General Scientific Premises of Measuring Complex Phenomena

H.M. Hubey, Professor

Computer Science

Montclair State University, Upper Montclair, NJ 07043

ABSTRACT

General scientific and logical premises lurking behind the art of measuring complex phenomena, specifically intelligence, are explored via fuzzy logic, probability theory, differential equations, thermodynamics, generalized dimensional analysis, philosophy and psychology.

KEYWORDS: *generalized dimensional analysis, path functions, fuzzy operators, fuzzy logic, thermodynamics, extensive variables, intensive variables*

0. THERMODYNAMICS, OSs, AND TURING

Thermodynamics is probably the classical and ideal example of a *system-theoretic* point of view, and one that is built on the twin concepts of *state* and *process*. Furthermore, it is probably the only link from physics to the study of living things, which are most likely the most complex things which humans will ever have to study. The physical sciences are the easy sciences; it is the life sciences that are the hard sciences.[1] Unfortunately, physical scientists work with powerful tools, and life sciences have restricted themselves to working with much less powerful tools[1].

Thermodynamics is a perfect example of a science whose development lead to the improvement of the measurement of a fundamental dimension of physics. It was not until Lord Kelvin saw some inconsistencies that the concept of an 'absolute' temperature scale was created. In measurements of things such as length, mass, or time we can easily envision the concept of 'zero'. But it is not so with temperature. Nobody knew what the lowest obtainable temperature was. In the arguments in the philosophy of science there exist *data-first* and *theory-first* schools. Here we have a case in which both are iteratively used. The problem of intelligence is most likely to follow this pattern of development. If the problem is in an area that has a well-developed theory, we must try to explain the phenomenon in terms of the developed theory. It is only when we cannot that we can start thinking about a new theory, and this requires datamining techniques.

An Operating System (OS) is a very complex object. It has been said that "I may not know what an OS is but I can recognize one, when I see one!". The same thing may be said about intelligence, (or cognitive ability or any of the other related words such as awareness, consciousness, or autonomy, or even life.) The Artificial life newsgroup (alife) skipped trying to define life or artificial life. The only serious effort in this direction was made by Alan Turing. He essentially formalized the saying about the OS into intelligence. We may not know what 'intelligence' is but we know how to recognize one when we see one. Apparently when we talk about intelligence, we are talking about 'human kind' or 'human type' or 'human level' intelligence, or at least 'living thing' kind (type/level) of intelligence. We can say things about this without being able to define it precisely. It is precisely about this intelligence that Turing was referring to when he wrote about what is now referred to as the 'Turing Test'. He understood all the problems that involve discussions of this thing called intelligence many decades ago and offered his 'Gordian Knot' solution. Sometimes thinkers are unable to break through the boundaries of what has been created. Whitehead claims that Aristotle hindered the development of science for 2,000 years because nobody was courageous enough to break through the boundaries of the box for the sum total of all knowledge for human kind.

1. MEASUREMENT THEORY I

Normally, in the physical sciences, the possibility that an instrument may be capable of high precision while not being able of high accuracy does not occur to people. It can only occur if the instrument is broken. If the instrument is a very simple one (such as a ruler) we'd see immediately if there was something seriously (or obviously) wrong. If the instrument is a highly complex one, then there would be various self-tests. However, in the social/life sciences creation of 'instruments' is an art. It is quite possible for the instrument to be *reliable* (precise) but not *valid* (not accurate) or vice versa. For example, a psychologist might decide to create a questionnaire which he claims measures 'hostility'. The same person taking this test (the questionnaire) might obtain different scores at different times. So habituated are we to measuring things in this modern age that we scarcely give thought to the possibility that what is being represented as a number may be meaningless. That is the validity of the measurement i.e. that the measurement or metric actually measures what we intend to measure. In physical measurements there is usually no such problem. Validity also comes in different flavors such as construct-validity, criterion-related validity, and content-validity. Reliability refers to the consistency of measurements taken using the same method on the same subject. *(Please see Figure 1)*

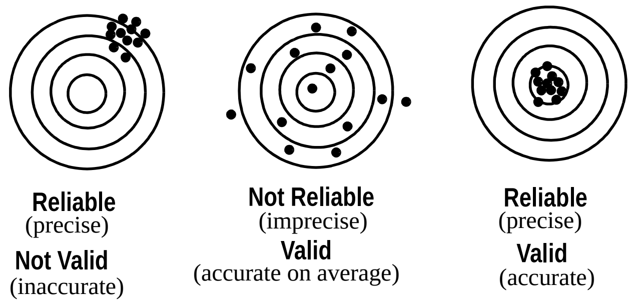

Figure 1: Reliability and Validity Analogy: One normally expects accuracy to increase with precision. However in the social sciences they are independent.

