A proposal about the meaning of scale, scope and resolution in the context of the interpretation process

GERARDO FEBRES*,**
* Departamento de Procesos y Sistemas, Universidad Simón Bolívar, Venezuela.
** Laboratorio de Evolución, Universidad Simón Bolívar, Venezuela.

When considering perceptions, the observation scale and resolution are closely related properties. There is consensus in considering resolution as the density of elementary pieces of information in a specified information space. Differently, with the concept of scale, several conceptions compete for a consistent meaning. Scale is typically regarded as way to indicate the degree of detail in which an observation is performed. But surprisingly, there is not a unified definition of scale as a description's property. This paper offers a precise definition of scale, and a method to quantify it as a property associated to the interpretation of a description. To complete the parameters needed to describe the perception of a description, the concepts of scope and resolution are also exposed with an exact meaning. A model describing the recursive process of interpretation, based on evolving steps of scale, scope and resolution, is introduced. The model relies on the conception of observation scale and its association to the selection of symbols. Five experiments illustrate the application of these concepts, showing that resolution, scale and scope integrate the set of properties to define any point of view from which an observation is performed and interpreted.

Key Words: interpretation; scale; fundamental scale; scope; resolution; information.

1. INTRODUCTION

The 'anatomy' of a description has been the subject of intense discussion. Three abstract entities have been recognized as essential [1] [2] for the construction of descriptions in any language or communication system: resolution, scale and scope.

Surprisingly, there is not a unified definition of scale of a description. If scale is treated as a quantifiable concept—one that can be managed by the computer—, the available definitions are even fuzzier.

Complex systems offer difficulty to the recognition of all their components. That is precisely a reason to call them complex systems. Depending on the focus of the observer, some parts of the system may be recognized as elements while other parts may be shadowed. Additionally the frontier separating the perceived elements may not be clearly defined due to overlapping or, on the contrary, due to empty meaningless sectors, which do not add information to the perceived system description. Whatever the situation is, the observer decides which ones are the elements forming his interpretation. When the observer associates some sectors of the description with some meaning or logical pattern, he decides which are symbols representing the elements forming his interpretation of the observed description. This suggests that the selection of portions of the description to convert them into symbols, thus building a language, is a crucial step in its proper interpretation. Since the number of symbols and their size are intimately related with the concept of scale, there seems to be a link between the meaning of scale and the set of symbols participating in the interpretation of a descriptive process.
The concept of scale, along with its relationship to emergence and complexity, have been subject of research and discussion. Heylighen [1] presented emergence as a measure of the change of dynamics after a system transition. This measure cannot be directly made over the system itself but over models or observations of it. Bar-Yam [2,3] has associated complexity with information profiles. In this sense, Bar-Yam identifies the relevance of the information that emerges when the system is observed from different detail levels, as the essential cause of complexity. Ryan [4] depicts the relationship of scope, resolution and self-organization, considering emergence as the apparition of novel properties that a system exhibits when it changes from a condition to another. The discussion focuses on scope and resolution, but scale is left as a slave property of resolution.

Prokopenko, Bochetti and Ryan [5] consider scale as parameter defining the emergence phenomena. However in their treatment, scale is almost the same as degree of detail or level of resolution, thus diminishing the degree of independence that scale, as a concept, should have in relation to scope and resolution.

Fernandez, Maldonado and Gershenson [6] indicate that any change of the system's structure is reflected on the quantity of information needed to describe the system before and after the change. The change of the amount of information is a measure of the emergence between any two states. Fernandez et al. [6] show how four numbers initially expressed in a sequence of binary digits, can be presented in a sequence of numbers expressed at different basis. The resulting entropy, computed for each string, clearly suggests there is an important impact of the base of the language used—the number of different symbols—, in the effort the reader must apply to interpret the message.

The treatment of this problem—the observation of a system description—has been typically restricted to the idea of considering the scale, as a representation of the selected way of grouping information elements into groups of regular shapes and equal size; in other words, grouping information elements into spaces topologically equivalent. This vision has proved to be of limited utility since it is a linear simplification of the resolution and therefore does not add freedom to model the consequences of varying parameters within the process of interpreting descriptions. This paper offers a novel conception of scale, establishing clear differences with the concepts of resolution and scope. These concepts are intrinsically involved in the description of systems, and not making the appropriate distinction between them may restrict the possibility of studying systems at several scales. Among the possible scales of looking at a system’s description, there is one we give special attention: The Fundamental Scale. Some formalization of the concept of Fundamental Scale, as the scale at which an observation can be interpreted with minimal entropy, is also an objective of this paper.

2. PROPERTIES OF DESCRIPTIONS: RESOLUTION, SCOPE, SCALE, ALPHABET AND ENCODING

2.1. Resolution

Resolution is a human created artifact to split system descriptions in regular, equally shaped and sized, pieces. It results from the process of discretizing the description of a system. The original description can be discrete or can be an analogous depiction directly taken from
physical reality. Resolution ends up being the number of equally sized pieces in which we divide the original description, and thus it refers to the size of the smallest piece of information of a description, or equivalently, to the density of information elementary pieces. It is commonly specified as the number of smallest information pieces that fit into each dimension of the description. Considering all dimensions conforming the description, resolution can be regarded as the total number of elementary information pieces included in the whole description. In this case we use the letter $R$ to refer to it. When resolution is specified as the density of information contained in a physical dimension, we use the number information pieces $r_j$ that fit into the dimension considered, thus $r_j = R_j/\text{Dim}_j$, where $\text{Dim}_j$ refers to the absolute size of the physical dimension used. As an example we can consider a 16" x 9" computer screen with 1920 pixels in the horizontal longitudinal dimension and 1080 pixels in the vertical longitudinal dimension. The resolution $R$ is regarded as 1920 x 1080 [pixels x pixels] and $r$ would be 120 x 120 [pixels/in • pixels/in]. If the description refers to a 60-seconds long sequence of 36000 very short sounds, then $R = 36000$ [sounds] and $r = 600$ [sounds/sec].

Resolution, as a concept, losses meaning when the mesh of information elements is not regular —an information structure formed by a set of symbols with a diversity of sizes—. Such a situation can hardly be described using the resolution as a characteristic parameter because the density of the resolution would not be a constant.

2.2. Scale

All of us have an intuitive notion of scale. Commonly, the term scale is associated with the distance from which the system is observed. Thus, the term ‘scale’ is typically used to mean that the system is being interpreted from a closer point of view —higher scale with finer detail—, or from a farther point of view —lower scale with less detail. This would work fine if scale were exclusively defined by the distance between the observer and the object. This not always the case. In fact, nature is not built with equally shaped and equally sized objects, but the dominance of the use of artificial devices to observe nature, has led use the word ‘level’ as a synonym of scale.

The frequent interchange of the terms ‘detail degree’, ‘level’, ‘observation distance’, and ‘scale’, promotes the confusion between the concepts of scale and resolution. I think this confusion of terms has delayed a clear conception of what scale actually is, and as a result, we are still lacking of a unified definition of scale of a description that can work in an information space with non-regular shaped symbols.

During the last decade, several studies reflected the relevance the concept of scale in our interpretation of descriptions. In 2004 Bar-Yam [2,3] presents complexity as a property intimately related to the scale. His treatment of scale as a variable capable of varying continuously rather than discretely, leads him to present scale profiles, where entropy decreases monotonically as the scale shows less detail. Even though these profiles are a good representation of the envelope of the complexity, they would only be constructible by representing scale as the result of the increase, or degradation, of the resolution with which the system is observed. Thus, these profiles illustrate the relationship between complexity and resolution, while the actual scale is left out of analysis.
Piasecki and Plastino [7] showed entropy as a function of scale length—the size of the group of pixels used to represent each object component—in a pattern of regularly distributed greyscale pixels. Their graphs show local minimal values of entropy when the scale length is a multiple of the characteristic size of the pattern, measured in pixels. Again, Piasecki and Plastino [7] actually work with resolution changes, but their experiment shows how, when the size of the objects being described fits with an integer number of pixels, the entropy radically drops. It seems that the observation scale is not exclusively dominated by the system properties. Even more, the scale is a product of our ‘interpretation processes’. When we consider a description, our brain probably scans several interpretations of the observed description. At each interpretation, we combine raw information by joining adjacent information elements and forming with them hypothetical larger symbols. Simultaneously we look for patterns which we can associate with previous experiences and learned notions, or even with our personal conception of beauty, thus giving certain meaning to a message that was initially abstract. Therefore, the scale is a property of the way the observer looks at the system’s description. Once the observed system’s conception is organized in our brain, a clear account of the symbols resulting from our interpretation, along with their frequency of appearance and their relative position, constitute our model of the system.

