SHANNONIAN, SEMANTIC AND PRAGMATIC GEOINFORMATION

Juval Portugali
ESLab (Environmental Simulation Lab), Department of Geography and the Human Environment,
Tel Aviv University, Tel-Aviv 69978, Israel, juval@post.tau.ac.il, http://eslab.tau.ac.il,
Tel: 972(+3)640-8661, Fax: 972(+3)640-6243

Abstract

The distinction between the notions of Shannonian vs. Semantic information is due to Haken in his book Information and Self-Organization. Semantic information refers to the quality, that is, meaning, afforded by a message; while Shannonian information to sheer quantity – to “information with meaning exorcised”. In a recent paper “The face of the city is its information”, Haken and Portugali (2003) have shown, first, that different elements in the city afford different levels of information. Second, that the Shannonian information afforded by the various urban elements can be measured by means of Shannon’s theory of information. Third, that the very ability to do so depends on a qualitative cognitive process of categorization that entails semantic information. Fourth, that the above three properties are part of humans’ cognitive capabilities. The aim of this paper is twofold: first, to extend our theory to geographical elements at large. Second, to explore the implications thereof to the domain of geoinformatics.

INTRODUCTION

The distinction between the notions of Shannonian information vs. Semantic information is due to Haken (1988/2000) in his book Information and Self-Organization. Semantic information refers to the quality, that is, meaning, afforded by a message; this is the way we commonly use to term information in everyday language. The notion of Shannonian information, on the other hand, refers to the usage of the term information as implied by Shannon’s theory of information (Shannon and Weaver 1949). According to this theory information is a sheer quantity – or as Haken (ibid) puts it: it is “information with meaning exorcised”. Shannon has suggested several ways by which information can be measured the most common one is by means of information bits:

\[ I = \log_2 Z \]  

when Z is the number of possible states the system can take. For example, is the case of rolling a die, Z=6 and I – the quantity of Shannonian information enfolded in the process of rolling of a die – is about 2.5 bits. As can be seen from this definition, Shannonian information is not related to any meaning.

Given such measures one can now determine the amount of information, going through a given information channel, processed by a computer or conveyed by a spoken or written message, regardless of the meaning of this information.

It is hard to exaggerate the importance of Shannon Information theory. It was central to the development of computer technology and science, communication and information...
sciences, cognitive sciences, to name but a few of the domains that are specifically relevant to this conference.

And what about Geoinformation? As can be judged by the variety of presentations in this conference, as well as from the literature at large, students of geoinformation tend to use the notion information as it is commonly employed in everyday language, that is, as data and as semantic information. The formalism of Shannon’s theory as well as the possibility that there are different forms of information (Shannonian, semantic and probably more) are usually not considered.

In a recent paper “The face of the city is its information”, H&P (Haken and Portugali, 2003) have shown, first, that different elements in the city afford different levels of information. Second, that the Shannonian information afforded by the various urban elements can be measured by means of Shannon’s theory of information. Third, that the very ability to do so depends on a qualitative cognitive process of categorization that entails semantic information. Fourth, that the above three properties are part of humans’ cognitive capabilities. The aim of this paper is twofold: first, to extend Haken and Portugali’s theory to geographical elements at large. Second, to explore the implications thereof to the domain of geoinformatics. In the process I’ll introduce a third notion of information, namely, pragmatic information that refers to action.

**SHANNONIAN INFORMATION**

H&P have developed their theory regarding Shannonian information of the city in light of two cognitive science’s applications of information theory: One is Miller’s (1956) seminal paper “The magic number seven plus or minus two: Some limits on our capacity for processing information”. As the title of Miller’s study indicates, he brings evidence demonstrating that there is a limit to humans’ capacity to process information in short term memory. Namely, it is about 2.5 bits.

The second cognitive science’s application is associated with attempts to apply information theory to Gestalt psychology in order to quantify figurative goodness. These studies indicate that different abstract forms transmit different quantities of information due to their very structure. In order to illustrate this property H&P have produced Figure 1 that illustrates and compares the Shannonian information conveyed by two similar patterns.

