# UNIVERSAL AND COGNITIVE NOTIONS OF 'PART'


José Hernández-Orallo$^{\perp}$

*Universitat Politècnica de València,*
*Dep. de Sistemes Informàtics i Computació,*
*C/ de Vera s/n E-46022, València, Spain*
E-mail: jorallo@dsic.upv.es



**Résumé.** Une question fondamentale de la Science des Systèmes est la notion de *partie* d'un système ou *subsystème*. Toutefois, cette notion s'emploie traditionellement d'une manière non-formelle ou substituée par l'idée de *subensemble*. On introduit une relation universelle de partie qui est basée sur la Complexité de Kolmogorov et une version cognitive basée sur une variante espace-temporelle. Malgré qu'il y a quelques inconvénients techniques, par exemple, la relation de partie n'est pas transitive, ces notions formales peuvent être utilisées pour des situations où l'intuition n'arrive pas, ou d'autres notions analytiques de partie ne sont pas applicables. Aussi, on derive quelques notions intéressantes: de plusiers versions sur l'idée de partition et des niveaux alternatifs i hiérarchiques de descriptions d'un système quelconque. Finalement, on commente leurs applications et l'on les compare avec d'autres notions de partie.



**Mots Clefs**: Partie, Subsystème, Complexité Kolmogorov, Distance Cognitive, Machines de Turing, Reconnaissance de Patrons, Cognition.



**Abstract.** A fundamental question of Systems Science is the notion of system *part* or *subsystem*. However, this notion is traditionally used in an informal way or substituted by the idea of *subset*. We introduce a universal relation of part based on Kolmogorov Complexity and a cognitive version based on a time-space variant. Although there are some technical inconveniences, e.g. the part relation is not transitive, these formal notions can be used in cases where intuition does not suffice or other analytical notions of part are not applicable. Thus, we derive some interesting notions: different versions on the idea of partition and alternative and hierarchical levels of description of any system. Finally we discuss its applications and we compare it with other notions of part.



**Keywords**: Part, Subsystem, Kolmogorov Complexity, Cognitive Distance, Turing Machines, Pattern Recognition, Cognition.


## 1. INTRODUCTION

The idea of part, and its relation with the whole, has motivated many controversial and paradoxical discussions since Aristotle. It is still a fundamental and current issue of investigation and debate in modern Systems Science, especially in dynamical and complex systems, where the claim "the whole is much more than the sum of its parts" is more relentless. However, the idea of part is usually substituted by informal and subjective notions or by other different concepts.

For instance, a collection is an object that can be described from its parts. In this case, the idea of subset is directly applicable. However, most objects are not just collections, and many

$^{\perp}$ Also at the Department of Logic and Philosophy of Science of the University of Valencia.

of their parts are discovered more or less easily a posteriori, by perception or investigation. In general, the idea of subset is not valid for complex systems because there is no unique collection of things that can describe the whole.

It is important to highlight that 'part' is a cognitive notion quite different from the notion of subset. For instance, given a black and white picture $p_{n\times m}$, its left upper quadrant $p(1..n/2, 1..m/2)$ is usually recognised as a part of it. However, if we extract a 'silhouette' of a person from the image, there is no Euclidean way to define that as a part. Even if we 'extract' the pixels into a subset, the pixels must be arranged in a very 'whimsical' way in order to reconstruct the silhouette. On the contrary, an ordered subset of $k$ pixels randomly extracted from the original picture would generally not be recognised as a part of it. The example provides more insights; for instance, if the picture $p_{n\times m}$, has at least a black pixel, is a black pixel picture $p_{1\times 1}$ a subpart of $p_{n\times m}$?

The idea of part is also different from the idea of substring and subsequence (a substring is any sequential subsequence). Some parts are not substrings, for example, given the sequence $x$ = "a000a000a00a0a00a0" , the sequence $y$ = "aaaaaa" is intuitively seen as a part of it but it is not a substring. On the other hand, the difference between a subsequence and a part also arises in many cases. For instance, given the sequence $x$ = "12345678901234567", is the sequence "11" a part of it?

Does this cognitive character of part mean that we must leave the idea of part to subjectivism? Is it relative in the end? Obviously, there is some relativism in the idea of part, as two people do not recognise the same parts in a picture, but there are also *coincidences* too. Here the following question arises: Is *only* cultural knowledge that makes us coincide in the recognition of the same parts, either physical (a leg is a part of a body, the Earth is a part of the solar system, Spain is part of Europe, a step is a part of a stair, ...) or logical (a chapter is a part of a book, a second is a part of a minute, the memories are a part of a person, ...)?

The non-existence of an analytical method to discern this does not mean that there are no absolute and intrinsical properties that are exploited in the cognitive process of recognising a part. In fact, this has been done during more than thirty years of pattern recognition (Watanabe 1972). The same rationale motivated a cognitive but absolute notion of distance, recently formalised in (Bennett et al. 1997); the distance between two objects $X$ and $Y$ can be defined as the maximum between the descriptive information necessary to convert $X$ to $Y$ and the descriptive information necessary to convert $Y$ to $X$. In this way, two negative images turn out to be extremely close from a cognitive point of view (just convert 1's into 0's and viceversa) while extremely far from an Euclidean point of view.