This argument lets me introduce statements about scale that does not contradict our previous intuitive notions: The Scale of a system, as it is observed, is directly related to the set of symbols used to create the system’s model. The scale can be numerically represented by the number of different symbols used in each interpretation, thus the numerical value of the scale works as the base of the language we use for our interpretation. The numerical system we use, for example, is regarded as a base=10 system because it consists of 10 different digits. If we use only two symbols, then we are reading the system at base=2; the system has not to be binary, but we are interpreting it with a binary language.

When symbols fit into a regular lattice of pixels, the number of pixels forming each one of these regular symbols, specify the shape and the size of them. If this were the case, saying the system exhibits rectangles formed by \( n \times m \) pixels, could be appropriate to specify how we are looking at the system’s description, in short, could be appropriate to specify our scale of observation. But if we, on the contrary, are looking at the countries of a map, there is no constant arrays of pixels that can be assigned to indicate that we are looking at countries. The symbols here should be the countries shown in the map, disregarding any number of pixels contained in any country; the base of the scale would be the number of different countries seen in the map. If we see the same map at the scale of continents, the symbols become the continents, and the scale’s base would be the number of them.

Finally, it should be emphasized the quantitative notion of scale, or the base of the scale, which value is identical to the symbolic diversity and therefore the designation of \( D \) is interchangeably used to refer to both, diversity and scale’s base, or simply scale.

2.3. Scope

The scope refers to the total number of information units contained in the description. When the description is done over arrays on elements regularly distributed, the scope equals de
product of the resolution of each dimension of the description. Up to this point, scope seems to be a redundant concept with resolution. However, when the information is contained in a non-regular sized mesh, the resolution, as parameter, losses its meaning and utility; nevertheless, we could still use scope to characterize the description, just by counting the number of information elements, or symbols, contained.

Once the size and shape of the symbols have been established by the selected scale, the entire description conveys an amount of information determined by the total number of symbols, repeated or not, included in the description. In this sense, the scope equals the length of the description measured as number of symbols and thus, length and scope are both represented by the letter $l$.

2.4. Symbols, alphabet and encoding

From our point of view, alphabet is only indirectly connected to a description. An exact definition of our conception of alphabet is justifiable to properly depict some notions about scale. The alphabet is the set of elementary symbols used to create the complex symbols that may be part of our interpretation of a message. In other words, a symbol may be formed by several alphabetic symbols—usually regarded as characters—. When reading an English text in a computer screen, the symbols of the description could be the words of the text, while the alphabet should be considered as the set of letters used to form larger symbols or words. While these larger symbols are useful due to their capacity to hold meaning or semantic content, the symbols forming the alphabet are only tools with no meaning by themselves. Even though alphabet symbols can be represented as the conjunction of even smaller symbols, at certain point of our interpretation, we are not interested in the sub-symbols of the alphabet symbols, thus considering them as elementary, at least at this observation scale.

The symbols of an alphabet are useful to create larger, and probably meaningful, symbols. We refer to the rules which allow the formation of alphabet symbols as the ‘encoding’. With this schema an English text can be regarded as a description formed by words, which in turn are formed by letters. These letters are elements of the set considered as the English alphabet. Now, each one of these letters may be written by a single drawing, as with a pencil and paper, or by another ‘letter forming system’ as a group of pixels with colors resembling the alphabetic symbol, or the set of dashes and dots if the Morse is the encoding system used.

This reasoning leads to treat alphabet and encoding as artifacts which enable us to construct an observation scale, or simply an observation. But the alphabet and the encoding are not directly linked to the description, though they are essential to any written communication system. Whenever the alphabet is confused with the set of symbols used in a description, the freedom to form independent interpretations of the message, disappears.

3. NOTIONS ABOUT ‘SCALE’.

3.1. An intuitive notion of scale

The scale of a description has been commonly considered as groups of the most elemental information component used within the description. If, for example, the description consists of an image projected on a computer screen, pixels would be the most elemental component
because each pixel shows the same color and therefore it does not make sense to divide it into smaller pieces. Now assume the screen is showing a bunch of letters, one after another forming an array of characters that fills up the screen. If we observe the letters, the scale would be represented by groups of adjacent pixels, each group forming images of larger size than a single pixel, and containing a graphical description of a letter. If we decide that we are observing characters instead of pixels—which we have never stopped seeing—then we can say that our observation is at the Character Scale. It is also possible to think of the same computer screen description, being represented by the binary codes of each characters instead of the characters themselves. In fact, any text being represented in a digital computer, as we know them today, is only the reflex of some binary code physically stored as the orientation of micro-magnets or the shine grade of micro-mirrors contained in a magnetic or optical disk. If we could magnify the disk where this message is stored, we would see the same message at the scale of bits.

3.2. The alphabet as a coding tool

People speak, read and write natural languages using words as symbols to pack any meaning they want to incorporate to a more complex entity: the idea. When communication occurs by means of patterns of sonic, graphic, chemical signals or any physical media, the meaningful ideas or instinctive triggers are associated with these patterns. No transcription is needed because each receiver already has an interpretation and meaning about the signal pattern received. It is worthwhile to mention that the absence of meaning—any meaning including abstract meaning—implies considering the involved set of symbols as noise. The written version of a message is just the coded version of the idea the sender intends to communicate. Here, I consider an alphabet as the tool that serves the elementary information pieces to allow for the coding of a message and to ‘write’ it onto a surface. Thus, for western natural languages, for example, the alphabet is the set of letters we use to form words and to embed them in a physical media. For western music, the alphabet is the set of figures used to represent a sound with a specified pitch with the position of the figure over the pentagram. The figure’s shape itself indicates the lasting of the sound. Even though contrasting the conventional conception of alphabet, for Asian languages, the alphabet is formed by symbols, as each one of those symbols represents a sound with meaning in its corresponding language.

This is important because we can analyze information by handling some sort of written version of any message. In order to operate and compute properties of a message or description expressed in a specific language, we do not use the original version of the message. Instead, we rely on a written version that allow us to split and group different parts of the message, and look for logical patterns which may build up or represent information.

We do not know how to directly deal with the perceptions of our senses. The writing systems, then, are crucial for our method because we are actually evaluating the properties of communication systems regarding their representations by means of a writing system. To illustrate this point, the Table 1 shows a non-exhaustive list of several languages with some associated writing systems and their alphabets. The proposal is then, to recognize those sequences of characters which can be used as information elementary units to form a message transcription.
Table 1: Examples of Languages of different nature and their corresponding writing and encoding systems.

| Nature/Dimensions | Language Group | Examples | Alphabet components | Encodings |
|-------------------|----------------|----------|---------------------|-----------|
| Natural Languages 1 Dimension | Alphabetic | English | letters | Conventional writing |
| | | Spanish | | Binary |
| | | Russian | | Braille |
| | | French | | Morse Code |
| | | Arabic | | ASCII code |
| | Syllabic | Chinese | | |
| | | Korean | | |
| | | Japanese | | |
| Musical | Western music | Chromatic scale | | Pentagram Notation |
| 1 Dimension | Indu-raga music | Major scale, Minor scales | | |
| | Chinese music | Pentatonic scale | | |
| | African Traditional Music | Other Scales | | |
| | Alarm and Warning signals | Sound Volume | | |
| | | Alteration signs | | |
| | | Rhythms | | |
| | | Tempos | | |
| | | Stridency of harmonies | | |
| | | Ascending or descending tonal patterns | | |
| Graphic 2 Dims. | 2D graphical expressions | Color spectrum | | Pixel Size |
| | | Brush Strokes | | Pixel Color |
| | Charts & Diagrams | Geometric Shapes | | |
| Mathematics 1+ Dimensions | Numerical system, 1 dimension | Numbers | Digits and decimal point | 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, ‘.’ |
| | Equations, 1+ dimensions | Math Expressions | Mathematical operators | Conventional Math |
| | | | Char strings declared as Symbols | ASCII code |
| Biology | Genomes | ADN | Adenine, Thymine, Cytosine, Guanine | A, T, C, G |
| | | RNA | Adenine, Uracil, Cytosine, Guanine | A, U, C, G |

3.3. The scale is a choice.

The observation scale does not have to be made at a regular space. When nature decides how to group its elements to constitute organisms and societies, it rarely pays attention to the regularity of the spaces that serve as frames for those compound entities. Of course, there are cases that can be seen as exceptions to this statement. The hexagonal pattern in a panel of bees or the fractal describing the appearance of a fern, could be considered regular spaces, but in a description combining panels of bees and ferns, the resulting space would not be regular any more. Thus, when observing the elementary information pieces of a description, for
example letters in an English-written text, the question is how to group those letters, in order to get the best possible understanding of the text; the answer would be: words.

