \ldots, \mathrm{P}_7\}$, where $P_1, \ldots, P_7$ are the 7 distinct points indicated in Figure 3; and (b) the variables $X, Y, \ldots$ range over D_2^* whose elements are the m_2-sets, that is, the sets of three P_i points, for $1 \leqslant i \leqslant 7$, lying on the sides, the bisectrices, and the circle inscribed in the triangle represented in Figure 3: $D_2^* = \{\{P_6, P_2, P_4\}, \{P_2, P_7, P_5\}, \{P_5, P_4, P_3\}, \{P_4, P_7, P_1\}, \{P_3, P_7, P_6\}, \{P_3, P_2, P_1\}, \{P_1, P_6, P_5\}\}$; we have that $\mathrm{D}_2 \cup \mathrm{D}_2^*$ is also the domain of a model of T, a model represented in Figure 3. We call such a model '$\mathcal{M}_2(\mathrm{T})$.'

To show that $\mathcal{M}_2(\mathrm{T})$ is a model of T, it is sufficient, using the interpretation just provided, to check that **Axioms 1 − 7** are true of the object represented in Figure 3.

If we, now, compare $\mathcal{M}_1(\mathrm{T})$ with $\mathcal{M}_2(\mathrm{T})$, we realize that, among other things: (1) $(\mathrm{D}_1 \cup \mathrm{D}_1^*) \cap (\mathrm{D}_2 \cup \mathrm{D}_2^*) = \emptyset$; (2) the elements of $\mathrm{D}_1 \cup \mathrm{D}_1^*$ are not homogeneous

[5] These axioms have been taken, with some minor alterations, from [16], §2.10, p. 30.

with the elements of $\mathrm{D}_2 \cup \mathrm{D}_2^*$; and that (3) $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$, are isomorphic to each other.

With regard to point (3) above, we notice that if f is the function $f : D_1 \to D_2$ such that:

$$\begin{aligned} f(A) &= P_6, \\ f(B) &= P_2, \\ f(C) &= P_5, \\ f(D) &= P_4, \\ f(E) &= P_7, \\ f(F) &= P_3, \\ f(G) &= P_1; \end{aligned}$$

and g is the function $g : D_1^* \to D_2^*$ such that:

$$\begin{aligned} g(X) &= g(\{x_i, x_j, x_k\}) \\ &= \{f(x_i), f(x_j), f(x_k)\} \end{aligned}$$

for $1 \leqslant i \leqslant j \leqslant k \leqslant 7$, then f induces a bi-univocal correspondence between D_1 and D_2, whereas g induces a bi-univocal correspondence between the set D_1^* (of m_1-sets) and the set D_2^* (of m_2-sets).

Now, it is clear that the function ψ, where $\psi : D_1 \bigcup D_1^* \to D_2 \bigcup D_2^*$ such that:

$$\psi(\lambda) = \begin{cases} f(x) & \text{if} \quad \lambda = x \\ g(X) & \text{if} \quad \lambda = X \end{cases}$$

shows that $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$ are isomorphic to one another. In fact, ψ induces a bi-univocal correspondence between $D_1 \bigcup D_1^*$ and $D_2 \bigcup D_2^*$ preserving the two (primitive) relations $\in$ and $\ni$, that is:

$$\begin{aligned} x \in X &\text{ iff } \psi(x) \in \psi(X) \\ X \ni x &\text{ iff } \psi(X) \ni \psi(x). \end{aligned}$$

The case relative to the existence of two isomorphic models $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$ of T brings out very clearly that the pattern described by the axioms and theorems of T is independent of the nature of the objects present in $\mathrm{D}_1 \cup \mathrm{D}_1^*$ (the first seven letters of the alphabet plus ...), and in $\mathrm{D}_2 \cup \mathrm{D}_2^*$ (the seven distinct points highlighted in Figure 3 plus ...). The pattern described by the axioms and theorems of T is an abstract mathematical structure realized by/present in both $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$.