This inspires a different approach for the relation of part: to measure the relative description of the part w.r.t. the description of the whole.

## 2. TOWARDS THE RELATION OF PART

In order to define a relation of part from a descriptive point of view we must first settle the notions of objects, descriptions, and minimal descriptions.

### 2.1. Preliminaries

We choose an alphabet $\Sigma = \{0, 1\}$. An object is any element from $\Sigma^*$, being $\cdot$ the composition operator, usually omitted. The empty string or empty object is denoted by $\varepsilon$. The term $l(x)$ denotes the length or size of $x$ in bits (or elements from $\Sigma$). The relation $<_{lex}$ between two objects denotes precedence in left-to-right lexicographic order, considering $0 <_{lex} 1$.

Under the Church-Turing thesis, any *description* of any object of reality can be converted in a description in a computational machine, so our idea of 'object' is, in this sense, universal.

It is important to realise that whatever perception or knowledge that we can acquire about an object or a system is given in a descriptive way.

The complexity of an object can be measured in many ways, one of them being its degree of randomness (Kolmogorov 1965), which turns out to be equal to the shortest description for it. Descriptional Complexity, Algorithmic Complexity $C(\cdot)$ or Kolmogorov Complexity $K(\cdot)$ formalise this idea, and it has been gradually recognised as a key issue in statistics, computer science, artificial intelligence, epistemology and cognitive science (see e.g. Li & Vitányi 1997).

DEFINITION 2.1. KOLMOGOROV COMPLEXITY

The *Kolmogorov Complexity* of an object $x$ given $y$ on a descriptional mechanism (or bias) $\beta$ is defined as:

$$K_\beta(x|y) = \min \{ \, l_\beta(p) : \phi_\beta(p|y) = x) \, \}$$

where $p$ denotes any "prefix-free" $\beta$-program, and $\phi_\beta(p|y)$ denotes the result of executing $p$ using input $y$.

The complexity of an object $x$ is denoted by $K_\beta(x) = K_\beta(x|\varepsilon)$. It can be seen elsewhere (e.g. Li & Vitányi 1997) that Kolmogorov Complexity is an absolute and objective criterion of complexity, and it is independent (up to a constant term) of the descriptional mechanism $\beta$. In other words, there is an invariance theorem that states that any universal machine can emulate another. For this reason many properties are proven just in an asymptotic way. Throughout the paper we will use the relation $<^+$ as the asymptotical extension of $<$, namely $a <^+ b$ iff there exists a positive constant $k$ such that $a < b + k$.

In the following, we will assume that $\beta$ is selected to observe that $\forall x, y \in \Sigma^*$ $K_\beta(x|y) \geq 1$. If there is no possible confusion, the $\beta$ subscript will be omitted. It is obvious to see that $\forall x \in \Sigma^*$ $K(x|x) <^+ 0$ because there is always a program of constant size of the form "*print the input*". It is also easy to see that $\forall x \in \Sigma^*$ $K(x) <^+ l(x)$ because there is always a program of size less than $l(x)$ plus a constant value of the form "*print x*". In the case that $K(x) \geq l(x)$ we say that $x$ is random. We denote $x^*$ as the first (in lexicographical order) $\beta$-program for $x$ such that $K(x)= l(x^*)$. In this way we can say that an object is random iff $l(x^*) \geq l(x)$. It is obvious that $K(x|x^*) <^+ 0$. However, since $K(\cdot|\cdot)$ is not computable, it is shown elsewhere (e.g. Li & Vitányi 1997) that the other way is just $K(x^*|x) <^+ \log l(x)$.

Finally, although $K(\cdot|\cdot)$ is an absolute measure of information it is not computable because it does not consider time. In this way, $K(\cdot|\cdot)$ does not reflect a cognitive view of information. There have been many proposals to incorporate time to Kolmogorov Complexity. The most appropriate way to weigh space and time of a program, the formula $LT_\beta(p_x) = l(p_x) + \log_2 \text{Cost}(p_x)$, was introduced by Levin in the seventies (see e.g. Levin 1973). In this way, a variant of $K(\cdot|\cdot)$ can be easily defined from it:

DEFINITION 2.2. LEVIN'S LENGTH-TIME COMPLEXITY

The *Levin Complexity* of an object $x$ given $y$ on a descriptional mechanism $\beta$:

$$Kt_\beta(x|y) = \min \{ \, LT_\beta(p) : \phi_\beta(p|y) = x) \, \}$$

This is a very practical alternative of Kolmogorov Complexity, because as well as avoiding intractable descriptions, it is computable. Later on we will come back to this variant to define the cognitive version of part.

## 2.2. Intuitive Properties of Part

Almost every dictionary gives at least two different entries for 'part'. The first one is usually something like "*1. Each thing that results from dividing another thing*" and the second one is like "*2. Thing that, jointly with other or others, composes a 'larger' whole*". From here, we can translate these three properties of a part to a descriptional context:

*a*. A part should be 'smaller' or less complex then the whole.

*b*. A part can be easily described or removed from the whole.

*c*. The whole is described more easily if some of its parts are given.

