1.3 Representations and Information as a Function of Scale

Yaneer Bar-Yam
(Dated: March 10, 2016)

A representation is a map of a system onto mathematical variables. More correctly, a representation should be understood as a map of the set of possible states of a system onto the possible states of mathematical variables. A faithful representation must have the same number of states as the system it is representing. This enables the states of the representation to be mapped one to one to the states of the system. If a model has fewer states than the system, then it can't represent everything that is happening in the system. If a model has more states, then it is representing things that can't happen in the system. Conventional models often do not take this into account and this results in a mismatch of the system and the model; they are unfaithful representations and do not properly identify the behavior of the system, and thus ultimately its response to environmental forces or interventions we might consider. Because we are interested in influencing the system, we only want to know the distinctions that matter. We have to focus attention on those states that are *distinguishable* at a particular scale of observation.

To formalize these ideas for complex systems, it is useful to understand information as related to scale. We define the complexity profile as the amount of information necessary to represent a system as a function of scale. Information theory defines the amount of information in a message as the logarithm (base 2) of the number of possibilities of the message—the number of bits needed to represent the set of possible messages. Thus, the complexity profile is given by the number of possible states of the system at a particular scale. Typically, the finer the scale of inquiry about a system, the more information is needed to describe it (Fig. 1). The complexity at the finest scales is finite because of quantum uncertainty and is equal to a universal constant, $1/k_B \ln(2)$, times the entropy for a system in equilibrium, where k_B is Boltzmann's constant.

A single real number has infinite possibilities in all of the infinite digits of its representation. Therefore it has the ability to represent an infinite amount of information. This would seem to indicate that we could use a single real number to represent a system. For example, a number that represents the density of a liquid has infinite information, but we know from studies of phase transitions that this single number isn't enough. Why doesn't this work? The problem is that the way the information is organized in scale in the real number does not correspond to the way it does in the system. A real number can represent the position of a point along one dimension. Let's say we start by knowing where the object is to a resolution of 1 unit of length. Increasing the resolution by a factor of two means we can distinguish which of the two possible segments that are 1/2 units of length it is in. Communicating this information requires a binary variable. For each 2 fold increase in resolution we have 2 additional possibilities to specify. The number of bits is the logarithm (base 2) of the scale of resolution. However, for a liquid at its critical point the number of bits increases differently with increasing resolution. As resolution increases we have to describe the fluctuations of density. The growth in the number of bits is more rapid than one bit per factor of two in resolution (see
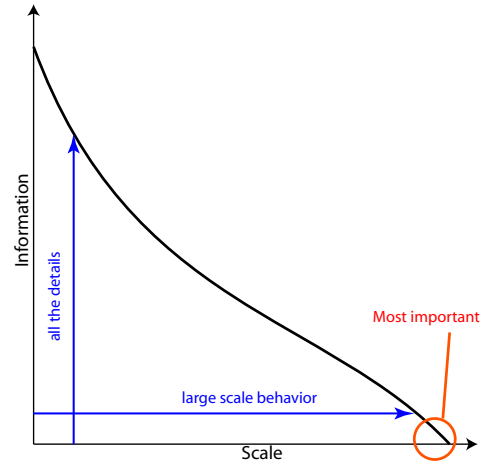

FIG. 1: The complexity profile is the amount of information that is required to describe a system as a function of the scale of description. Typically, larger scales require fewer details and therefore smaller amounts of information. The most important information about a system for informing action on that system is the behavior at the largest scale.

Fig. 2).

A sufficient representation, therefore, is one that has a set of possible states corresponding to the set of distinguishable states of the system at each level of resolution, down to the level we need to describe the properties we care about—the relevant parameters—and no further. When considering interventions that affect the large scale properties of the system, rather than accumulating details about the system, we should start with the largest scale pattern of behavior and add additional information only as needed. According to the complexity profile, each piece of information about a system has a size—the largest scale at which we can begin to detect that piece of information.

$$x = 0.d_1 d_2 d_3 d_4 d_5 d_6 ...$$

$$
\begin{aligned}
y = 0.d_1 d_2 d_3 &d_4 d_5 d_6 ... \\
&e_1 e_2 e_3 e_4 e_5 ... \\
&f_1 f_2 f_3 f_4 \cdots \\
&g_1 g_2 g_3 \cdots \\
&h_1 h_2 ...
\end{aligned}
$$

FIG. 2: A single real number (x, top) has infinitely many digits, which increase the amount of information available at a rate that is two possibilities for every change of scale by a factor of two. Real numbers are not good representations of systems for which the amount of information grows differently with scale (y, bottom). The number of digits as a function of scale is characterized by the complexity profile Fig. 1.