![Figure 1: The Shannonian information conveyed by two similar patterns. Source: Haken and Portugali (2003), figure 6.](image)

On the right hand side we see a pattern of dots that can be perceived in one way only, that is, one out of one possibility. Its Shannonian information I is thus:

\[ I = \log_2 1 = 0. \]

On the left-hand side we see the same spatial distribution, but of triangles instead of dots. Here, however, there are three different ways to perceive that pattern, that is, one out of three possibilities. Its Shannonian information is therefore:
In their paper, H&P further show that the same applies to the urban landscape: different urban elements in the city convey different levels of information that can be quantified by means of Shannon’s theory of information. For example in Figure 2:

![Figure 2](image)

This property applies to geographical objects in large: a landscape, the elements of which are identical or similar to each other, transmits very little information to the navigator, for example. The sea is a case in point and so is a landscape of dunes where all hills are similar. On the other hand, when all objects are different from each other, information is high, but our cognitive ability to make use of this information is low because of Miller’s “magic number 7”. For example, one can easily get lost in a forest where all trees and local situations are different from each other and never repeat themselves. Venice is a case in point too.

A more general way to measure Shannonian information can be obtained by introducing an index \( j \) that distinguishes among the different objects under consideration. Thus, if all the objects are the same (e.g. as in Figure 2, row 1), the index \( j \) is the same, if they are different (Figure 2, row 2), the index \( j \) is different. In the case of geographical information, \( j \) refers to the way the geographical objects are being typologied: by height, architectural style, land uses and the like. Once index \( j \) is chosen, one may attribute to each index \( j \) the relative frequency of the occurrence of the corresponding objects \( (p_j) \):

\[
p_j = \frac{N_j}{N}
\]

where \( N \) is the total number of objects in the system of interest, and \( N_j \) the number of objects of the same kind. From here follows a somewhat different form of Shannonian information \( i \) that is given by

\[
I = \log_2 3 \approx 1.5.
\]
\[ i = - \sum_{j=1}^{N} p_j \log_2 p_j \]  \hspace{1cm} (3)

when the relative frequencies \( p_j \) are subject to the normalization condition
\[ \sum_{j=1}^{N} p_j = 1 \]  \hspace{1cm} (4)

Note that \( i \) as defined here (3) differs from \( I \) as defined above. While \( I \) (1) increases with the number \( N \) of objects, \( i \) (3) is the information per object, that is to say, \( i = I/N \) in the limit of large \( N \) (i.e. > 100). Thus, \( i \) is independent of the number of objects, but reflects their variety and character (The details concerning the transition from the definition of Shannonian information are given Appendix B of H&P paper).

Now, if all objects are equal (Figure 2, row 1), then there is only one kind and only one value of the index \( j \), namely \( j=1 \), and \( N_1 \) is equal to total number of objects \( N \). According to the definition (2), \( p_1 = 1 \), and \( i = 0 \). In other words, there is no information at all. On the other hand, if all objects are unique and differ from each other, then there are as many values for index \( j \) as there are objects, i.e. \( j \) runs from 1, 2... till \( N \). In this case \( N_j = 1 \),
\[ p_j = \frac{1}{N} \]  \hspace{1cm} (5)

and the information becomes
\[ i = \log_2 N \]  \hspace{1cm} (6)

In the case of Figure 2, row 2, \( i \) is about 1.5 bits, but for a city of thousands of buildings, each unique and different from the rest, \( i \) might be a rather large number. Note, however, that what counts here is not the size of the city per se, but the diversity of its objects (buildings, roads, parks, train stations, etc.) that determine the number of indices \( j \).

**SEMANTIC INFORMATION ENTERS IN DISGUISE**

One of the surprising outcomes of H&P study is the finding that, contrary to what was previously assumed, semantic and Shannonian notions of information are interrelated via the choice of the above indices \( j \) to which, implicitly or explicitly, a specific meaning and context must be attached. In other words, in order to determine the objective Shannonian value of specific objects, each must first be a member in a group. That is, one among many instances that were given the identity and meaning termed ‘a building’, ‘a street’, ‘a tree’, ‘a mountain’. Without that qualitative contextual or categorical property, the Shannonian quantity cannot be determined.