In the communication process there is always present pressure to make the language used more effective. As Zipf [8] stated in his “Principle of the Least Effort”, we tend to reserve shorter words to apply them for the more frequent ideas we need to express. Starting from Kolmogorov’s “Shortest Description Length”, in 1978 Rissanen [9] introduced the Minimal Description Length Principle (MDL. He used the name “Shortest Data Description”) to build minimal description length models of an observed sequence of characters. The MDL has been widely studied afterwards (Hansen and Yu [10]). Many, if not most, studies are centered in the problem of the compressibility of strings, approaching the MDL problem as the recognition of probability distribution model which best resemble real probability distributions of symbols within certain description.

The typical approach consider the binary alphabet as the base for all descriptions. Thus considering the binary alphabet as the base for the descriptions whose length is measured and is eventually compressed. For the computer science field this is justifiable because in digital computers all files, strings, models, etc., are at some point, stored and handled through binary codes. But we are concerned with the language itself, and therefore this proposal deals with the search of the alphabet in which the description was naturally written. In doing so we are highlighting the fact that any language has evolved around its adaptation in order to produce short descriptions.

A hypothesis is then, that for every language, there is a set of symbols which reduce at a minimum the lengths of the descriptions built with it, thus minimizing their entropy based on the frequency of the symbols. Along with their relative frequencies, this set of symbols constitute a good model of the language, and represents what the language itself has chosen as its underlying structure; it is an appropriated scale to use when interpreting a description. Since this is the only one scale (set of symbols) that minimizes the entropy of a description (expressed with a language), the distinctive word ‘fundamental’ is added. Some examples reinforce this explanation.

Example 1 (E1): Assume there is a character sequence that we consider a description: “a ab abc abcd abcde abcdef abedefg”. Using $W$ to represent a random variable that can take a specific symbol as value, and focusing on single characters as the symbols of this description, leads to the following probabilities based on their frequencies. We refer to this as the interpretation (E1.I) of Example 1 (E1.I): $P(W = 'Ø') = \frac{7}{35}$, $P(W = 'a') = \frac{7}{35}$, $P(W = 'b') = \frac{6}{35}$, $P(W = 'c') = \frac{5}{35}$, $P(W = 'd') = \frac{4}{35}$, $P(W = 'e') = \frac{3}{35}$, $P(W = 'f') = \frac{2}{35}$, $P(W = 'g') = \frac{1}{35}$. The scale base is $D = 8$ and the computed entropy is $h = 0.937$. Each ‘space’ in the message have been represented as ‘Ø’ in the probability expressions.

Example 2 (E2): A second interpretation of the same description assumes the message was expressed with a natural language, and therefore regards as symbols those character strings which look as words, those are the symbol strings beginning or ending with a ‘space’ (‘Ø’). Under this conditions the symbols of the description would lead to the following probabilities, computed with Eq. (3): $P(W = ' Øa') = \frac{1+2}{35}$, $P(W = ' Øab') = \frac{1+3}{35}$, $P(W = ' Øabc') = \frac{1+4}{35}$, $P(W = ' Øabcdefabcdefg') = \frac{1+5}{35}$. The scale base is $D = 8$ and the computed entropy is $h = 0.937$. Each ‘space’ in the message have been represented as ‘Ø’ in the probability expressions.
Example 3 (E3): Another interpretation looks for sequences of characters to form symbols that lead to a reduction of computed entropy. Under this conditions the symbols of the description lead to the following probabilities based on their frequencies and sizes: 

\[ P(W = 'a') = \frac{2+2}{35}, \quad P(W = 'b') = \frac{1+1}{35}, \quad P(W = 'c') = \frac{5+4}{35}, \quad P(W = 'd') = \frac{1+1}{35}, \quad P(W = 'e') = \frac{2+2}{35}, \quad P(W = 'f') = \frac{1+1}{35} \]

The scale base is \( D = 7 \) and the computed entropy is \( h = 0.895 \).

The notion of the minimal entropy, being independent of the alphabet used, supports the validity of the concept of Fundamental Scale as a property intimately related to the structure of the language used, despite of alphabet which serves as a mere instrument to build the written version of the message.

3.4. The Fundamental Scale

When the communication system works on unknown rules, it is not possible to decide a priori the scale to interpret the description. That is the case of the texts of any file containing recorded music. There are no words in the sense we are used to, and the characters we see do not indicate any meaning for us. Thus, we cannot even be sure about the meaning of the space character “ “. In natural languages, a space is used as delimiter for words, but in music a space does not mean a silence. Fortunately, even having no idea about the ‘grammar’ of a communication system, we still can rely on the Minimal Description Length Principle (MDL) to reveal the symbols with which that communication system is built.

Therefore, whenever a description is expressed in a language with unknown grammar, the MDL principle can be applied by means of the Fundamental Scale Algorithm [11], and obtain measurements of the quantity of information broadcast with the description.

4. THE INTERPRETATION PROCESS

All along our treatment, the idea about scale as being a consequence of the observer’s choice of the symbols have been emphasized. Generally people speak, read and write natural languages using words as symbols to pack any meaning they want to incorporate to a more complex structure: the idea.

The question is: Once we have the minimal entropy set of symbols for a message, that is the fundamental scale, what could happen if we express the message with a different encoding? Specifically, an encoding of lower base as compared to the original one.

The encoding system is just a convenient way of assembling the symbols to represent a description. The encoding system is usually linked to a set of elementary symbols referred to as alphabet. Intuitively the encoding should not change the Fundamental Scale of a description, nor should it change its quantity of information or entropy. The treatment of this conjecture is as follows:
Consider a description $\mathcal{M}$ as the sequence of symbols $A_i$, all of them contained in the alphabet $\mathcal{A} = \{A_1, A_2, A_3, \ldots, A_n\}$. Thus, $\mathcal{M}$ can be seen as a specific sequence of the letters (or characters) of alphabet $\mathcal{A}$. A descriptor of $\mathcal{M}$ is the probability distribution $P(A_i)$ where $A$ is the random variable associated with the frequency with which each letter of alphabet $\mathcal{A}$ is encountered within message $\mathcal{M}$. Then, using $f_{A_i}$ to represent the number of times the letter $A_i$ appears within the message, we can model message $\mathcal{M}$ writing:

$$\mathcal{M}: P(A) = \{P(A = A_1), P(A = A_2), \ldots, P(A = A_i), \ldots, P(A = A_n)\} = \{f_{A_1}, f_{A_2}, \ldots, f_{A_n}\}. \quad (1)$$

Then, the entropy $h_{\mathcal{A}n}$ associated to the message $\mathcal{M}$ when observed as the $n$ different letters of alphabet $\mathcal{A}$, equals the entropy of distribution $P(A)$,

$$h_{\mathcal{A}n} = - \sum_{i=1}^{n} f_{A_i} \log_2 f_{A_i}. \quad (2)$$

The figure computed in Expression (2) holds for the entropy obtained when reading the message one letter at the time; each letter would be considered as a symbol and no symbol would be allowed to be formed by more than a single letter. This is what I call ‘reading at the scale of letters’. Intuitively we know this would be a very ineffective way of reading. Fortunately, there are better ways of reading message $\mathcal{M}$ by grouping neighbor characters $A_i$ to form $D$ different multi-character symbols $Y_j$. This approach may be effective in the sense of reducing the associated entropy. Even more, assume the symbols $Y_j$ are such that the resulting entropy is minimal, we can write:

$$h^*_{\mathcal{A}D} = - \sum_{j=1}^{D} P(Y_j) \log_2 P(Y_j), \quad (3)$$

where $h^*_{\mathcal{A}D}$ represents the minimum possible entropy associated to the process of interpreting the message encoded by means of alphabet $\mathcal{A}$. In order to comply with the neighborhood condition at the time of forming words $Y_j$,

$$Y_j = A_i A_{i+1} A_{i+2} \ldots A_{i+S-1}, \quad (4)$$

where $S$ is the length of symbol $Y_j$. We claim this is a better way of reading $\mathcal{M}$. Therefore, we replace its former model with:

$$\mathcal{M}: P(Y) = \{P(Y = Y_1), P(Y = Y_2), \ldots, P(Y = Y_j), \ldots, P(Y = Y_D)\}. \quad (5)$$