2. MEASUREMENT THEORY II

We often need to make things comparable to each other. We call this *normalization*. That is most easily done if we use numbers. For example, one way to normalize test grades is simply to divide every grade by the highest grade in class. This guarantees that the highest grade in class is 1.0. In order to be able to compare one boxing match to another a standard scoring system is used in which the same number of referees are used to score the bout, and for each round at least one boxer must be given 10 points. In Rasch measurements, we use

$$\frac{P}{1-P} = e^{\alpha - \delta} \qquad (1)$$

where P=Prob{answering correctly}, α =ability, and δ =difficulty of question. However, this is not scale-free. It would probably be better to use something like

$$\frac{P}{1-P} = \frac{\alpha}{\delta} \quad \text{or} \quad \frac{P}{1-P} = 1 + \ln\left(\frac{\alpha}{\delta}\right) \qquad (2)$$

In this case it is only necessary that both α and δ be measured on the same scale (somehow). Obviously, it would be best for all purposes to use numbers in the standard interval [0,1].

3. MEASUREMENT THEORY III

Before we try to normalize quantities we should know what kinds of measurements we have. They determine if we can multiply those numbers, add them, or can merely rank them etc. Accordingly measurements are classified as: (i) Ratio scale, (ii) Interval scale, (iii) Ordinal scale, or (iv) Nominal scale.

Absolute (Ratio) Scale**:** The highest level of measurement scale is that of ratio scale. A ratio scale requires an absolute or nonarbitrary zero, and on such a scale we can multiply (and divide) numbers knowing that the result is meaningful.

Interval Scale: The Fahrenheit and Celsius scales are interval scales. The differences on these scales are meaningful but ratios are not. That is what Kelvin found out, and that is what the absolute temperature scale is about. When measuring things such as intelligence, consciousness, awareness, or even autonomy, or hostility, we have no guarantee that we are measuring any of these on an absolute scale. There must be some other guidelines. One of the guidelines is obviously the study of various scales. In the intelligence game, psychologists have mainly relied on the *central limit theorem* in ‘hoping’ that intelligence is a result of many many different things adding up to create a Gaussian density. Thus they have contrived to make sure that test results are Gaussian.

Ordinal Scale: The next level on the measurement scale is the ordinal scale, a scale in which things can simply be ranked according to some numbers but the differences of these numbers are not valid. In the ordinal scale we can make judgements such as A>B. Therefore if A>B and B>C, then we can conclude that A>C. In the ordinal scale there is no information about the magnitude of the differences between elements. It is possible to obtain an ordinal scale from questionnaires. One of the most common, if not the most common is the multiple-choice test, called the *Likert scale*, which has the choices: extremely likely/agreeable, likely/agreeable, neutral, unlikely/disagreeable, and extremely/very unlikely/disagreeable.

Nominal Scale: The lowest level of measurement and the simplest in science is that of *classification* or *categorization*. In categorization we attempt to sort elements into categories with respect to a particular attribute. It ranks so low on the scale that it was added to the measurement scales later. Even an animal that can tell food from nonfood can be said to have learned or can be said to know about set operations instinctively.

The most basic and fundamental idea underlying these scales which is not even mentioned, and which is extremely important for measurement of complex phenomena in the life sceinces, is that in the final analysis, it is the human sensory organs that are the beginnings of all measurement. In the measurement of temperature, although a difference scale was easy to set up via the human sensory organs (and induction), it took theory and scientists to obtain an absolute scale for temperature. To obtain a difference scale the only thing necessary was for humans to note that the liquid in the glass went up when it was hotter. There was no way to know which was more hot and which less hot except via our naked senses.