At this point a legitimate problem that might arise is 'How is the structure common to $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$ given to us?' and another is 'What sort of thing is this structure?' Let us address the second question first.

A structure/pattern is an ordered pair the first element of which is the domain of the structure — in our case $\mathrm{D}_1 \cup \mathrm{D}_1^*$ or $\mathrm{D}_2 \cup \mathrm{D}_2^*$ — and whose second element is a set of relations defined on this domain — in our case the relations are $\in$ and $\ni$ — relations the basic properties of which are implicitly defined by the axioms.

With regard to the question concerning the reality of the structure instantiated by $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$, consider that if objects **a** and **b** exist and, therefore, are real then also the structure they realize exists and, therefore, is real.

The answer to the first question is more complicated, because there is no unique way in which a pattern, even a mathematical one, becomes salient to an observer. However, it is certainly the case that necessary conditions for seeing a certain object as the realization of the pattern/mathematical structure we have been talking about in this paper are: (1) the observer's acquaintance with object **a** or with object **b**, (2) the observer's knowledge of T, and (3) the observer's knowledge of the appropriate interpretation of T.

Another non-psychological way of addressing the question 'How is the structure common to $\mathcal{M}_1(\mathrm{T})$ and $\mathcal{M}_2(\mathrm{T})$ given to us?' consists in transforming object **b** into an object **c** isomorphic to object **b** such that object **c** is clearly isomorphic to object **a** (see on this Figures 4-6). For, since isomorphism is a transitive relation this would show that object **b** is isomorphic to object **a**.

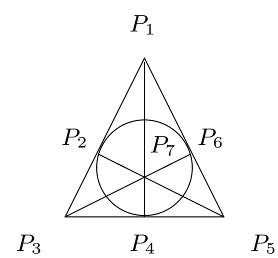

Fig. 4. Object **b**

P_6	P_2	P_5	P_4	P_3	P_3	P_1
P_2	P_7	P_4	P_7	P_7	P_2	P_6
P_4	P_5	P_3	P_1	P_6	P_1	P_5

Fig. 5. Object **c**

A	B	C	D	E	F	G
B	C	D	E	F	G	A
D	E	F	G	A	B	C

Fig. 6. Object **a**

Notice that the procedure illustrated above is non-psychological, because, although we always assume that the observer finds himself in 'normal conditions', the procedure acts on the objects observed and not on the observer. Indeed, in constructing object **c**, we have simply 'opened' **b** in such a way as to obtain a 3×7 matrix which has as columns the sets of points contained in each line. (The order in which the points actually occur in the respective lines is not relevant for our purposes.)

Several are the things that interest us in this example. We shall briefly comment on some of them in the next section.[6]

4 Some comments on the case study

Among the necessary conditions for 'seeing a certain object as ... ' that we have mentioned in the previous section the first is the observer's acquaintance with object **a** and/or with object **b**. Now, the possibility for an observer of being acquainted with **a** and/or **b** depends, among other things, on:

> [T]he particular pattern-recognition machinery hard-wired in our visual systems — edge detectors, luminance detectors, and the like ... [T]he very same data (the very same streams of bits) presented in some other format might well yield no hint of pattern to us ([6], p. 33).

Other important conditions upon which the possibility of an observer being acquainted with **a** and **b** depends are the size and position of objects **a** and **b** relative to the observer. To see this, immagine that objects **a** and **b** are microscopic and the observer is an average human being without any support provided by technology; or that **a** and **b** are too far from the observer to be surveyable by him, etc.

Secondly, in the absence of the formal system T and of the relevant interpretations of T, the observer cannot see the pattern/structure instantiated by **a** and **b**. This is because, in the absence of the formal system T and of the relevant interpretations of T, he is in no position for making the observations concerning the salient features of the pattern/structure in question, observations such as those which have to do with the part/whole distinction, etc. This shows that T, together with the relevant interpretations, does not simply power a deductive engine, but is also a system of representation.