Notice that now symbols $Y_j$ are of different sizes. Therefore the probability $P()$ of encountering a symbol $Y_j$ within the message, should account for the frequency of appearances $f_{Y_j}$, the size of the symbol $S_{Y_j}$ and the size $L_{\mathcal{A}}$ in characters of the message. Thus, this probability can be expressed in terms of the frequency $f_{Y_j}$ and the number of times the symbol of size $S_{Y_j}$ fits into a $L_{\mathcal{A}}$ characters long message, as

$$P(Y_j) = \frac{f_{Y_j} \cdot S_{Y_j}}{L_{\mathcal{A}}}. \quad (6)$$

Equation (3) for minimal entropy can now be rewritten in terms of the symbols of language $Y$ built over the characters of alphabet $\mathcal{A}$. We obtain:
Expression (7) can be seen as the contributions of each symbol to the message's total entropy. Centering our attention on some consecutive symbols, we can study the effects of changing the set of symbols selected. From any arbitrary set of symbols forming the message, three cases are recognizable as possible ways to modify the symbol selection with reference to previous symbol set:

**Case 1. Symbol Splitting:** Splits symbol \( Y_j \) into two smaller symbols. In this case the object symbol \( Y_j \) is replaced by new symbols \( Y'_j \) and \( Y'_{j+1} \). The sub-indexes of all successive symbols will increase by one. The new symbolic diversity also increases from \( D \) to \( D + 2 \). One could think that symbol \( Y_j \) as the only instance of that symbol in the entire text. If that were the case, by replacing it with \( Y'_j \) and \( Y'_{j+1} \), would generate only one additional symbol. However, this seems very unlikely because, in order to be part of a minimal entropy set of symbols, \( Y_j \) should appear more than once in the message. Another possibility is that new symbols \( Y'_j \) and \( Y'_{j+1} \) already exist within the message. This is also unlikely because the minimal entropy criterion has already decided to represent the sequence \( Y'_j Y'_{j+1} \) with the 'compacted' form \( Y_j \). Concluding, for Case 1 we can use \( \mathcal{V} = D + 2 \) as the symbolic diversity for this message interpretation. An example can illustrate this situation.

Example 4 (E4): assume \( W \) is a minimal entropy symbol found within the message of Example (E3). Thus, \( W \) is an element of the Fundamental Scale set, specifically, take \( W_3 = \{ \emptyset abc \} \) from the interpretation of Example 3. Then, symbol \( W_3 \), could be split into two symbols; let’s call them \( W_{31} \) and \( W_{32} \). Making \( W_{31} = \{ \emptyset ab \} \) and \( W_{32} = \{ c \} \) would be a valid replacement for \( W_3 \) since they add up the original \( W_3 \) with no character superposition. In this case the emergence of symbols \( W_{31} = \{ \emptyset ab \} \) and \( W_{32} = \{ c \} \), which did not appear in the previous interpretation, would increase the symbolic diversity from \( D \) to \( V = D + 2 \), and the resulting entropy would raise up in comparison to the formerly computed value.

**Case 2. Symbols Boundary Drifting:** The boundary between two adjacent symbols \( Y_j \) and \( Y_{j+1} \) moves making one symbol larger and the other smaller and resulting in new symbols \( Y'_j \) and \( Y'_{j+1} \). The sub-indexes of symbols does not shift, but the new symbolic diversity increases from \( D \) to \( D + 2 \) due to the birth of new symbols \( Y'_j \) and \( Y'_{j+1} \) which did not exist before. Again, as reasoned for Case 1, if symbols \( Y'_j \) and \( Y'_{j+1} \) had existed before, they would likely appeared as minimal entropy symbols \( Y_j \) and \( Y_{j+1} \) prior to the boundary drifting. Thus, for Case 2 we can use \( V = D + 2 \) as the symbolic diversity for this message interpretation.

Example 5 (E5): assume \( W \) is a minimal entropy symbol found within the message of Example (E3). Thus, \( W \) is an element of the Fundamental Scale set, specifically, take \( W_4 = \{ \emptyset abc \} \) and \( W_4 = \{ d \} \) from the interpretation of Example (E3). Then we could see the sequence \( W_3 W_4 \) (\( \emptyset abcd \)) as the sequence of two new symbols \( W_{31} = \{ \emptyset ab \} \) and \( W_{41} = \{ cd \} \), leaving the original message intact. The emergence of symbols \( W_{31} = \{ \emptyset ab \} \) and \( W_{41} = \{ cd \} \) may compensate for frequency reduction of symbols \( W_3 = \{ \emptyset abc \} \) and \( W_4 = \{ d \} \), which may extinct after the
replacement of these instances. In any case, after each symbol boundary drifting step, the resulting symbol diversity will remain the same or increase at most from \( D \) to \( V = D + 2 \).

**Case 3. Symbol Joining:** Two adjacent symbols \( Y_j \) and \( Y_{j+1} \) are replaced by one symbol larger symbol \( Y'_j \). Depending on the number of instances of symbols \( Y_j \) and \( Y_{j+1} \), the symbolic diversity may reduce zero to two symbols. The new diversity will be in the interval \( V = [D - 2, D] \).

There are many other ways of modifying the selection of symbols. But all of them can be understood as the result of combining cases 1, 2 and 3.

**The interpretation process:** Based on this schema, the interpretation process can be understood as a recursive process of symbol selection. Starting from an arbitrary symbol set to integrate the message, the process of interpretation steps over the splitting, joining and the boundary drifting of symbols, pursuing a probably unconscious reduction of the entropy associated to the selected symbols. This process continues until our interpretation evolves and we are able assign an intelligible meaning to the whole message, or until we are satisfied with the sensations triggered by the message, or until we reject simply paying more attention to it. When there are obvious meaningful patterns, our interpretation quickly fixes those patterns and submits to variations the selection of other symbols to ‘explain’ the remaining sub-spaces within the description’s scope. This is the case of written or spoken words in natural languages. When we can communicate with a certain natural language, we rapidly recognize the symbols of a description, thus the interpretation process is not as complex as it may appear; as we are knowledgeable of the language, we can even detect—and correct or compensate for—deviations that may exist with respect to the language accepted forms. We can imagine this acceleration of the interpretation process, not only for natural languages, but also for other radically different and means of communication. Take, for example the case in which we see a pattern of bands of seven different colors: red, orange, yellow, green, blue and violet. In that order. For most of us that pattern has an already assigned semantic meaning; in spite of its apparent arbitrary and abstract condition, whether we read it or see it, this pattern refers to a rainbow. If the pattern pf colors is shaped as a bow, the association to a rainbow is intensified. Sound signals may also form patterns of symbols we can interpret according to a very similar process, just working with sonic symbols instead of colors or graphical patterns.

