The Basic Problem with Language

Our goal is to capture in our minds the essence of molecular information and thereby capture the essence of the living universe. We intend to develop a language designed specifically to do this, but there is one gymungus problem that currently stops us dead in our tracks, call it the information relativity problem. Here it is in a nutshell:

The essence of information is counting. Before we can count we must have things to count. Now, define "things to count," or at least try to name the things you intend to count.

Noun = Thing
Thing = Set member

A set is merely a container. We can define any set any-'ol-way we want by merely making a list of requirements that must be met before anything can belong to a particular set and therefore take its name. We define the container (set) and then put things into it. By virtue of a thing's existence within a set, the thing takes on the properties of the container. In other words, the thing contains a certain amount of information simply by virtue of its membership in a particular set. Unfortunately, the same thing can be put into many different containers simultaneously.

Set A = Big, Mammal, Male, American.
Set B = Small, Human, Male, American.

I fit into both sets. I am a Human, Male, Mammal, 5'10", 170lbs, and I was born in Indianapolis, which makes me a proud American. By Mammal standards, I am Big. By American Male standards, I am Small. Obviously, I am now simultaneously Big and Small. So now what is the information value of knowing my relative size? It should be one bit because there are only two choices, Big and Small. However, we now have two bits of information and they are contradictory by virtue of me being in both of these sets. So, we have an obvious problem of defining and quantifying information in general.

I am Big is a true statement. I am Small is also a true statement. There should be one bit of information in knowing my size in any context, but it takes multiple bits of information to know the correct context for knowing my size bit. I must have many size bits just to handle the binary question of whether I am Big or Small, or to be able to utter a true statement about my relative size. However, it is even worse than it may first appear. Recall that "American" is also a relative term, and Americans have been around for centuries. Compared to all Americans in all of time, I am Big. This doesn't even address my age, which compared to those much younger or much older I am again Big. So now we must add more than one timeline to our ever-expanding system of naming.

The obvious solution is to store all of the information about me in a file, and then compare it to other things only as needed. It is a relational database kind of thing. We must first decide how precise the weight and height data must be, and that will determine the number of bits we should set aside. That literally is the "value" of "knowing" my

size, but is it really? Consider that I might be compared to a virus particle or the solar system, so we probably need many bits of storage. Let's say we set aside a kilobyte for storing my size data, which should be more than adequate. Next, we need lists to which I am to be compared. Last, we need algorithms to make the comparisons.

How much information is contained in my size?
The answer is: it still depends.

Where now does the actual information of my size actually exist?
The answer is: everywhere.

We could make eight billion comparisons of me to other things, and so we derive a gigabyte of relative size information. We could even rank my size within each set used to make the comparison, which would turn each size bit into many gigabytes of information. So, the information value of my size depends on my physical parameters, the storage of the physical parameters of all other things, the algorithms that take the measurements, store the data and – most importantly – it depends on the systems that make the comparisons. The "reality" of my size is far less important than the codes that give us the information about my size.

How can this concept be conveyed in simple words, for all things, every time?
I don't know, but I do know that it won't be easy.

Note: We have only talked about nouns, yet the same problem exists – worse, really - with verbs, adjectives, adverbs, etc. So, this is going to take a huge measure of out-of-the-box thinking.

This is the simplest way I can think of to define the problem at hand. It is a really nasty problem, and it only gets worse the deeper we go. This is fundamentally the problem for which we seek a new language – the problem of context. It is a problem of language – all language - and there are no simple solutions; unfortunately, this is the heart of the problem when we begin to describe the activities of molecules at the heart of life.