From the considerations above, we can conclude that necessary conditions for pattern recognition in mathematics are the existence of: (1) an observer O; (2) a domain of objects D; and of (3) a system of representation Σ, i.e. $(\mathrm{O}, \mathrm{D}, \Sigma)$.[7]

Thirdly, the mathematical structure which becomes salient when we observe objects **a** and **b** *through* T depends not only on T, but also on **a** and **b** — this is where the realism concerning mathematical structures comes in. In fact,

[6] A discussion of whether mathematics as a whole is conceivable as a science of patterns/structures can be found in: [14], [10], [11], [15], [Resnik, 2001], [12], [13], [1]

[7] Actually, the system of representation Σ is an ordered pair $\Sigma = (\mathrm{T}, I)$, where T is a set containing (as a subset) a recursive set of axioms $\mathcal{A}$ and all the logical consequences of $\mathcal{A}$, and I is an interpretation of T on to D.

given that we can prove in T that there exist exactly seven elements in D and seven m-sets, if, for instance, the number of letters of the Italian alphabet we considered as elements of our matrix were different from seven, the matrix could not be a model of T (the same applies *mutatis mutandis* to the number of points of intersection of three lines in **b**).

Fourthly, we have a criterion of identity for the structure/pattern described by T, criterion of identity represented by model isomorphism, i.e., **a** and **b** instantiate the same structure, because they are isomorphic models of T. This is a very important condition, because it guarantees that the concept of structure is well defined.

Fifthly, we should notice that the definition of structure we offered in §4 — a structure $\mathcal{S}$ is an ordered pair whose first element is a domain of objects D, and second element is a set $\Re$ of relations defined on D — together with the criterion of identity for structures (isomorphism) provide both a rigorous characterization of what falls under the concept of pattern in mathematics, and the possibility of operating a natural generalization of this concept to fields different from mathematics.

With regard to the second point above, notice that both the examples of patterns examined in §3 can be accounted for in terms of structures. In the philosophy of mind and action case, the structure $S_1 = (\mathrm{D}_1, \Re_1)$ is such that D_1 contains beliefs, whereas $\Re_1$ contains relations defined on D_1 such as $\models_{\mathrm{pd}}$ — the plausible deontic consequence relation, where $B_1, \ldots, B_n \models_{\mathrm{pd}} B$ means: someone who believes $B_1, \ldots, B_n$ plausibly ought to believe B. (The turnstile $\models_{\mathrm{pd}}$ is typical of a non-monotonic logic.)

The case of a structuralist account of patterns displayed by crystals of snowflakes (see Fig. 1) is even simpler than that discussed above. The pattern/structure $S_2 = (\mathrm{D}_2, \Re_2)$ of a crystal of a snowflake consists of a domain D_2, the elements of which are the molecules of water contained in the snowflake, and of a set $\Re_2$ whose elements are the physical laws determining how the molecules of water in D_2 are related to one another in the crystal.

But, of course, if the definition of structure we offered in §4 is applicable to both the examples of patterns examined in §3, so does also the identity condition for structure: structure isomorphism.

From here on, as a consequence of what we have been arguing so far, we are going to consider the two words 'pattern' and 'structure' as synonyms.

5 Patterns' morphogenesis and cognitive architectures

If we consider the pattern/structure instantiated in object **a** (Fig. 2), we realize that this is a complex entity composed out of simpler entities. The simplest, or atomic entities, are the first 7 letters of the Italian alphabet $\mathrm{A}, \mathrm{B}, \ldots, \mathrm{G}$, and then we have the molecular entities represented by the subsets of three elements of the set $\{\mathrm{A}, \mathrm{B}, \ldots, \mathrm{G}\}$ which appear as the columns of the 3×7 matrix in Fig. 2.