### 4.1. The relationship among alphabet, scale, interpretation and entropy

The alphabet is a difficult to describe artifact. It is a set of elementary (non-divisible) symbols used to encode larger symbols. We say the symbols of an alphabet are non-divisible meaning the encodings based on an alphabet do not need to represent any their elementary symbol in terms of others, presumably smaller, symbols. However, we know that any letter in an alphabet can be represented by a decimal or hexadecimal number, just as those numbers can be expressed in binary basis. All those representations require several symbols contained in another alphabet, which again, are considered elementary symbols with respect to the alphabet they belong to. Thus the alphabet is a set of symbols we convene not to split into smaller symbols, while we are encoding within the rules of that alphabet.
We want to explore the convenience of changing the alphabet used to encode the message. Then I propose looking at the change entropy while the alphabet used for the encoding is replace from $\mathcal{A}$ to $\mathcal{B}$. In Equation (5) showed the description of message $\mathcal{M}$ encoded with alphabet $\mathcal{A}$ and somehow interpreted to result in the minimal entropy indicated in Equation (7). If we now encode message $\mathcal{M}$ using an alphabet $\mathcal{B}$ of $m$ symbols forming $E$ different multi-character symbols $B_k$, message $\mathcal{M}$ could be described as:

$$\mathcal{M}: P(B) = \{P(B = B_1), P(B = B_2), ..., P(B = B_k), ..., P(B = B_E)\}.$$  (8)

Expressing the probabilities $P(B = B_k)$ in terms of the frequencies $f_{B_k}$ and the number of times the symbol of size $S_{B_k}$ fits into $L_B$, we write:

$$h_{BE} = - \sum_{k=1}^{E} f_{B_k} \cdot S_{B_k} \log_{E} \frac{f_{B_k} \cdot S_{B_k}}{L_B}.$$  (9)

Notice that in Equation (9) no asterisk is not shown on the entropy token $h_{BE}$, indicating the possibility of entropy of message $\mathcal{M}$ being not minimal if interpreted with symbols formed with alphabet $\mathcal{B}$.

As mentioned above, the base of alphabet $\mathcal{A}$ is $n$ and the base of alphabet $\mathcal{B}$ is $m$. Thus we can approximate the length of symbols $S_{B_k}$ encoded in terms of alphabet $\mathcal{B}$, in terms of the length of the same symbol expressed with alphabet $\mathcal{A}$. Similarly, the length of the total description $L_B$ in terms of $L_A$ as follows:

$$S_{B_k} \approx \left(\frac{S_{Y_j}}{L_A}\right)^c, \quad L_B \approx \left(\frac{L_A}{m}\right)^c$$  (10a)(10b)

where $\frac{n}{m}$, is the ratio of the alphabets’ bases.

We have a way to study the effects of changing the alphabet used to encode the message, at least around the minimal entropy symbol set for alphabet $\mathcal{A}$. We now proceed to evaluate the change of the entropy when the alphabet $\mathcal{A}$ is replaced by $\mathcal{B}$. As for the interpretation process, here we see several cases:

**Case 4. Change of alphabet. Invariant symbol selection:** Let’s consider the contribution of symbol $Y_j$ to the total entropy, as indicated by Equation (8). We refer to it as $h_{AV_j}$. Keep in mind $h_{AV_j}$ is not an entropy; it is just symbol’s $Y_j$ contribution to it when encoded using alphabet $\mathcal{A}$.

$$h_{AV_j} = - \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \log_{D} \frac{f_{Y_j} \cdot S_{Y_j}}{L_A}.$$  (11)

Now, representing symbol $Y_j$ using alphabet $\mathcal{B}$, we can write

$$h_{BV_j} = - \frac{f_{B_j} \cdot S_{B_j}}{L_B} \log_{D} \frac{f_{B_j} \cdot S_{B_j}}{L_B},$$  (12)
where \( B_j \) refers to the same symbol \( Y_j \), but this time encoded with alphabet \( B \), while \( S_{Y_j} \) and \( S_{B_j} \) are the corresponding lengths of symbol \( Y_j \) when encoded with alphabets \( A \) and \( B \). Under these circumstances, clearly \( S_{B_j}/L_B \) equals \( S_{Y_j}/L_A \) and \( f_{B_j} \) equals \( f_{Y_j} \) as well. Thus, a change of the encoding’s alphabet applied to a fixed interpretation, does not change the entropy contributions of the selected symbols, nor does it change the total entropy associated with the interpretation.

**Case 5. Change of alphabet. Symbol boundary drifting:** Consider the contribution \( h_{AY_{Y+1}} \) to the total entropy of two consecutive symbols \( Y_j \) and \( Y_{j+1} \) of the minimal entropy interpretation of message \( M \) encoded with alphabet \( A \); it can be estimated by as:

\[
h_{AY_{Yj+1}} = - \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \log_D \left( \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \right) - \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \log_D \left( \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \right).
\]

(13)

Now, changing the encoding to alphabet \( B \) and focusing on the segment of the message which corresponds to symbols \( Y_j \) and \( Y_{j+1} \), thus time represented by symbols \( B_k \) and \( B_{k+1} \), the entropy contribution \( h_{BB_kB_{k+1}} \) is estimated as:

\[
h_{BB_kB_{k+1}} = - \frac{f_{B_k} \cdot S_{B_k}}{L_B} \log_B \left( \frac{f_{B_k} \cdot S_{B_k}}{L_B} \right) - \frac{f_{B_{k+1}} \cdot S_{B_{k+1}}}{L_B} \log_B \left( \frac{f_{B_{k+1}} \cdot S_{B_{k+1}}}{L_B} \right).
\]

(14)

Notice the base of the algorithms is now \( V \), which considers the possibility of the symbol diversity increasing from \( D \) to \( V = D \cdot V = D + 1 \) or \( V = D + 2 \), depending on the existence of instances of the new symbols \( B_k \) and \( B_{k+1} \) coincident with symbols \( Y \) in the former interpretation, and the number of remaining instances of the replaced symbols \( Y_j \) and \( Y_{j+1} \); in any case, the new symbol diversity \( V \) does not depend on the alphabets’ base relationship \( c \). The variation of these entropy contributions is then \( \Delta h_{AB} = h_{AY_{Yj+1}} - h_{BB_kB_{k+1}} \). Expanding it we obtain:

\[
\Delta h_{AB} = - \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \log_D \left( \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \right) - \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \log_D \left( \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \right)
\]

(15)

\[
+ \frac{f_{B_k} \cdot S_{B_k}}{L_B} \log_B \left( \frac{f_{B_k} \cdot S_{B_k}}{L_B} \right) + \frac{f_{B_{k+1}} \cdot S_{B_{k+1}}}{L_B} \log_B \left( \frac{f_{B_{k+1}} \cdot S_{B_{k+1}}}{L_B} \right).
\]

(16)

Whichever the number of characters the boundary between \( B_k \) and \( B_{k+1} \) moves, the following relation must hold:

\[
S_{B_k} + S_{B_{k+1}} = S_{Y_j}^c + S_{Y_{j+1}}^c \Rightarrow S_{B_{k+1}} = S_{Y_j}^c + S_{Y_{j+1}}^c - S_{B_k}.
\]

(16)

Applying Equations (10b) and (16), Equation (15) can be rewritten as:

\[
\Delta h_{AB} = - \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \log_D \left( \frac{f_{Y_j} \cdot S_{Y_j}}{L_A} \right) - \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \log_D \left( \frac{f_{Y_{j+1}} \cdot S_{Y_{j+1}}}{L_A} \right)
\]

(17)

\[
+ \frac{f_{B_k} \cdot S_{B_k}}{L_B} \log_B \left( \frac{f_{B_k} \cdot S_{B_k}}{L_B} \right).
\]
Taking the derivative of the entropy contribution \( \Delta h_{AB} \) in expression (17) with respect to \( c \) we obtain the effect of changing the alphabet basis with symbol boundary drifting.

\[
\frac{\partial \Delta h_{AB}}{\partial c} = -c \cdot \frac{f_{Bk} \cdot S_{Bk}}{L_a c} \log_{L_a c} f_{Bk} \cdot S_{Bk} + \frac{L_a c}{f_{Bk} \cdot S_{Bk}} \cdot \frac{1}{\ln V} \cdot f_{Bk} \cdot S_{Bk} \]

\[
+ \frac{f_{Bk+1} \left( \log f_{Bk+1} \right)}{L_a c} \cdot f_{Bk+1} \cdot \frac{S_{Bk+1} - S_Bk}{S_{Bk+1} - S_Bk} \]

\[
+ \frac{S_{Bk+1}}{L_a c} \cdot \frac{1}{\ln V} \cdot f_{Bk+1} \cdot \frac{S_{Bk+1} - S_Bk}{S_{Bk+1} - S_Bk} \]

rearranging terms,

\[
\frac{\partial \Delta h_{AB}}{\partial c} = -c \cdot \frac{f_{Bk} \cdot S_{Bk}}{L_a c+1} \log_{L_a c+1} f_{Bk} \cdot S_{Bk} + \frac{L_a c}{f_{Bk} \cdot S_{Bk}} \cdot \frac{1}{\ln V} \cdot f_{Bk} \cdot S_{Bk} \]

\[
+ \frac{f_{Bk+1} \left( \log f_{Bk+1} \right)}{L_a c} \cdot f_{Bk+1} \cdot \frac{S_{Bk+1} - S_Bk}{S_{Bk+1} - S_Bk} \]

\[
+ \frac{S_{Bk+1}}{L_a c} \cdot \frac{1}{\ln V} \cdot f_{Bk+1} \cdot \frac{S_{Bk+1} - S_Bk}{S_{Bk+1} - S_Bk} \]