Property *a* is straightforward under Kolmogorov complexity. Given two objects $\forall X, Y \in \Sigma^*$, $X$ is a part of $Y$ if $K(X) \leq K(Y)$. The other properties are more difficult to define and are addressed in section 3.

## 2.3. 'Desirable' Technical Properties of a Relation of Part

Inspired in the notion of set inclusion, we want to define an alternative structural notion for part, that we will denote with $\subseteq$. Some properties that would be convenient for this relation of part $\subseteq$ are:

1. $\forall X \in \Sigma^*,\ \varepsilon \subseteq X$ *(the empty object is part of any object)*
2. $\forall X \in \Sigma^*,\ X \subseteq X$ *(reflexive)*
3. $\forall X, Y, Z \in \Sigma^*,\ (X \subseteq Y \wedge Y \subseteq Z) \Rightarrow (X \subseteq Z)$ *(transitive)*
4. $\forall X \in (\Sigma^* - \varepsilon),\ X \not\subseteq \varepsilon$ *(the empty object has no parts but itself)*
5. $\forall X, Y \in \Sigma^*,\ (X \subseteq Y \wedge Y \subseteq X) \Rightarrow (X = Y)$ *(antisymmetric)*

However, not all of these properties are equally important. In particular, properties 1, 2 and 4 are basic, but property 3 is sometimes not so intuitive. For instance, a person is a part of a country but an ear is not a part of a country. Property 5, although almost always true for physical objects, is not true in some mathematical objects (e.g. the fractals).

Finally, there is another property which is given in any lattice and it could also be studied: the existence of a maximal element.

# 3. UNIVERSAL PART

In this section we present five different definitions of part based on Kolmogorov Complexity. The first two are based on property b), and the third one is based on property c). In the end we present a corrected version that integrates properties a), b) and c). In all cases, we study their technical properties and their meaning.

## 3.1. Removable Versions

If we recall property b) "a part can be easily *described* or *removed* from the whole", there are many ways to measure this *easiness* of description. Relative Kolmogorov Complexity $K(X|Y)$ seems the best way to compare $X$ with $Y$. If we compare it with an absolute value, i.e. a constant, we would have that any small object would be a part of anything. If we make the comparison w.r.t. the complexity of the whole, any small object (wrt. the whole) would be a part of it, although it could be completely unrelated.

As a result, $K(X|Y)$ must be compared with the complexity of $K(X)$:

DEFINITION 3.1. REMOVABLE UNIVERSAL PART

An object $X$ is a removable universal part of an object $Y$ in $\beta$, denoted by $X \overset{\downarrow}{\trianglelefteq}_{\beta} Y$, iff:

$$K_\beta(X|Y) \leq \log K_\beta(X)$$

The meaning of the definition is that the part is significantly easier to describe if we have the whole that if we must describe the part from scratch. The use of the logarithm is a very convenient way for measuring this "significantly easier", although other function could be used (e.g. the square root).

Many results from these definitions are intuitive. For instance, given an image with sufficient complexity, any characteristic part can easily be described from it. However, the notion of substring does not always match with that of part. As an example, take the object X = $0^n1^n$ where $n$ is a very long and highly compressible number (e.g. $64^{64}$) and take the substring Y of X composed of 56231451 consecutive 0's and 924515 consecutive 1's. The object X is not useful for describing Y, so $X \not\trianglelefteq^{\downarrow}_{\beta} Y$.

It is interesting to compare the definition of part with the notion of subset. Here, we will show that definition 3.1 has some technical inconveniences:

**Theorem 3.2.** The empty object is not always a part of any non-empty object. In other words, property 1 does not hold for $\trianglelefteq^{\downarrow}_{\beta}$.

**Proof**. There are universal description mechanisms where the length $l(p)$ of the shortest program that outputs the empty string (more properly it halts) for every input $X$ is greater than $k$. Consequently, $\exists X \in \Sigma^*$, $K_\beta(\varepsilon|X) \geq l(p)$, but also $\log K_\beta(\varepsilon) = \log K_\beta(\varepsilon|\varepsilon) < \log l(p)$. Obviously, $\exists X \in \Sigma^*\ K_\beta(\varepsilon|X) \not\leq K_\beta(\varepsilon)$. □

**Theorem 3.3.** Some objects are not parts of themselves. In other words, property 2 does not hold for $\trianglelefteq^{\downarrow}_{\beta}$.

**Proof**. Just choose $X = \varepsilon$. Thus, $K_\beta(\varepsilon|\varepsilon) = K_\beta(\varepsilon) \not\leq \log K_\beta(\varepsilon)$. □

Although counterexamples for property 2 (and 4) only happen for very simple objects, the problems for property 1 (and 3) are much more general because for most universal machines, $K_\beta(\varepsilon|X) = K_\beta(\varepsilon)$.

By looking at the counterexamples, a first idea to solve them can be to make the comparison asymptotical, i.e.:

DEFINITION 3.4. ASYMPOTICAL UNIVERSAL PART

An object $X$ is an asymptotic universal part of an object $Y$ in $\beta$, denoted by $X \subseteq_\beta^+ Y$, iff:

$$K_\beta(X|Y) <^+ \log K_\beta(X)$$

As a result, properties 1 and 2 hold:

**Theorem 3.5.** The empty object is always a part of any non-empty object. In other words, property 1 does hold for $\subseteq_\beta^+$.