Notice that the atomic entities mentioned above can be thought as patterns/structures of points, as is shown by observing the obvious isomorphism existing among any two of the following different objects:

$$\text{A}, \text{A}, \text{A}, \text{A}, \text{A}.$$

Moreover, molecular expressions such as $\{\text{A}, \text{B}, \text{D}\}, \{\text{B}, \text{C}, \text{E}\}, \ldots, \{\text{G}, \text{A}, \text{C}\}$ (the columns of the matrix) can also be seen as patterns of patterns. Indeed, the structural rôle of these three-element sets (of patterns) is revealed by the fact that they are obviously isomorphic to the following three-element sets of patterns: $\{a, b, d\}, \{b, c, e\}, \ldots, \{g, a, c\}$.

All these considerations lead us, in a very natural way, to speak of a morphogenetic process which, starting from atomic patterns $\text{A}, \text{B}, \ldots, \text{G}$ (patterns of type 0), produces molecular patterns $\{\text{A}, \text{B}, \text{D}\}, \{\text{B}, \text{C}, \text{E}\}, \ldots, \{\text{G}, \text{A}, \text{C}\}$ (these are patterns of type 1, because their elements are patterns of type 0), molecular patterns which then give origin to the pattern realized in object **a** (Fig. 2). (The latter is a pattern of type 2, because its elements are patterns of type 1).

Now, from the brief account of the patterns' morphogenetic process described above, it should be clear that such a process is capable of generating patterns of arbitrarily large complexity. Therefore, to answer the problem 'How do we come to know patterns?' on the part of a finite cognitive system which is dependent on a limited amount of resources, resources for which he is in competition with other finite cognitive agents, we are going to suggest that such an agent must be endowed with a biologically inspired cognitive architecture (described in §6) which consists of different systems for the representation and the manipulation of information.

To see this, let $\mathcal{A}, \mathcal{B}, \ldots, \mathcal{G}$ be the shortest neural network algorithms for the recognition of $\text{A}, \text{B}, \ldots, \text{G}$, within the set of the alphabet letters $\{\text{A}, \text{B}, \ldots, \text{Z}\}$.[8] The shortest neural network algorithm for the recognition of $\{\text{A}, \text{B}, \text{D}\}$ will have a length much longer than the sum of the lengths of $\mathcal{A}, \mathcal{B}$ and $\mathcal{D}$, because, among other things, lacking a concept of set, our neural network will have to treat $\{\text{A}, \text{B}, \text{D}\}$ as a plurality of individual patterns and, if we exclude pluralities containing repetitions of letters such as $\{\text{A}, \text{A}, \text{B}\}$, etc., our algorithm will have to deal with a domain D represented by the power set of $\{\text{A}, \text{B}, \ldots, \text{G}\}$ which contains 2^7 elements.

Furthermore, the next step, that is, the recognition of **a**, becomes already computationally onerous. For, if $\mathcal{ABD}, \mathcal{BCE}, \ldots, \mathcal{GAC}$ are the shortest neural network algorithms for the recognition of, respectively, the following patterns: $\{\text{A}, \text{B}, \text{D}\}, \{\text{B}, \text{C}, \text{E}\}, \ldots, \{\text{G}, \text{A}, \text{C}\}$, the length k of the shortest neural network algorithm for the recognition of **a** will be quite formidable, because, having to recognize **a** out of 7! possible 3×7 matrices the columns of which are the possible permutations of $\{\text{A}, \text{B}, \text{D}\}, \{\text{B}, \text{C}, \text{E}\}, \ldots, \{\text{G}, \text{A}, \text{C}\}$, k will be much greater than the sum of the lengths of $\mathcal{ABD}, \mathcal{BCE}, \ldots, \mathcal{GAC}$.

[8] We mention here neural network algorithms, because such algorithms are so far the most basic biologically inspired general procedures for pattern-recognition.