By inspecting (18b), those positive and negative terms can be recognized. Therefore we can write a condition for \( \frac{\partial \Delta h_{AB}}{\partial c} > 0 \). That is

\[
\frac{\partial \Delta h_{AB}}{\partial c} > 0 \text{ if and only if } G + H < P + Q + R \text{ , } \]

where

\[
G = c \cdot \frac{L_a c \cdot f_{Bk+1} \left( S_{Bk+1} c - S_{Bk} \right)}{L_a c+1} \log_{L_a c+1} f_{Bk+1} \cdot S_{Bk+1} \cdot S_{Bk} \]

\[
H = -c \cdot \frac{L_a c \cdot f_{Bk+1} \left( S_{Bk+1} c - S_{Bk} \right)}{L_a c+1} \frac{1}{\ln V} \]
Evaluating Expression (19) by parts we can conclude that for non-small values of $L_d$ and $c > 1$, $G$ and $H$ grow faster than $Q$ and $R$ respectively. $P$ is always positive. Hence, any drifting of the boundary between two minimal entropy symbols will produce an increase of the resulting interpretation entropy which cannot be compensated by encoding the message with a lower-based alphabet. Therefore, condition (19) shows that deviating our attention from a minimal entropy symbol set, increases the symbolic entropy. Additionally, changing the encoding of message $M$ from an alphabet $A$ of base $n$, producing $D$ different symbols of minimal entropy, to an alphabet $B$ based on $m$ characters producing $E$ different symbols, will have no effect on the resulting symbolic entropy. On the other hand, any deviation from the minimal entropy symbol set, brings an increase of the entropy, independently of the new encoding alphabet.

4.2. Scale Downgrading

Once the observer’s interpretation settles onto a specific scale, a selection of the relevant symbols forming the description has been stablished. Accounting for the frequency of each symbol leads to the construction of the symbol probability distribution which characterizes the description. If these probabilities are ordered, a so called symbol probability profile can be obtained. These profiles are by themselves descriptions of the system observed, represented by a shape of $D – 1$ degrees of freedom, where $D$ is the symbolic diversity. When the symbol selection obeys the minimal entropy criterion —that is when the observation is performed according to the fundamental scale—the profile shape may serve for identity purposes due to the uniqueness condition of the minimal entropy probability distribution.

Figure 1: The profile representing a MIDI version of Beethoven’s 9th Symphony 3rd Movement. Figure 1a shows the full fundamental scale version at 2828 different symbols. Figures 1b and 1c represent the degraded versions of the same profile at 513 and 65 symbols respectively.
In order to use the ranked probability profiles for identification purposes or to compare systems observed on different scales, it may be convenient to smooth the frequency profile while preserving its overall shape.

It is possible to smooth a frequency profile while preserving its overall shape. This is done by removing some points from the profile. The selection of points to be removed must be done considering the density of points in each profile segment, in order to keep the same level of detail in all sectors of the profile; this is the basis of the formulation of scale downgrading first presented in [12].

This procedure for downgrading the language scale is useful given the frequent requirement of expressing text descriptions at the same scale or to drive any description’s profile shape, to any arbitrarily selected scale. Figure 1 shows the profile shape representing the Beethoven’s 9th Symphony 3rd Movement, recorded as a MIDI file. Figure 1a shows the frequency profile of the 2828 fundamental symbols which describe the MIDI file at the lowest entropy. Figures 1b and 1c show the profile with scale downgraded to 513 and 65 symbols respectively.

5. SOME TESTS WITH DIFFERENT LANGUAGE EXPRESSIONS

The following sections present tests where the information of several descriptions are computed at different scales of observations. These tests allow for the comparison of the symbolic information resulting from different interpretations of the same descriptions. After recognizing the scope $L$, resolution $R$, and the scale $D$ of the situation, the entropy $h$ corresponding to each one of these observations is computed and presented in tables to draw conclusions.

Figure 2: Entropy $h$ vs. description length $L$ in symbols. Graphs show the relationship between entropy $h$ and scope $L$ (symbol length) for descriptions expressed in: a. English, b: Spanish. Graphs show the impact of the observation scale over the resulting entropy of each observation.

5.1. Natural languages

Figure 2 shows the relationship between entropy and message length measured in symbols, for 128 English speeches and 72 Spanish speeches. Details about the text's contents and properties were published by Febres and Jaffe [12].
We know that representing each one of these speeches in a binary scale, will show about the same number of 0’s and 1’s. As a result the computed symbolic entropy would be near 1 and therefore, not grouping symbols —as the computers would do by forming Bytes of seven or eight bits each— it would be practically impossible to extract information.

The minimal entropy is produced when reading the message at the Fundamental Scale. That does not surprise. The symbols were identified having the minimization of entropy as the objective. But it is worth to mention the dramatic difference between the entropy of these observations at the Fundamental Scale and the observations of the same speeches at any other scale.

5.2. Same symbolic structure. Different perceptions

Both mosaics shown in Figure 3 are built with identical number of pixels. Each is an array of 60x60 pixels some of which are dark or light colored. For both mosaics there are white or grey pixels inducing the interpretation of the figures in one or other manner. But the number of pixels and different colors used are the same, implying their symbolic information is the same.

![Figure 3: Two perceptions of a 2D mosaic with a resolution 60 x 60 pixels. Mosaic (a) shows pixels with 4 different colors. Same color pixels are grouped and separated by white pixels forming triangles. Mosaic (b) shows vertical and horizontal white lines dimmed to grey.](image)

When estimating the account for information expressed in each type, however, there are different accounts which depend on the type of information considered. Interpreting the mosaics as groups of 3136 pixels (56x56) each one represented with a color out of four possible colors. However, changing the focus from single pixels to the larger tiles suggested by the arrangements, Triangles in Figure 3a and bands in Figure 3b, the distribution of the types of information settles on the amounts shown in Table 2. Notice that resolution, scope and scale have different values for the three interpretations of these mosaics, but entropy equals one for all of them due to the uniform distribution of symbol frequencies.
Table 2: Properties of each interpretation of 2D patterns shown in Figure 3.

| Scale name       | Data representation | Fig 3a | Fig.3a | Fig.3b |
|------------------|---------------------|--------|--------|--------|
| Resolution Rhorz | Pixels Symbols      | 56     | 3      | 6      |
| Resolution Rvert | Pixels Symbols      | 56     | 3      | 1      |
| Resolution Rangle| - i                | 4      | 4 ii   | 1 ii   |
| Resolution Rcolor| 4                   | 4 ii   | 2      |
| Scope (Length) L | 3136                | 36     | 6      |
| Scale (Diversity) D| 4                  | 4 ii   | 2 +    |
| Entropy h [0-1]  | 1                   | 1      | 1      |
| Specific diversity d| 0.001             | 0.111  | 0.333  |

1 This degree of freedom doesn’t exist for single pixels

ii Only four angular positions are required.

iii Triangles: light blue, light green, dark blue, dark green.

+ Light band, dark band.

5.3. Partial changes of resolution and scope

Figure 4 uses a 2D example to illustrate the same picture observed for different combinations of resolution and scope. Fig. 4a shows a set of 2D symbols over a ‘surface’ of 27 x 46 pixels. Here the squares have the role of elementary information and each of them may have one of two values: black or white, therefore the number of possible states for each square is 2. There are 27 squares per side, thus the resolution can be approximated to \( R = 27 \times 27 \text{ px} \).