This is/was as basic as knowing the difference between which of two sticks is longer than the other or which of two weights is the heavier one. Similarly in the measurement of intelligence, the final arbiter is still the naked human senses. Humans must make up the tests and decide which is more intelligent, say a chimpanzee or a dog. There can be no other way to proceed. The genius of Turing was that he realized this immediately. Therefore, Turing’s basic intuition is correct. We might not know what intelligence is but we can recognize it when we see it. Secondly, we should probably turn to nature to find examples and a hierarchy or scaling of intelligences. It would not be off the mark to accept that all living things are intelligent to a degree, and that EI (Encephalization Index) is basically a good scale on which to compare the intelligences of at least some living organisms.[2]

4. MEASUREMENT THEORY IV

Before we can even think about whether our measurements are on an absolute or difference scale we have to make sure that the objects that we deal with are *quantifiable* in some way and that we can measure them (with numbers naturally). Our handle on the problem is that the things we measure in physics (and hence engineering) come in *fundamental dimensions.* For example, dimensions of that particular branch of physics called mechanics consists of M {mass}, L {length}, and T {time}. For electrical phenomena we need one more dimension, Q (charge), and for thermal phenomena we need θ (temperature).

Then we can entertain the thought of using *dimensional analysis* for complex phenomena which is a method of reducing the number and complexity of experimental variables which affect a given physical phenomenon, using a sort of compacting technique. If a phenomenon depends upon n dimensional variables, dimensional analysis will reduce the problem to only k dimensionless variables, where the reduction $n - k$ =1,2,3 or 4 depending on the problem. Since these new dimensions are products/ratios of the old variables to various powers, the new dimensionless space has nonlinearly twisted and compacted the old problem in a way in which we can see regularity.

These ideas have been put to good use in biology [3]. For example, the mass of an animal grows proportional to L^3 but its surface area is only proportional to L^2. Thus, as animals get larger they have to have larger cross-sections of bones to support all that weight. So an elephant does not look just like a large sheep. These ideas have to be taken into account when prototypes, say, airplanes are tested in wind tunnels. Many other things having to do with scaling of living things such as metabolism, oxygen consumption, heat exhaustion, cooling etc. can be found in Schmidt-Nielsen[3]. For example, one way to make different animals's brains comparable is to compare not their brain capacities but the ratio of their brain mass, b, to their body mass B. Until recently, there was no method that could cluster the variables in similar ways as above so that nonlinear dimensional compaction was not available, but now there is a generalized data-driven method.[4]

5. PHILOSOPHY

Why do we do philosophy? One reason is because we do not want to 're-invent the wheel'. If philosophers have already thought about this topic, we should at least be aware that thought has been expended and results have been achieved.

Operationalism: The problem of what is being measured in quantum mechanics was solved during the early part of this century by 'operationalism' an idea (by Bridgeman) that the operations that are being executed define what is being measured. As long as everyone does the same thing, we are guaranteed that we all measure the same thing. In the measurement of something like intelligence, obviously, the problem of validity remains.

Quality vs Quantity: Thermodynamics, gave us the concept of *extensive* and *intensive* variables. It is often remarked in narratives that a fundamental difference exists which can be characterized by the words 'quantitative' vs. 'qualitative'. Often what is meant by the word qualitative is "intensive" since concepts often characterized as a quality can also be quantified. If a system consisting of a lot of 10,000 TVs is split into two sets at random, the quality of the two subsystems will equal each other and the quality of the TVs of the whole original system. A state of a system is characterized by a set of parameters. If we split a thermodynamic system (say a container of gas) in half some of the parameters will obey $X_1 + X_2 = X_s$ and others will obey $x_1 = x_2 = x_s$. The former (upper case) are *extensive* parameters, and the latter *intensive* parameters.

Open vs Closed: The concepts open vs closed (*endogeneous vs exogeneous*) are obviously very closely related to each other. In a closed system there can be no such thing as an exogeneous variable. At the same time, in general there is really no accurate or clear definition of what an open system is. In thermodynamics from where these ideas are probably borrowed, an open system is one which exchanges mass with its surroundings. A closed system may exchange heat, and do work on its surroundings, or have work done on it by its surroundings. Additionally, heat and work are processes. In other words, they are not *point functions*, but *path functions*.