**Proof**. Whatever universal description mechanism there is always a program of constant size $l(p)$ such that outputs the empty string (more properly it halts). Consequently, $\forall X \in \Sigma^*$, $K_\beta(\varepsilon|X) \leq l(p)$. Obviously, there exists a $k = l(p) + 1$ such that $\forall X \in \Sigma^*\ K_\beta(\varepsilon|X) < \log K_\beta(\varepsilon) + k$. □

**Theorem 3.6.** Any object is always a part of itself, i.e., property 2 does hold for $\subseteq_\beta^+$.

**Proof**. Whatever universal description mechanism $\forall X \in \Sigma^*$, there is always a program of constant size $l(p)$ of the form "PRINT THE INPUT", so $K_\beta(X|X) = l(p)$. Once again, there exists a $k = l(p) + 1$ such that $\forall X \in \Sigma^*\ K_\beta(X|X) = l(p) < \log K_\beta(X) + k$. □

However, if we fix the constant of the asymptotical relation,

**Theorem 3.7.** Property 3 does not hold for $\subseteq_\beta^+$.

**Proof**. Select $X, Y, Z$ random, namely $K_\beta(X) = l(X)$, $K_\beta(Y) = l(Y)$ and $K_\beta(Z) = l(Z)$, and $Z = rYs$ and $Y = tXu$ such that $r,s,t$ and $u$ are also random and $l(r) < \log l(Y)$, $l(t) < \log l(X)$, $l(r) + l(t) > \log l(X)$, and $l(s) + l(u) > \log l(X)$. In this case we have that $X \subseteq_\beta^+ Y$ because $K_\beta(X|Y) = K_\beta(X|tXu) <^+ l(t) < \log l(X) = \log K_\beta(X)$ and $Y \subseteq_\beta^+ Z$, because $K_\beta(Y|Z) = K_\beta(Y|rYs) <^+ l(r) < \log l(Y) = \log K_\beta(Y)$.

But we have that $Z = rtXus$ and $r, s, t, u, X, Y$ and $Z$ are random so $K_\beta(X|Z) \geq min(l(r) + l(t)\ ,\ l(u) + l(s)) > \log l(X)$ and hence $X \not\subseteq_\beta^+ Z$. □

It is easy to see that property 4 does not hold either. Definition 3.4 may be interesting for infinite objects, e.g., mathematical entities, however, we are more concerned with finite objects. This will discard any asympotical solution in the following.

Without more discussion for the moment on the applicability of the first definition (3.1), we are going to exploit the other side of the coin, property *c*:

## 3.2. Constructive Versions

A direct formalisation of property *c* would make that a part would be anything that helps to construct an object, although this would entail that a superobject would be a part of any object. For this reason we must be more restrictive in this property:

DEFINITION 3.8. CONSTRUCTIVE UNIVERSAL PART

An object $X$ is a constructive universal part of an object $Y$ in $\beta$, denoted by $X \overset{\uparrow}{\unlhd}_{\beta} Y$, iff:

$$K_\beta(Y|X) \leq K_\beta(Y) - K_\beta(X) + \log K_\beta(X) + 1$$

This refines property *c*, because the cost of constructing the whole diminishes in the same amount of the given part. The intention of the term $-K_\beta(X) + \log K_\beta(X) + 1$ is to force that almost all the information of $X$ is used for constructing $Y$. Moreover, this is done without much adaptation cost, a cost that is represented by $\log K_\beta(X) + 1$.

It can be contrasted with many examples that definition 3.8 behaves in a similar way as definition 3.1. This is justified by the following theorem:

**Theorem 3.9.** Definitions 3.1 and 3.8 are equivalent up to an additive term $O(\min\{\log K(X), \log K(Y)\})$.

**Proof**. The proof is direct from the following theorem (Li & Vitanyi 1997):

$K_\beta(X) - K_\beta(X|Y) = K_\beta(Y) - K_\beta(Y|X) \pm O(\min\{\log K(X), \log K(Y)\}$

Under this additive term it holds from definition 3.8, $K_\beta(Y|X) \leq K_\beta(Y) - K_\beta(X) + \log K_\beta(X) + 1$ and the previous theorem, that $K_\beta(Y) + K_\beta(X|Y) - K_\beta(X) \leq K_\beta(Y) - K_\beta(X) + \log K_\beta(X)$ and definition 3.1. $K_\beta(X|Y) \leq \log K_\beta(X)$ follows immediately. ☐

However, since we are interested in finite objects, it is interesting to study the properties of definition 3.8 in more detail. The properties do not hold for any universal machine.

### Transparent Universal Machines

Fortunately we can *wrap* any universal machine such that $\forall X \in \Sigma^* : k > K_\beta(\varepsilon|X) > K_\beta(\varepsilon)$.