![Figure 4](image-url)

**Figure 4:** Effects of changes of resolution and scope over a 2D representation of polygons. Graphic representation of a language scale downgrading from scale \( D \) to scale \( S \) (\( S < D \)). The total number of symbols at scale \( D \), representing \( D \) different symbols on the top graphs, are transformed in \( S \) different symbols when the language is represented at the scale \( S \), as in the bottom graphs. Also, graphs on the right exhibit greater scope than those on the left.
Using the picture in Figure 4a as reference, Figure 4b increases the scope, Figure 4c reduces the resolution and Figure 4d increases the scope and reduces the resolution simultaneously. The impact of these variations of resolution and scope is presented in Table 3.

Table 3: Balance of information for the 2D example presented Figure 4.

| Figure | Fig. 4a | Fig. 4b | Fig. 4c | Fig. 4d |
|--------|---------|---------|---------|---------|
| Scale name | Pixels | Symbols | Pixels | Symbols | Pixels | Symbols | Pixels | Symbols |
| Data representation | 0's & 1's | Polygons | 0's & 1's | Polygons | 0's & 1's | Polygons | 0's & 1's | Polygons |
| Resolution $R_{horz}$ | 28 | 28 | 46 | 46 | 13 | 13 | 23 | 23 |
| Resolution $R_{vert}$ | 28 | 28 | 28 | 28 | 13 | 13 | 13 | 13 |
| Resolution $R_{angle}$ | - $^i$ | 8 $^{ii}$ | - $^i$ | 8 $^{ii}$ | - $^i$ | 8 $^{ii}$ | - $^i$ | 8 |
| Scope (Length) $L$ | 784 | 8 | 1288 | 10 | 169 | 8 | 299 | 8 |
| Scale (Diversity) $D$ | 2 | 5 $^{ii}$ | 2 | 3 $^+$ | 3 | 4 $^{**}$ | 3 | 5 $^{***}$ |
| Entropy $h$ [0-1] | 0.970 | 0.928 | 0.999 | 0.646 | 0.934 | 0.813 | 0.940 | 0.861 |
| Specific diversity $d$ | 0.003 | 0.625 | 0.002 | 0.300 | 0.018 | 0.500 | 0.010 | 0.625 |

$^i$ This degree of freedom doesn’t exist. Single pixels’ angular position concept degenerates.
$^{ii}$ Only eight angular positions are required to describe symbols represented.
$^{iii}$ Square, triangle, trapezoids, large rectangle, small rectangle.
$^+$ Square, triangle, trapezoids.
$^{**}$ Large Triangle, rectangle, and noise: trapezoids, stairs-like polygon.
$^{***}$ Large Triangle, small triangle, wedge, and noise: trapezoids, stairs-like polygon.

5.4. The impact of reorganizing

This section presents a little experiment. The purpose is to assess the impact of organizing the symbols with which we interpret a description. Figure 5 shows arrays of 30 squares colored with five different intensities; the lightest named as 1 and the darkest named as 5. In the leftmost array, Figure 5a, the squares are randomly organized while Figure 5b shows the colored squares ordered with the darkest at the top-left corner of the array and the lightest at the bottom-right corner. Figures 5c and 4d show the organized distribution of squares indicating different symbols formed by grouping several squares into each type of symbol.

![Figure 5](image.png)

**Figure 5.** Four interpretations of the same distribution of 30 squares colored with five different tones of blue. Number indicate each tone used. The lightest is represented by 1 and the darkest with 5. Each tone appears with the same frequencies in the three graphs. (a) Shows the 30 squared randomly ordered. (b) The squares are ordered according to the rule indicating that no darker square can appear below or at the right of another square. (c) Shows groups of symbols formed by regular shaped lattice of 1 x 6 bricks. (d) Shows with black borders the groups of squares forming symbols to reduce the entropy of this interpretation.
In the leftmost array, Figure 5a, the squares are randomly organized while Figure 5b shows the colored squares ordered with the darkest at the top-left corner of the array and the lightest at the bottom-right corner. Figures 5c and 5d show the organized distribution of squares indicating different symbols formed by grouping several squares into each type of symbol.

This test shows the impact of the interpretation over the distribution of the different types of information. Interpreting Figure 5b as an array of 30 squares, requires just as much symbolic information as the disorganized squares presented in Figure 5a. Despite our unavoidable tendency to appreciate order in Figure 5b, if we consider all squares as independent single symbols, transmitting this information would require the same effort as for transmitting Figure 4a. But if we let our brain to group the squares in repeated patterns by degrees of color intensity (see Figure 5c), we reduce the number of symbols we have to consider and the spatial information associated to them. The possibility for associating semantic information also appears along with the variety of different symbols that now can be arranged. This transference of information from one type to another may be augmented (as illustrated in Figure 5d) or diminished with the grouping of the squares to form symbols of any shape.

### Table 4: Properties of several interpretations of the 2D patterns shown in Figure 5.

| Figure | Fig. 5a | Fig. 5b | Fig. 5c | Fig. 5d |
|---|---|---|---|---|
| Scale name | Symbols | Symbols | Symbols | Symbols |
| Data representation | Single squares | Single squares | Organized squares | Organized squares |
| Resolution $R_{hors}$ | 6 | 6 | 2 | 3 |
| Resolution $R_{vert}$ | 5 | 5 | 5 | 3 |
| Resolution $R_{color}$ | 5$^i$ | 5$^i$ | Varies$^{ii}$ | Varies$^{iii}$ |
| Scope (Length) $L$ | 30$^{iii}$ | 30$^{iii}$ | 10 | 16 |
| Scale (Diversity) $D$ | 5$^i$ | 5$^i$ | 7$^*$ | 6$^{**}$ |
| Entropy $h_{[0-1]}$ | 0.943 | 0.943 | 0.970 | 0.812 |
| Symb. Info. $\mathfrak{H}$ [bits] | 0.167 | 0.167 | 0.700 | 0.375 |

$i$ Different colors for $1 \times 1$ array of squares.
$ii$ Approximation of different combinations of ordered $3 \times 1$ squares.
$iii$ Different combinations of $2 \times 2$, $2 \times 1$, $1 \times 2$ and $1 \times 1$ arrays of ordered squares.
$^*$ $5^4/4 + 10 + 10 + 5$.
$^{**}$ 554 | 543 | 332 | 322 | 321 | 211 | 111.

5.5. Music

Table 5 shows a comparison of information balance of two segments of music recorded in .MP3 format. The two segments correspond to the same fraction of Beethoven’s 5th symphony 1st mov. They are different in the instruments used to play them. One is the full orchestra this piece was written for. The other is played with a piano solo. For each version of the music segment analyzed, three observation scales are used: binary, characters and The Fundamental. The characters’ scale consists of splitting the music-text in single characters. Each character exhibits
a frequency with which entropy is computed. The binary observation can be obtained by substituting each character with its corresponding ASCII number expressed in the binary base. The Fundamental Scale is obtained applying the Fundamental Scale Algorithm [11] which finds the sequences of characters that minimize the overall entropy of the text as interpreted.

The results clearly show that none of the two pieces analyzed is based on characters. While the entropy for both descriptions, computed with a character-based language is close to the maximum, indicating no symbolic information associated, a substantial reduction of entropy was reached when descriptions were read at the Fundamental Scale. Reading these texts was possible only using the Fundamental Scale as a director indicating where to focus. The results indicate that, despite not recognizing any meaning for the Fundamental Symbols encountered, there is way of reading these descriptions, so that some structure implying information emerges from initially unreadable texts.

Figure 7: A tiny fraction of the text which constitutes the Beethoven's 5th symphony 1st movement segment interpreted with orchestra. Extracted from a sound recorded file.