In general in mathematical modeling via differential equations, the surroundings (*forcing* or *source* term) is everything that does not have the system variable in it and usually put on the rhs. However, when these concepts are specifically applied to intelligence, we have to clarify what it is that the system exchanges with its surroundings. The concept can apply to both exchanging data and or information with its surroundings. At the same time, the word "open" may be used to refer only to the problem at hand (i.e. if the problem is "open-ended"), but then it is not about generalized intelligence but about a specific problem. To generalize it we will then be forced to think about what little we know about how the brain does its work or how to generalize from the mathematical methodology that presently exists (i.e. logic, probability theory, etc). [1]

Many-as-One: The most fundamental such concept according to modern math is 'set' and forms the basis of logic, where philosophers are at home. This idea is the building block of all systems. A body is not just a parts list although it is comprised of many subsystems thus is not merely a set. We have many ways in mathematics of treating many things as one. A *tensor* is a general object of any degree. A zero dimensional tensor is a *scalar.* A one dimensional tensor is a *vector* or an *n-tuple.* A two dimensional tensor is called a *matrix*. In addition to this, from computer science we have the latest, and more flexible concept of hierarchical ordering via OOP (object-oriented programming) in which an object is a set of parameters without necessarily being merely a set or a vector.

Parallel vs Serial (sequential): This is one idea that occurs quite often. Some problems are parallelizable. For example, to dig a large ditch if we hire 100 workers as long as they do not interfere with each other, the ditch-digging will go at a rate 100 times as fast as before. However, if I want to send a message with a messenger, it does not matter if we use 100 messengers. The increase in the number of messengers might increase the *reliability* but will not affect the speed of the delivery. But parallelity also has to do with simultaneity (not always in time), choices, and substitutability, and logic.[7]

Trade-offs and Logic: We can sometimes trade-off something for something else in which case these things are substitutes of some kind. This idea shows up in logic as a logical-OR (co-norm). In the psychology and cognitive science literature, many different components of intelligence are posited. It is quite possible that some of these intelligences are composed of other more primitive types. If so, then are some of these substitutes for each other?

6. PSYCHOLOGY & COGNITIVE SCIENCE

Obviously throughout most of the century those who have worked on the nature and measurement of intelligence (almost always human intelligence) have been psychologists. They have had recourse to and benefited from methods and argumentation in both philosophy and physics. The kinds of questions with which they have toiled can be summarized in modern (and mathematical) terms as:

i) What kind of a quantity is intelligence? Is it *binary* or measurable on some scale? What kind of a scale is appropriate? Is it an *ordinal*, *interval*, or an *absolute* (ratio) scale?

ii) Is it an *additive function* of its constituents, the most important ones for purposes of simplification being hereditary (nature) and environmental (nurture)? Or is it a *multiplicative function*? Is it logarithmic function, an exponential function or a polynomial function of its variables?

iii) Is it a *vector/tensor* or a *scalar* (Spearman's g)? In other words, can a single number be produced from many numbers which is meaningful? Is there a hierarchy of intelligences, some of which subsume some of the others?

iv) Is it a *state* or a *process* ? In other words is it a *point function*, or a *path function*? Is it a *quality* or a *quantity*? In other words, is it an *extensive* variable or an *intensive* variable?

v) The *nature vs nurture* problem: Are the differences in intelligence among humans due mostly to heredity or environment?

There is a related (and incorrectly stated) version of (v) which is "Is intelligence mostly genetic?" The answer is quite plainly that intelligence is mostly genetic if intelligence is discussed in its most general form, that is including machine intelligence and animal intelligence. However the answer to (v) is much more complicated.[5]

An almost perfect example of a vector of cognitive science is color. We all know what colors are but they would be virtually impossible to explain to someone who was congenitally blind. If we did attempt to "explain" colors by explaining that "black is the absence of color and white is a mixture of all the colors" it is likely that the blind person would think of colors as what we call "gray scale". The analogical question is whether the components of intelligence that psychologists have posited are like colors in that they 'seem' as if they are 'unique' objects or is there a single number which we may obtain from the components.[8] Is this single number like colors or is it like the gray-scale?

7. COMPLEXITY AND HIERARCHY

The concept of layering or hierarchy is one of the most basic in the universe. Whereas hierarchy requires more detailed explanation the concept of layering is easier to envision and observed all over the world, at a very coarse-resolution. We use pictures of all sorts (as in Figure 2).