To further elaborate this point lets us return to what has been said above regarding equations 2-4, namely, that in a city of thousands of buildings, all unique and different from each other, \( i \) might be a rather large number. While such a city offers just one specific configuration of objects and therefore a perfect means for recognition and orientation, it requires an enormous amount of memory. In face of the above noted constraint on humans’ short-term memory by the “magic number 7”, this situation of a very high \( i \), might be a major disadvantage.
What one needs, therefore, is an optimal solution that will lie in between the extreme cases, all objects different and all equal. It is here where Lynch’s (1960) five elements (landmarks, paths, …) come in. Take for example the element District: When all buildings are identical (i.e. \( j = 1 \) and \( i = 0 \)), subdividing the city into districts increases the number of the indices \( j \) and by implication the information conveyed by the city. When all buildings are different (\( j \) is very large and \( i \) cannot be memorized), subdividing the city into districts entails information compression that reduces the number of indices \( j \) into a few, easily remembered, districts.

As can be seen, the choice of \( j \) determines the semantic information of the city – its qualitative structure – that in its turn defines the boundaries within which the Shannonian information of the city can be measured. The other elements of Lynch do the same job, but to see how we need to develop the concept of pragmatic information.

**PRAGMATIC INFORMATION**

Pragmatic information is an aspect of semantic information (This notion was first introduced to me by H. Haken in a personal communication). While semantic information refers to the meaning conveyed by a message or object, pragmatic information refers to the action it affords, in other words, to its action-related meaning. For example, the semantic information conveyed by a certain object in the environment might be that it is a rock, a tree, a river, a road, etc., while the pragmatic information afforded by these objects is that the rock is seatable (or not), the tree is climable (or not), the river is swimable and crossable (or not), the road is drivable, walkable, seatable (or not) and so on. Pragmatic information is intimately related to Gibson’s (1979) notion of affordance.

Now, take Lynch’s second element – path, or to be more specific, take roads in the city of Tel-Aviv (Figure 3). As can be seen, all roads afford the action of driving, some roads afford also moving in between focal points in the city, while only a few afford driving, moving between focal points and crossing the city from one side to the other. As a consequence, when we wish to categorize the road network of a city (Tel-Aviv in this case) by means of the index \( j \), we’ll have an index (category) \( j_1 \) that consist of a large number of relatively short roads, a much smaller number of \( j_2 \) roads, and a few \( j_3 \) roads. As in most cities, in Tel-Aviv each road is unique and differ from all other roads by its specific location, landscape and name. But due to the magic number 7, people will not be able to memorize all the short roads that were categorized as index \( j_1 \), but will have no problem memorizing each of the roads that were categorized and indexed as \( j_2 \) and \( j_3 \). This was recently confirmed by an empirical study conducted by Omer et al., (Forthcoming).

From the above follows that, compared to the roads in \( j_1 \), the roads of category \( j_2 \) afford less Shannonian information but a higher level of pragmatic information. There is a process of circular causality here: Drivers’ action in the city is influenced (determined) by the pragmatic information afforded by its road network – they drive more on the \( j_2 \) and \( j_3 \) roads. As a consequence they are exposed more to the specific structure and uniqueness of these roads and consequently remember them better, and so on in circular causality. Such roads thus become paths in Lynch’s sense. The same can be said on the rest of Lynch’s elements: landmarks, nodes and junctions.
To generalize: in a landscape (or any set of geographical objects) were all objects are different, the Shannonian information is very high, but the pragmatic information (the action afforded by this landscape or data) is low. A possible solution is to compress the information by means of categorization and the indices j. It is common in such a process that the number of objects in one (or few) of the categories j is small (about 7) to the extent that is affords high pragmatic information. Such object them become elements in Lynch’s sense. This is illustrated above in Figure 2 (rows 3, 4) and below in Figure 4 that is taken from a study currently under way.
CONCLUDING NOTES: INFORMATION AND SELF-ORGANIZATION

Semantic and pragmatic information are phenomena typical of open and complex systems that exhibit also properties of self-organization, non-linearity, chaos and the like. Such properties and phenomena imply that the potential states the system can take are infinite and cannot be determined a-priory. Shannon information, per contra, makes sense only in the context of closed systems where the number of the system’s possible outcomes is finite and a-priory known. As shown by Haken (1988/2000), the process of self-organization as described by his theory of synergetics is also a process of information compression: By means of self-organization a system with infinite possible outcomes is being compressed into a finite number of possible states. The notion self-organization refers to processes such as pattern recognition and categorization that in their turn give rise to semantic and pragmatic information and the effect of information compression. Some of these processes are innate to humans’ cognitive system, e.g. categorization of colors. Others are related to technology and the media that we use – in the spirit of McLuhan and Fiore (1967) in their book *The Medium is the Massage*. Still others are related to culture, economy, society and politics.

REFERENCES