But, of course, in order to individuate the relevant structure realized in **a**, we should now concatenate to our neural network algorithm for the recognition of **a** another neural network algorithm of length k^* for the individuation of the isomorphism inducing function $\psi : D_1 \bigcup D_1^* \to D_2 \bigcup D_2^*$ (see §3). And, since both $D_1 \bigcup D_1^*$ and $D_2 \bigcup D_2^*$ contain 14 elements each, our algorithm will have to recognize ψ out of a set of 14^{14} functions. A tall order indeed!

All these considerations make us suspect that if a finite cognitive agent dependent on a limited amount of resources, resources for which he is in competition with other finite cognitive agents, has in its cognitive architecture systems for the representation of information which use only neural networks, it could not go very far in its pattern recognition activity. And this would not be a consequence of the fact that there are certain patterns for which in principle there is no neural network based algorithm capable of recognizing them, but of the consideration that these algorithms, if they exist, would have to be unfeasibly long, given the computational limitations of our agent.

6 The cognitive architecture. An evolutionary account.

Given what we said in the previous section about the connection existing between patterns' morphogenesis and the cognitive architecture of a finite cognitive agent who is dependent on a limited amount of resources, resources for which he is in competition with other finite cognitive agents, in what follows in this section we are going to illustrate a cognitive architecture (see figure 7) consisting of three levels of information-representation: a *subconceptual level*, in which data coming from the environment (sensory input) are processed by means of a neural network based system; a *conceptual level*, where data are represented and conceptualized independently of language; and, finally, a *symbolic level* which makes it possible to manage the information through symbolic/linguistic representations and computations.

Notice that all three levels for the representation and processing of information mentioned above are present in humans, and that the first two levels may be found in most higher animals, etc.

We have already come across the sub-conceptual level of representation (the Sub-conceptual Tier) in §5 when we discussed the possibility of recognizing type 0 patterns (atomic patterns) by means of algorithms based on neural networks. What we need to do now is providing a brief description of the conceptual and symbolic levels of representation of the cognitive architecture sketched in Figure 7.

The conceptual level of the cognitive architecture of our agent consists of the so-called 'Gärdenfors conceptual spaces'. According to Gärdenfors, conceptual spaces are metric spaces which represent information exploiting geometrical structures rather than symbols or connections between neurons. This geometrical representation is based on the existence/construction of a space endowed with a number of what Gärdenfors calls 'quality dimensions' whose main func-

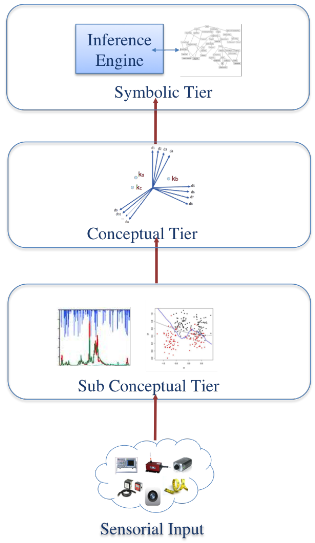

Fig. 7. A sketch of the cognitive architecture

tion is to represent different qualities of objects such as brightness, temperature, height, width, depth.

Moreover, for Gärdenfors, judgments of similarity play a crucial role in cognitive processes and, according to him, the smaller is the distance between the representations of two given objects (in a conceptual space) the more similar to each other the objects represented are.

For Gärdenfors, objects can be represented as points in a conceptual space, points which we are going to call 'knoxels',[9] and concepts as regions (in a conceptual space). These regions may have various shapes, although to some concepts — those which refer to natural kinds or natural properties — correspond regions which are characterized by convexity.[10] According to Gärdenfors, this latter type of region is strictly related to the notion of prototype, i.e., to those entities that may be regarded as the archetypal representatives of a given category of objects (the centroids of the convex regions).

Finally, the symbolic level (the Symbolic Tier) of the cognitive architecture consists, instead, of language-based systems of information representation and computation.

[9] The term 'knoxel' originates from [7] by the analogy with "pixel". A knoxel k is a point in Conceptual Space and it represents the epistemologically primitive element at the considered level of analysis.