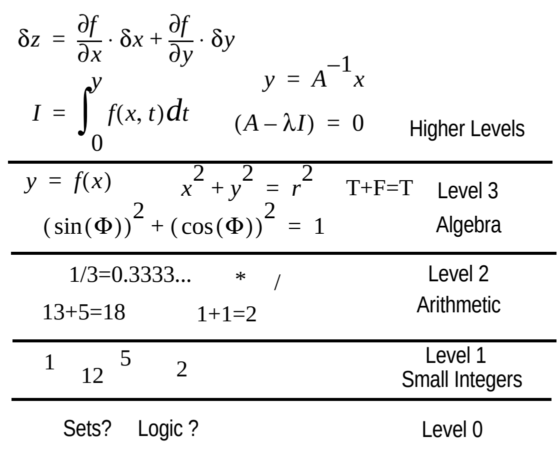

Figure 2: Highly-suggestive Layering in Mathematics: Knowledge is built-up in layers. New knowledge is built on top of old knowledge. This has significance for intelligence testing.

What better example than knowledge? *Data is raw. Information is data that is meaningful to an intelligent entity. Knowledge must be compressed information.* The only way to compress information is via exploiting regularities and patterns. Since mathematics is the study patterns, and regularities of all kinds, it is clearly the best tool with which to do science. Many more examples of layering can be found [1],[5],[6].

Thus the *scientificity* (intensity) of knowledge must be mathematics. Is it possible to measure intelligence separate and apart from knowledge? Do we want to weight some kinds of knowledge more heavily than others?

8. DISTANCE & MEASUREMENT

The main problem here is whether, after having gone through the problem of identifying the various components of intelligence, we should multiply them or add them to create a single number called intelligence. Therefore two prototypical choices for distance are

$$d(\vec{x},\vec{y}) = \left(\sum_{i=1}^{n}(\alpha_i x_i - \beta_i y_i)^{2m}\right)^{\frac{1}{2m}} \quad (3)$$

$$d(\vec{x},\vec{y}) = \left[\prod_{i=1}^{n} x_i^{\alpha_i} \bullet y_i^{\beta_i}\right]^{\frac{1}{\Omega}} \quad (4)$$

Obviously, in Eq (4) every component must be nonzero. There are good reasons why it is so. If normal functioning of a human depends on having absolutely no genetic defects, and if the intelligence of a human is determined by n genes, then if any of them is defective it should effect the score in the same way that the reliability of a composite is the product of the reliabilities of its components. In this sense, then the factors are analogous to probabilities.

This is also how we humans apparently tend to evaluate intelligence, as can be seen in the schizoid labeling of the condition known as *idiot-savant.* Being apparently superhuman in one aspect of intellectual activity is not sufficient to escape the label 'idiot'. It is said that *an expert knows everything about nothing whereas a generalist knows nothing about everything.* In an extension of this, then, today's experts (i.e. engineers) are idiot-savants. Their social IQ is said to be low. Programs like Maple, then, are also idiot-savants.

9. AVERAGE-IZATION

Consider the problem of being a juror in a beauty pageant. We will be forced to use a kind of scale in Eq. (5) (below)

$$B(\vec{x}) = 1 - \left(\prod_{j=1}^{n}\left[\{x_j - \mu_j\}^{\alpha_j}\right]\right)^{\frac{1}{\Omega}} \quad (5)$$

where the μ_i are the means. For example, the features/properties (of the vector x) may be nose length, skin color, lip thickness, fatness, etc. We will not want to vote for those with lips too thin or too thick, with noses that are too long, or too short, legs too thin or too thick, skin too pale or too dark. In other words, we are not looking for the minimum or the maximum but rather the most perfect average there is (with some caveats). This is a different kind of logic, triage logic [10].