DEFINITION 3.10. *K*-TRANSPARENT UNIVERSAL MACHINE

Given any universal machine $\phi$, we can make a *k-transparent* universal machine $\beta^\phi$ from it, with $k \geq 5$, in the following way:

| Program | Input | Output |
|---|---|---|
| $1x$ | $\varepsilon$ | $\varepsilon$ |
| $1x$ | $y \neq \varepsilon$ | *does not halt* |
| $01x$ | $y$ | $\varepsilon$ |
| $0001x$ | $y$ | $y$ |
| $0^{k-1}1x$ | $y$ | $\phi(x|y)$ |
| *any other program does not halt.* | | |

It is easy to show from it that $\forall X \in \Sigma^*, X \neq \varepsilon, 2 = K_\beta(\varepsilon|X) > K_\beta(\varepsilon) = 1$ and $K_\beta(X|X) = 4$, which allows to prove the following three properties for definition 3.8:

**Theorem 3.11.** The empty object is always a part of any non-empty object. In other words, property 1 does hold for $\overset{\uparrow}{\unlhd}_{\beta}$.

**Proof**. $K_\beta(Y|\varepsilon) = K_\beta(Y) \leq K_\beta(Y) - K_\beta(\varepsilon) + \log K_\beta(\varepsilon) + 1$ since $K_\beta(\varepsilon) = 1$. ☐

**Theorem 3.12.** Any object is always a part of itself. In other words, property 2 does hold for $\overset{\uparrow}{\subseteq}_\beta$ if $\beta$ is at least 8-transparent.

**Proof**. If $X \neq \varepsilon$ then $K_\beta(X|X) = 4 \leq K_\beta(X) - K_\beta(X) + \log K_\beta(X) + 1 = \log K_\beta(X) + 1$ and this holds because $K_\beta(X) \geq 8$ then $\log K_\beta(X) + 1 \geq 4$. If $X = \varepsilon$ then $K_\beta(\varepsilon|\varepsilon) = 1 \leq K_\beta(\varepsilon) - K_\beta(\varepsilon) + \log K_\beta(\varepsilon) + 1 = \log K_\beta(\varepsilon) + 1 = \log 1 + 1 = 1$. □

**Theorem 3.13.** The empty object has no parts but itself. In other words, property 4 does hold for $\overset{\uparrow}{\subseteq}_\beta$ if $\beta$ is transparent.

**Proof**. if $X \neq \varepsilon$ then $K_\beta(\varepsilon|X) = 2 \nleq K_\beta(\varepsilon) - K_\beta(X) + \log K_\beta(X) + 1 = 2 - K_\beta(X) + \log K_\beta(X)$ since $\forall X \in \Sigma^*\ K_\beta(X) \geq 1$. □

It is not difficult to find counterexamples for the transitive relation, however, we can see that if $X \overset{\uparrow}{\subseteq}_\beta Y$, $Y \overset{\uparrow}{\subseteq}_\beta Z$, then $K_\beta(Z|X) <^+ K_\beta(Z|Y) + K_\beta(Y|X) \leq K_\beta(Z) - K_\beta(Y) + \log K_\beta(Y) + K_\beta(Y) - K_\beta(X) + \log K_\beta(X) = K_\beta(Z) + \log K_\beta(Y) - K_\beta(X) + \log K_\beta(X)$. That is to say, $\overset{\uparrow}{\subseteq}_\beta$ is sometimes not transitive but just about $\log K_\beta(Y)$ and a small constant.

## 3.3. Joint Versions

Although theorem 3.9 says that definitions 3.1 and 3.8 are equivalent up to an additive term $O(\min\{\log K(X), \log K(Y)\})$, they can be quite different for small objects. It would be interesting to use both definitions in a unified way. We have seen that the constructive version is compliant with properties 1,2, and 4 if we choose a $\beta$ to be 8-transparent. However, the removable version is not compliant with them. The easiest way to make it follow properties 1,2, and 4 is:

DEFINITION 3.14. CORRECTED REMOVABLE UNIVERSAL PART

The relation "corrected removable universal part", denoted by $X \overset{\downarrow}{\subseteq}'_\beta Y$, is defined as:

$$X \overset{\downarrow}{\subseteq}'_\beta Y = X \overset{\downarrow}{\subseteq}_\beta Y \text{ iff } X \neq Y,\ X \neq \varepsilon,\ Y \neq \varepsilon$$

1. $\forall X \in \Sigma^*,\ \varepsilon \overset{\downarrow}{\subseteq}'_\beta X$
2. $\forall X \in \Sigma^*,\ X \overset{\downarrow}{\subseteq}'_\beta X$
4. $\forall X \in (\Sigma^* - \varepsilon),\ X \overset{\downarrow}{\nsubseteq}'_\beta \varepsilon$

Finally, we are able to define a joint version:

DEFINITION 3.15. JOINT UNIVERSAL PART

The object $X$ is a joint universal part of an object $Y$ in $\beta$, denoted by $X \subsetneq_\beta Y$, iff $X \overset{\uparrow}{\subseteq}_\beta Y$ and $X \overset{\downarrow}{\subseteq}'_\beta Y$

It is obvious that this definition follows properties 1, 2 and 4.

One of the nice properties of this joint definition is that it avoids that the program $x^*$ would always be a part of $x$. In the case of $\overset{\uparrow}{\subseteq}$, it would be very likely, because $K(x|x^*) = 0 + k$, with $k$ being very small. However, according to $\overset{\downarrow}{\subseteq}'$, the disequality $K(x^*|x) < \log l(x) + k'$ is very close to the real limit and the constant is large; this would rarely be less than $\log K(x^*)$ and it would not be a part.