[10] A set S is *convex* if and only if whenever $a, b \in S$ and c is between a and b then $c \in S$.

To see the three levels of the cognitive architecture at work, and assess their relative merits, consider the following problem: to recognize the pattern exemplified by object **A**.

If we assume that the algorithms that follow can all be expressed in a given language $\mathfrak{L}$, then the advantage of using algorithms based on tools characteristic of the sub-conceptual level (neural networks) to solve the problem above is that ... an algorithm is better than nothing! On the other hand, the obvious disadvantage is that neural network based algorithms can be relatively long.

Imagine now a 2-d Gärdenfors conceptual space, CSA, related to the letters of the alphabet. This is a CSA tessellated by means of prototypes of such letters using the well-known Voronoi's procedure. The pattern-recognition algorithm relating to **A** is quite simple: determine to which of the finitely many points belonging to CSA which represent the prototypes of the letters of the alphabet the point representing **A** in CSA is nearest.

Although the use of conceptual spaces is able to produce pattern recognition algorithms much more compressed than neural network based algorithms for the recognition of the same patterns, it has a serious defect: conceptual spaces are 'in the head' in the sense that they ultimately have perceptual space as a 'vehicle'. And, therefore, a finite cognitive agent dependent on a limited amount of resources, resources for which he is in competition with other finite cognitive agents, will have difficulties in exploiting the full potential of conceptual spaces.

However, the following 'symbolic algorithm': (1) list the letters of the alphabet; (2) check whether **A** is an instance of the first letter; if yes (3) stop; if no (4) check whether **A** is an instance of the second letter; ... (n) stop; is certainly shorter (and safer) than the CSA-algorithm mentioned above.

Other advantages of stepping up to the symbolic level are that:

1. language enables many minds to be connected in what we might call a 'world wide web' overcoming in this way the computational limitations of every single mind;
2. language is not 'in the head', in the sense that language allows:
   - **2.1** the storing of portable information in the form of articles, books, inscriptions, etc. information which, among other things, no longer needs to occupy storing space in individuals' minds;
   - **2.2** objectivity in the treatment of information, because in language information is conveyed by assertions for which there exist public criteria of correctness which we all learn when we learn the language;
3. language extends our representational and computational capabilities. To see this consider the natural number $10^{10^{10}}$. There is no chance that we are able to represent within our visual perceptual space such a multiplicity and distinguish it, for example, from a multiplicity of $10^{10^{10}} \pm 7$ elements. And yet, within number theory, not only there are many things we can prove about such multiplicities, but we can also use their cardinal numbers in our ordinary arithmetical computations. These considerations apply even more so to transfinite cardinal numbers such as $\aleph_0, \aleph_1, \ldots$ and their arithmetic.

Many more are the things that could be said in favour of the great importance of language for pattern-recognition. However, those which have already been mentioned in this section are sufficient to show the crucial rôle that the symbolic level has in the cognitive architecture of a finite cognitive agent who is dependent on a limited amount of resources, resources for which he is in competition with other cognitive agents.

But, before ending this section and the paper, we need to spend a few words to justify the cognitive architecture here presented. To this end, let us consider, as we have repeatedly said, that our cognitive agent is finite, dependent on a limited amount of resources, and engaged in a constant struggle for life with nature and other cognitive agents, and that:

> Owing to this struggle for life, any variation, however slight and from whatever cause proceeding, if it be in any degree profitable to an individual of any species, in its infinitely complex relations to other organic beings and to external nature, will tend to the preservation of that individual, and will generally be inherited by its offspring. ([5], Chapter III, p. 40.)

From this we have that, as a consequence of natural selection,[11] our cognitive agent not only develops a hard-wired pattern-recognition machinery in his visual system — edge detectors, luminance detectors, and the like (see on this the quotation from [6] on p. 7 of this article) — but also a multi-level cognitive architecture for the representation and manipulation of information.