Then, the human-kind of intelligence, if it is going to resemble what we humans normally think about perfection (apparently) should be measured via

$$I(\vec{x}) = 1 - \left[\prod_{i=1}^{n}\{x_i - \mu_i\}^{\alpha_i}\right]^{\frac{1}{\Omega}} \quad (6)$$

where the {x} are the various attributes of intelligence. The Turing test is probably for this kind of intelligence. For example, a machine that can solve differential equations and multiply 20 by 20 matrices in a jiffy (such as Maple, a Computer Algebra System) would flunk the Turing test. A human would know that a normal human (or maybe even an abnormal human) cannot do that. Therefore, the machine that could pass the Turing test would either have to be designed dumbed-down or it would have to learn to deceive. There are other things machines can do very quickly that humans cannot accomplish.

Thus the 'measure' above would show that such an entity could not be human (ceteris paribus, of course). In other words, as long as the machine is able to do the other things more or less as a human, then overachieving (outdoing humans) in one of the dimensions of the vector space would mark it as a machine.

Exactly the same would apply in some other capability such as being able to lift a few tons, swimming or running at superhuman speeds etc. For machines, then locomotion, would also be treated as part of intelligence. However, since even lower animals (less intelligent than us) can move around, it should not contribute much to the measurement of intelligence.

There are some pyschologists who want to include many human capabilities, such as physico-kinetic intelligence (i.e. physical ability) in the intelligence equation. Therefore, this 'autonomy' capability of animals/machines may also be considered to be a part of intelligence. We may take those that have been posited by psychologists as a starting point keeping in mind that some of them may really be substitutes for each other so that the measurement might be more complicated.

10. MORE SOPHISTICATION

Consider the simple problem of nutrition. Suppose we can create a balanced diet from the few foods available from three separate food groups; meat (protein), carbohydrates, and vegetables as shown below.

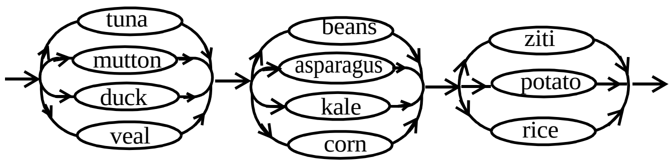

Figure 3: Parallel or Serial Choices. The problem is actually about multiplication vs addition. Diagrams such as this occur in electrical circuits, Boolean circuits [9], or choice making.

In terms of circuit analysis (which can be thought of in terms of Boolean algebra, [9] it is clear that the parallel lines are about choices (and thus lack of constraints) and therefore represent logical-OR (disjunction), whereas the seriality/sequentiality denotes a logical-AND (conjunction). Probably the first thing a statistician would do if faced with the problem of determining the relationship between food groups and a balanced diet would be to try correlation-regression analysis which would be nothing more than

$$N = \alpha_0 + \alpha_1 t + \alpha_2 m + \ldots + \alpha_n c \tag{7}$$

where t=tuna, m=mutton, c=corn etc. This is really the same kind of valuation of the problem as a weighted average. However, if we think logically then we should be considering a function of form;

$$N = MCV \tag{8}$$

since we need to ingest food from all the groups. Furthermore, since these food groups may be instantiated via specific examples, then using fuzzy logic, we should be regressing one of

$$N = (t + m + d + v)(b + a + k + c)(z + p + r) \tag{9a}$$

$$N = (t + m + d + v)^{\alpha}(b + a + k + c)^{\beta}(z + p + r)^{\gamma} \tag{9b}$$

Obviously, the latter form (Eq. 9) is not only correct but will result in many products (possibly to various powers). It is exactly this kind of products that dimensional analysis produces however it works only for problems with physical dimensions. However, there are methods that will produce similar equations for any problem if sufficient amount of data is available [4]. If intelligence-measurement is at least as complex as that of proper nutrition, then the simple weighted average kind of methods which are additive will not work. In other words the regression in Eq (7) is something like a combination of logical (or fuzzy) ORs and ANDs. A question that comes to mind is if there are fuzzy operators which are neither OR nor AND but something like both and exactly like neither. The special functions [11]

$$H_h(x, y) = \frac{1}{2} \cdot (x + y)^{h+1} \tag{10a}$$

$$M_m(x, y) = 2^m \cdot \left(\frac{(x-y)^2}{2[(x-y)^2]^{1/2}} \right)^{m+1} \tag{10b}$$

or others similar to these can be used in cases in which we are not sure if additive or multiplicative models should be used. One can show that [11]

$$Max(x, y) = H_o(x, y) + M_o(x, y) \tag{11a}$$

$$Min(x, y) = H_o(x, y) - M_o(x, y) \tag{11b}$$

Therefore the operator (fuzzy *t-co-norm*)