On the other hand, it avoids that $x$ would be always a part of the program $x^*$. According to $\overset{\downarrow}{\subseteq}'$, it would be very likely, because $K(x|x^*) = 0 + k$, with $k$ being very small. However, according to $\overset{\uparrow}{\subseteq}$, $K(x^*|x) < \log l(x) + k'$. In order to be less than $K(x^*) - K(x) + \log K_\beta(x) + 1$, since the first two terms are similar, $x$ should be almost random to make $K(x) > l(x)$. In this case, $x^*$ is "PRINT $x$" and $K(x^*|x)$ is close to 0. This is highly intuitive since a program "PRINT $x$" has as part "$x$".

Now, we will address property 5, antisymmetry. A first idea is to add to definition 3.15 the property a), i.e. $K(x) \leq K(y)$. With this solution there would be still many pairs of objects such that $X \subsetneq_\beta Y$ and $Y \subsetneq_\beta X$ (for instance two images one the negative of the other) and $X \neq Y$. With or without property a), these objects could be said to be isomorphic, denoted by $X \equiv Y$.

However, $\equiv$ is not and cannot be made transitive because all the objects with the same size would collapse in the same equivalence class.

A better way to make $\subseteq_\beta$ antisymmetric without using property a) is to define $X$ as a part of $Y$ iff $X \subseteq_\beta Y$ and (if $Y \subseteq_\beta X$ then $((K(X|Y) < K(Y|X) \vee (K(X|Y) = K(Y|X) \wedge X \leq_{lex} Y))$.

Finally, we answer negatively to the question whether $\subseteq_\beta$ has a maximal object:

**Theorem 3.16.** There is no maximal object $\Omega$ such that $\forall X \in \Sigma^* \; X \subseteq_\beta \Omega$.

**Proof**. Using $\overset{\downarrow}{\trianglelefteq}$, if $\Omega$ is finite then it is obvious since a random and independent string $x$ much greater than $\Omega$ can be chosen. If $\Omega$ is infinite every $X$ requires an index or program to use different information from $\Omega$. By a simple counting argument, this is impossible, because all the objects of length less than $n$, we would require $2^n$ different indexes, which can only be expressed with a mean length $n$, which is not less then $\log K(X) \leq \log n$. The constructive version gives the same result. ☐

It can be derived from this theorem that very small objects are difficult to be part of very large objects. The rationale is the fact that is necessary to index 'where' the part is, and this may be quite large for a large object.

# 4. PARTITIONS

One of the first concepts that can be derived from the notion of part is the notion of partition. The following definitions can be particularised for the different relations we have been seeing with slightly different interpretations. We will refer generically as any of them with the relation symbol $\subseteq$.

First of all, we require a useful relation:

DEFINITION 4.1. PROPER UNIVERSAL PART

An object $y$ is a proper part of an object $x$ in $\beta$, denoted by $y \subset_\beta x$, iff $y \subseteq_\beta x$ but $x \not\subseteq_\beta y$.

From here we are able to define the notion of partition and reduced partition:

DEFINITION 4.2. PARTITION

A set of objects $Y = \{ y_1, y_2, ..., y_m \}$ is a partition of an object $x$ in $\beta$ iff $\forall y_i$: $1 \leq i \leq m$: $y_i \subseteq x$ and there exists an ordering $o_j$ of $Y$ such that $x \subseteq y_{o1} \cdot y_{o2} \cdot ... \cdot y_{om}$.

DEFINITION 4.3. REDUCED PARTITION

The set of objects $Y$ is a reduced partition of an object $x$ in $\beta$ iff it is a partition of $x$ in $\beta$ and $\neg\exists Y'$ subset of $Y$ such that $Y'$ is a partition of $x$.

It is easy to prove that for every object $X$ there is always a reduced partition $Y = \{ X \}$ because $X \subseteq X$. However, this is not true for a proper partition:

DEFINITION 4.4. PROPER PARTITION

A set of objects $Y = \{y_1, y_2, ..., y_m\}$ is a proper partition of an object $x$ in $\beta$ iff $\forall y_i$: $1 \leq i \leq m$: $y_i \subset x$ and there exists an ordering $o_j$ of $Y$ such that $x \subseteq y_{o1} \cdot y_{o2} \cdot ... \cdot y_{om}$.