Table 5: Effects of different observation scales over the quantity of information of segment of Beethoven's 5th Symphony versioned with a full orchestra and piano solo.

| Beethoven's: 5th Symphony.1stMov. Segment.Orch | 5th Symphony.1stMov. Segment.Piano |
|-----------------------------------------------|-----------------------------------|
| Scale Name: | Binary | Characters | Fundamental | Binary | Characters | Fundamental |
| Data representation | Zeros and ones | Letters, Punct. and other signs | Recognized min. entropy symbols | Zeros and ones | Letters, Punct. and other signs | Recognized min. entropy symbols |
| Resolutn. R [symb./sec] | 188948 | 23618 | 4517 | 192669 | 24084 | 4241 |
| Scope $L_0$ | 5668432 | 708554 | 135519 | 7514080 | 939260 | 165387 |
| Scope $L_2$ | 5668432 | 5668432 | 1084152 | 7514080 | 7514080 | 1323096 |
| Scale value Diversity $D$ | 2 (0 & 1) | 252 | 4635 | 2 (0 & 1) | 257 | 13808 |
| Symbolic entropy $h$ | near 1 | 0.990 | 0.893 | near 1 | 0.990 | 0.722 |
| Specific Diversity $d$ | 0.049 | 0.00036 | 0.03420 | 0.049 | 0.00027 | 0.08349 |
6. DISCUSSION

6.1. Implications of scale, scope and resolution

The term scale is commonly used in a qualitative manner. Expressions like “individual scale”, “massive scale”, “microscopic scale”, “astronomical scale” and many other similar ones, are typically used to characterize the type of interpretation that should be given to certain descriptions. However, their utility relies on our subjective criteria to adequately apply those expressions. Subsequently, this rather diffused conception of scale is of little use for our purposes. We then propose a quantitative conception of scale. The scale of a system equals the scale of the language used for its description; the scale of the language equals the number of different symbols which constitute the language.

Interestingly, the system’s description scale is determined, in first place, by the observer, and in a much smaller degree by the system itself. The presumably high complexity of a system, functioning with the actions and reactions of a large number of tiny pieces, simply dissipates if (a) the observer, or the describer, fails to see the details, (b) the observer is not interested in the details, and prefers to focus on the macroscopic interactions that regulate the whole system’s behavior, or (c) the system does not have sufficient component diversity, which play the role of symbols here, to refer to each type of piece. It is clear that observing a system at a specified scale implies the use of a certain number of symbols. Hence, the number of different symbols used in a description is linked with our intuitive idea of scale. Therefore, the term ‘scale’ can be used as a descriptor of the combination language-observation by specifying the number of different symbols required to describe the language used for any observation. English, for example, is a 600 thousand word language if described in terms of words, but a 26 letter language if described in terms of letters.

Another important concept with a close relationship to scale is resolution. Resolution is the density of symbols, repeated or not, that participate in a description. Therefore, the resolution separates the description space in many smaller space segments, as many as indicated by the resolution itself. Each space segment must be occupied with a symbol, in other words, even an empty space is a valid symbol to be considered. In spite of the general use of the term resolution, the space-segments need not to be equally sized.

6.2. The Fundamental Scale as a language descriptor

For natural languages reading at the Fundamental Scale allows for a much lower entropy as compared with any of the other scales of reading. This raises the question: Why do we read at the scale of words? A possible answer is that instead of reading at the Fundamental Scale, which exhibits a lower entropy, we prefer the words as symbols due to the amount of non-symbolic information stored on each word. All along their lives, human beings learn, assign and associate meaning to rather complex character-groups —words— built over a selected alphabet of elementary signs. These symbols work afterwards as symbolic units containing a much more information than just the symbolic information linked to Shannon’s symbolic entropy. Agreeing a meaning for each word becomes, therefore, an effective way of coding messages to transmit ideas just by paying the cost of sending a low amount of symbolic information. For natural languages, most of the information, the semantic component, is
already stored in the receiver's mind. Thus, the entropy of the descriptions viewed at the scale of words is high as compared to the entropy computed when the observation is at the Fundamental Scale. This can be explained by observing that words are the type of symbols we humans have selected to minimize the effort needed to convey an idea, taking advantage of the fact that most of the idea is already on the receivers’ mind in the form of semantic information. In contraposition, the structure unveiled by the Fundamental Scale, shows the place where the language itself has decided to settle. It contains information about its grammar structure.

We cannot recognize sounds from the texts representing music (see Figure 7 for a segment of music represented as text). Thus we cannot rely in the concept of word to look for some structure beneath a text file containing the sounds of a musical piece. Analyzing these text files, which we know contain information because we can listen to them, is a painstaking task. Independently of the limited number of notes, rhythms, instruments and any other dynamic effect represented in the music record, the resulting sounds when the piece is played by the performer, are of an enormous, practically unlimited, diversity; music is the result of the superposition of an incredibly large number of components and is full of subtleties which we can hear, but which are not present in the music sheet. The full description of the musical phenomena produces files that, when are discretized, are so diverse in terms of symbols, that looking for musical word equivalents, is fruitless. Yet, applying the Fundamental Scale Algorithm we are able to recognize the symbols building a musical description, and their relative frequencies; the parameters we need to construct a graphical profile with which music can be represented and identified.

7. CONCLUSIONS

Independent and mutually exclusive notions of scale, scope and resolution are provided:

**Scale:** The set of different symbols used in a description. The scale can be numerically expressed as the symbolic diversity of the system’s description interpretation.

**Scope:** The total number of symbols used in a description.

**Resolution:** The density of symbols (alphabet-symbols or encoded symbols) used to create the symbols used in a description.

A quantitative conception of scale is introduced. This conception allows for generalizing Shannon’s information expression to include transmission systems based on more than two symbols; an expression that may be useful when evaluating the convenience of transmitting information by means of non-binary communication systems.

The scale, scope and resolution form an interesting basis to describe the dynamics of our interpretation processes when facing complex descriptions. Considering scale as a model's component, reinforces the idea about our capacity of thinking as linked to our ability to build internal languages with useful symbols. This is a motivating thought since it paves the road to conceive intelligence as the capacity of finding observation scales— or interpretation scales— which lead to a more effective way of understanding nature and its phenomena. Intelligence, or at least part of it, can be regarded as the ability to find the set of symbols which better interprets a message by means of reducing its complexity and thus leading us to a better
understanding of our environment. The association of the concept of scale with the number of symbols of the language employed by the observer to interpret a description, offers advantages for the effective representation of complex systems. It also serves as indication of the mechanisms triggered for our understanding of systems.

The scale concept, along with the methods for scale downgrading, are also promising ways to compare different interpretations of the same systems, as well as comparing different systems at the same scale. The Fundamental Scale allows for the construction of near minimal entropy profiles which may serve as systems identifiers. These tools allow for innumerable experiments which may help in our understanding about the interpretation process.

REFERENCES

[1] F. Heylighen, Modelling Emergence, World Futur. J. Gen. Evol. 31 (1991) 89–104. doi:10.1080/02604027.1991.9972256.
[2] Y. Bar-Yam, A mathematical theory of strong emergence using multiscale variety, Complexity. 9 (2004) 15–24. doi:10.1002/cplx.20029.
[3] Y. Bar-Yam, Multiscale Complexity/Entropy, Adv. Complex Syst. 7 (2004) 47–63. doi:10.1142/S021952590400068.
[4] A. Ryan, Emergence is coupled to scope, not level, Complexity. 13(2) (2007) 67–77. http://arxiv.org/abs/nlin/0609011.
[5] M. Prokopenko, F. Boschetti, A.J. Ryan, An information-theoretic primer on complexity, self-organisation and emergence, Complexity. 15 (2008) 11–28. doi:10.1002/cplx.20249.
[6] N. Fernandez, C. Maldonado, C. Gershenson, Information Measures of Complexity, Emergence, Self-organization, Homeostasis, and Autopoiesis, Springer, Berlin. Heidelgerg, 2014.
[7] R. Piasecki, A. Plastino, Entropic descriptor of a complex behaviour, Physica A. 389 (2010) 397–407.
[8] G.K. Zipf, Human Behavior and the principle of least effort: An introduction to human ecology, Addison-Welesly, New York, 1949.
[9] J. Rissanen, Modelling by the shortest data description, Automatica. 14 (1978) 465–471.
[10] M.H. Hansen, B. Yu, H.H. Mark, Y. Bin, Model Selection and the Principle of Minimum Description Length, J. Am. Stat. Assoc. 96 (2001) 746–774. doi:10.1198/016214501753168398.
[11] G. Febres, K. Jaffe, A Fundamental Scale of Descriptions for Analyzing Information Content of Communication Systems, Entropy. 17 (2015) 1606–1633. doi:10.3390/e17041606.
[12] G. Febres, K. Jaffe, Calculating entropy at different scales among diverse communication systems, Complexity. 0 (2015). doi:10.1002/cplx.21746.