At this point it is clear that questions like 'Why does the cognitive architecture have three different levels?', 'How do conceptual spaces come about in the cognitive architecture?', etc. can only be give an 'evolutionary answer', that is, the cognitve architecture we have illustrated above is the consequence of variations which come about in the system of representation and manipulation of information of human beings. These are variations which have been preserved as a consequence of their being greatly profitable for the crucially important pattern-recognition activity of humans.

7 Conclusions

In this paper we intended to give a contribution to the foundations of pattern-recognition theory; and, to do so, we decided to address two central questions: (a) 'What is a pattern?' and (b) 'How do we come to know patterns?'

Dealing with question (a), we produced a definition of mathematical pattern which we then generalized to fields different from mathematics (philosophy of mind and action, physics). But, when it came to answering question (b), we thought of presenting a cognitive architecture for a finite cognitive agent who is dependent on a limited amount of resources. This is a cognitive architecture

[11] 'This preservation of favourable variations and the rejection of injurious variations, I call Natural Selection.' ([5], Chapter IV, p. 51).

which is, in principle, able to cope with some of the basic demands posed by the process of pattern-recognition; and has developed as a consequence of Darwinian natural selection.

References

1. Bombieri, E.: 2013, 'The shifting aspects of truth in mathematics', *Euresis*, **vol. 5**, pp. 249–272.
2. Chaitin, G.: 1975, 'Randomness and Mathematical Proof', *Scientific American*, Vol. CCXXXII, pp. 47-52.
3. Chella, A., M. Frixione, and S. Gaglio. A cognitive architecture for artificial vision. Artif. Intell., 89:73111, 1997.
4. Dales, H.G. & Oliveri, G. (eds.): 1998, *Truth in Mathematics*, Oxford University Press, Oxford.
5. Darwin, C.: 1859, *On the Origin of Species*, edited by M. T. Ghiselin, Dover Publications, 2006, Mineola, New York.
6. Dennett, D.: 1991, 'Real Patterns', *The Journal of Philosophy*, Vol. 88, No. 1, pp. 27-51.
7. Gaglio, S., P. P. Puliafito, M. Paolucci, and P. P. Perotto. 1988. Some problems on uncertain knowledge acquisition for rule based systems. Decis. Support Syst. 4, 3 (September 1988), 307-312. DOI=10.1016/0167-9236(88)90018-8 http://dx.doi.org/10.1016/0167-9236(88)90018-8
8. Gärdenfors, P.: 2004, *Conceptual Spaces: The Geometry of Thought*, MIT Press, Cambridge, Massachusetts.
9. Gärdenfors, P.: 2004. 'Conceptual spaces as a framework for knowledge representation'. *Mind and Matter* 2 (2):9-27.
10. Oliveri, G.: 1997, 'Mathematics. A Science of Patterns?', *Synthese*, **vol. 112**, issue 3, pp. 379–402.
11. Oliveri, G.: 1998, 'True to the Pattern', in: [4], pp. 253–269.
12. Oliveri, G.: 2007, *A Realist Philosophy of Mathematics*, College Publications, London.
13. Oliveri, G.: 2012, 'Object, Structure, and Form', *Logique & Analyse*, **vol. 219**, pp. 401-442.
14. Resnik, M.D.: 1981, 'Mathematics as a Science of Patterns: Ontology and Reference', *Noûs* **XV**, pp. 529-550.

[Resnik, 2001] Resnik, M.D.: 2001, *Mathematics as a Science of Patterns*, Clarendon Press, Oxford.

15. Shapiro, S.: 2000, *Philosophy of Mathematics. Structure and Ontology*, Oxford University Press, Oxford.
16. Tuller, A.: 1967, *A Modern Introduction to Geometries*, D. Van Nostrand Company, Inc., Princeton, New Jersey.
17. Wittgenstein, L.: 1983, *Philosophical Investigations*, Second Edition, transl. by G. E. M. Anscombe, B. Blackwell, Oxford.