From here we can define the notion of cohesion of an object. Intuitively, an object abaabbaaabbbaaaabbbb.. $^n$..$a^k b^k$ is very cohesive, because any partition usually requires more information than the whole. On the contrary $1^n 0^m$ is not cohesive, because it is natural to divide it into $1^n$ and $0^m$. The following three definitions give three distinct degrees of cohesion:

DEFINITION 4.5. EASY PARTITION

A proper partition $Y = \{y_1, y_2, ..., y_m\}$ of $X$ is easy iff there exists an ordering $o_j$ of $Y$ :

$$K(y_{o1} \cdot y_{o2} \cdot ... \cdot y_{om} \mid x) < \log K(x)$$

DEFINITION 4.6. NATURAL PARTITION

A proper partition $Y = \{y_1, y_2, ..., y_m\}$ of $X$ is natural iff for every ordering $o_j$ of $Y$ :

$$K(y_{o1} \cdot y_{o2} \cdot ... \cdot y_{om} \mid x) < \log K(x)$$

DEFINITION 4.7. CHARACTERISTIC PARTITION

A proper partition $Y = \{y_1, y_2, ..., y_m\}$ of $X$ is characteristic iff:

$$K(y_1) + K(y_2) + ... + K(y_m) \leq K(x)$$

For instance, for $x=1^n$, we have that $K(x) = \log n + c$, being $c$ small. In this case definition 4.5 and 4.6 coincide and any partition of $1^n$, once concatenated, is equal to $x$, so it makes that $K(y_{o1} \cdot y_{o2} \cdot ... \cdot y_{om} \mid x) = c' < \log K(x) = \log (\log n + c)$ which is quite probable if $n$ or $c$ are large. On the contrary, it is difficult to find a characteristic partition for $1^n$. Finally, it is easy to see that $1^n0^m$ has also an easy and natural partition, and for many descriptional mechanisms it has also a characteristic partition $1^n$ and $0^m$. As a conclusion from these examples, definitions 4.5 to 4.7 should be understood asymptotically or under a particular descriptional mechanism, in order to make more sense from them.

Finally, there are other many notions that could be extracted and studied in detail from the notion of partition. For example, two alternative levels of separation could be defined if there are two partitions $Y$ and $Y'$. They would be two disjoint levels of separation if none of the elements $Y$ or $Y'$ are parts of an element of the other partition. On the other hand,

DEFINITION 4.8. THINNER PARTITIONS

A partition $Y = \{y_1, y_2, ..., y_m\}$ of $X$ is thinner than a partition $Y' = \{y'_1, y'_2, ..., y'_n\}$ of $X$ iff:

$$\forall y_i \in Y \ \exists y'_j \in Y' : y_i \subseteq y'_j$$

This last definition makes it possible for a hierarchisation of ontological levels of a system.

# 5. COGNITIVE PART

The definitions of section 3 are said to be universal because the time is not considered. However, any cognitive system is resource-bounded, especially by time. A part of an object could be discovered after a long observation; if this time is extremely long, it would not probably considered as a *cognitive* part. For this reason, we can easily adapt the definitions of section 3 to the time-considering version of $K(\cdot)$, Levin Complexity: $Kt(\cdot)$.

DEFINITION 5.1. REMOVABLE COGNITIVE PART

An object $X$ is a removable cognitive part of an object $Y$ in $\beta$, denoted by $X \,{}^{t}\!\overset{\downarrow}{\subseteq}_{\beta} Y$, iff:

$$Kt_\beta(X|Y) \leq \log Kt_\beta(X)$$

The interpretation of this definition is slightly different from definition 3.1, because the part is significantly easier to describe in space and *time*. The inclusion of time makes clearer some phenomena that already appeared in definition 3.1. For instance, there is a low probability to find a removable cognitive part $X$ of $Y$ with $l(X) + \log l(X) < \log l(Y)$. The rationale is asympotical. Since the term $Kt_\beta(X|Y)$ includes the time of reading $Y$, then $Kt_\beta(X|Y) > \log l(Y)$. By definition $Kt_\beta(X) <^{+} l(X) + \log l(X)$. So we have that $Kt_\beta(X|Y) >^{+} Kt_\beta(X)$.

Although the major advantage of definition 5.1 is that it is computable, it has the same problems with properties 1,2 and 4.

In contrast, the constructive variant is free from most of these problems:

DEFINITION 5.2. CONSTRUCTIVE COGNITIVE PART

An object $X$ is a constructive cognitive part of an object $Y$ in $\beta$, denoted by $X \,{}^{t}\!\overset{\uparrow}{\subseteq}_{\beta} Y$, iff:

$$Kt_\beta(Y|X) \leq Kt_\beta(Y) - Kt_\beta(X) + \log Kt_\beta(X) + 1$$

Again we must prove the subsequent properties by the use of transparent machines:

**Theorem 5.3.** The empty object is always a part of any non-empty object. In other words, property 1 does hold for ${}^{t}\!\overset{\uparrow}{\subseteq}_{\beta}$.

**Proof**. $Kt_\beta(Y|\varepsilon)=Kt_\beta(Y) \leq Kt_\beta(Y)-Kt_\beta(\varepsilon)+\log Kt_\beta(\varepsilon) +1$ since $Kt_\beta(\varepsilon) = 1$. □

**Theorem 5.4.** Any object is always a part of itself. In other words, property 2 does hold for ${}^{t}\!\overset{\uparrow}{\subseteq}$ if $\beta$ is 32-transparent.

**Proof**. If $X \neq \varepsilon$ then $Kt_{\beta}(X|X) = 4 + \log 4 = 6 \leq Kt_{\beta}(X) - Kt_{\beta}(X) + \log Kt_{\beta}(X) + 1 = \log Kt_{\beta}(X) + 1$ and this holds because $Kt_{\beta}(X) \geq 32 + \log 32$ then $\log Kt_{\beta}(X) + 1 \geq (\log 37) + 1 > 6$. If $X = \varepsilon$ then $Kt_{\beta}(\varepsilon|\varepsilon) = 1 + \log 1 = 1 \leq Kt_{\beta}(\varepsilon) - Kt_{\beta}(\varepsilon) + \log Kt_{\beta}(\varepsilon) + 1 = \log Kt_{\beta}(\varepsilon) + 1 = \log (1 + \log 1) + 1 = 0 + 1 = 1$. □

**Theorem 5.5.** The empty object has no parts but itself. In other words, property 4 does hold for $^{t}\overset{\uparrow}{\subseteq}$. if $\beta$ is transparent.

**Proof**. if $X \neq \varepsilon$ then $Kt_{\beta}(\varepsilon|X) = 2 + \log 2 \not\leq Kt_{\beta}(\varepsilon) - Kt_{\beta}(X) + \log Kt_{\beta}(X) + 1 = 1 + \log 1 - Kt_{\beta}(X) + \log Kt_{\beta}(X) + 1$ since $\forall X \in \Sigma^* \; Kt_{\beta}(X) \geq 1$. □

Once again, property 3 does not hold for $^{t}\overset{\uparrow}{\subseteq}$. However, we can see that if $X \,^{t}\overset{\uparrow}{\subseteq}\, Y$, $Y \,^{t}\overset{\uparrow}{\subseteq}\, Z$, then $Kt_{\beta}(Z|X) <^{+} Kt_{\beta}(Z|Y) - \log l(Y) + Kt_{\beta}(Y|X)$ because in $Kt_{\beta}(Z|X)$ the object $Y$ must only be read once, and $l(Y)$ units of time are saved. Then we have $Kt_{\beta}(Z|Y) - \log l(Y) + Kt_{\beta}(Y|X) \leq Kt_{\beta}(Z) - Kt_{\beta}(Y) + \log Kt_{\beta}(Y) + Kt_{\beta}(Y) - Kt_{\beta}(X) + \log Kt_{\beta}(X) + 2 - \log l(Y) = Kt_{\beta}(Z) - Kt_{\beta}(X) + \log Kt_{\beta}(X) + (\log Kt_{\beta}(Y) - \log l(Y) + 1)$. Since $\forall Y$, $Kt_{\beta}(Y) < l(Y) + \log l(Y) + c$, we have that $\log Kt_{\beta}(Y) - \log l(Y) + 1 < \log (l(y) + \log l(y) + c) - \log l(Y) + 1 < \log \log l(Y) + c'$. This makes $^{t}\overset{\uparrow}{\subseteq}$ transitive up to a term $\log \log l(Y)$ being $Y$ the *intermediate* object.

## 6. CONCLUSIONS

Different definitions of part have been introduced based on variants of descriptional complexity. We have studied their properties and we have shown those particular cases where they do not match with intuition. However, they are the first general and formal definitions of part which only require the description of the whole to discern what and what is not a part of it. Of course, they must be better studied with more space and specialised to particular frameworks in order to see their efficacy.

In the immediate future, it should be interesting to study in more detail the different versions of partitions and their application to Systems Science. Definition 5.1 could be refined by adding $\log l(Y)$ and properties 1, 2 and 4 should be studied after this inclusion under specific machines.

Finally, we envisage a very appealing definition for computer science:

DEFINITION 6.1. SUBPROGRAM (OR SUBTHEORY)

The object $y$ is a subprogram of an object $x$ in $\beta$ iff $y \subseteq_{\beta} x$ and $\phi(y) \subseteq_{\beta} \phi(x)$

## REFERENCES


Balcázar, J.; Díaz, J.; Gabarró, J., *Structural Complexity*, Springer Verlag, vol.1., 1988.

Beneyto, R., Status Epistemológico de la Sistémica, 1994

Bennett, C.H.; Gács, P.; Li, M.; Vitanyi, P.M.B.; Zurek, W.H. Information Distance, *IEEE Trans. Information Theory*, IT-44:4, 1407-1423, 1998.

Bertalanffy, L.V., *Teoria generale dei sistemi*, Istituto Librario Internationale, Milano 1971.

Blum, M., A machine-independent theory of the complexity of recursive functions, *J. ACM* 14, 4:322-6, 1967.

Chaitin, G. J., *Algorithmic Information Theory*, fourth printing, Cambridge University Press, 1992.

Forrester, J.W., *Principles of Systems*, Wright-Allen, Cambridge 1968

Herken, R. *The Universal Turing Machine*, Oxford University Press, 1994.

Kolmogorov, A.N., Three Approaches to the Quantitative Definition of Information, *Problems Inform. Transmission*, 1(1):1-7, 1965.

Kolmogorov, A.N., Logical basis for information theory and probability theory, *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 662-664, sept. 1968.

Levin, L.A., Universal search problems, Problems Inform. Transmission, 9:265-266, 1973.

Li, M.; Vitanyi, P., *An Introduction to Kolmogorov Complexity and its Applications*, 2nd Ed. Springer, 1997.

Watanabe, S., Pattern recognition as information compression, in: Watanabe (ed.) *Frontiers of Pattern Recognition*, New York: Academic Press, 1972.