$$F(x, y) = H_o(x, y) + (2\xi - 1)M_o(x, y) \tag{11c}$$

is neither a norm (intersection) or conorm (union) but a fuzzy operator or a fuzzy norm since it is a norm for $\xi = 0$ and a conorm for $\xi = 1$. Some of the present day attributes of intelligence posited by psychologists probably are substitutes for each other and thus Eq (6) might distort the measurement. Therefore, something like Eq (9) where the additions are fuzzy unions and fuzzy intersections will probably give better results. The equations are readily and intuitively comprehensible in terms of theory of reliability based on probability. Fuzzification of the norm-conorm can be done for any fuzzy logic. For example, the simple product/sum logic given by

$$i(x, y) = xy \tag{12a}$$

$$u(x, y) = x + y - xy \tag{12b}$$

can be easily fuzzified via

$$F(x, y) = \rho xy + (1 - \rho)(x + y - xy) \tag{12c}$$

11. HUMAN INTELLIGENCE

The main problem today in human intelligence tests (and genetics) is calculating how much of intelligence is ‘inherited’ and how much of it is learned. There are several ways in which the model for this may be derived. One way would be to point out general conditions which the ‘intelligence function’ must satisfy. It should be multiplicative. It should display the increase of intelligence in time from the time of birth. It should converge on some limit on average for the people while being allowed to fluctuate about the average rate of increase and the limit of human intelligence. The equation

$$\frac{dx}{dt} = \lambda(\alpha - x) \qquad (13)$$

increases exponentially, and converges to a limit which is a good approximation. We need to know what the parameters mean, and this can be gleaned from the behavior of the solution. In Fig (4a) we see several trajectories. Some converge to above average intelligence, and some to less than average. Obviously the coefficient α determines this limit.

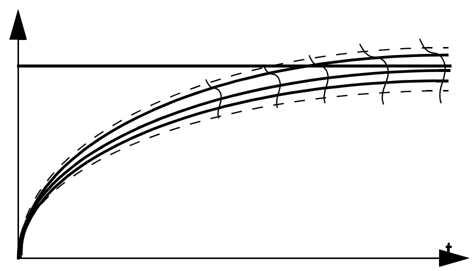

Figure 4a : Variations in α of the Intelligence Model .

In Fig (4b) we see a fluctuation in the rate of increase of intelligence, and this is controlled by the coefficient λ .

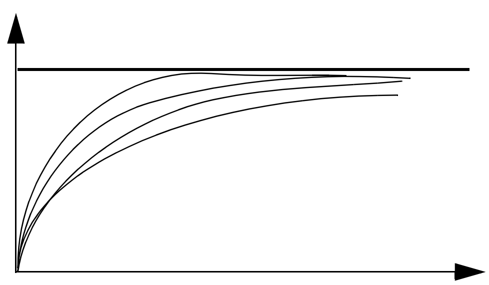

Figure 4b: Fluctuations in λ of the Intelligence Model

Logically both of these parameters then should be a function of both genetics and environment. Since we have determined that multiplicativity is important, the model should be

$$\frac{d}{dt}I(t) + \lambda G^{\eta} E^{\varepsilon}(t) I(t) = \lambda\alpha G^{h+\eta} E^{e+\varepsilon}(t) \qquad (14)$$

Integrating it once and rearranging terms we obtain the integral equation

$$I(t) = K(t) - \lambda G^{\eta}\int_0^t E^{\varepsilon}(\sigma) I(\sigma)\, d\sigma \qquad (15a)$$

$$\text{with } K(t) = \alpha\lambda G^{h+\eta}\int_0^t E^{e+\varepsilon}(s)\, ds \qquad (15b)$$

which is exactly what most researchers claim, that is, intelligence at time t, that is I(t), is a function of the past interaction of intelligence with environment summed up over time from time zero (birth) to the present time t. The interaction is multiplicative as it should be, and the equation is a reasonably good approximation over time of how living things (especially humans) learn. The solution is

$$I(t) = \Gamma e^{\lambda G^{\eta}\int_0^t E^{\varepsilon}(\tau) d\tau} \int_0^t E^{e+\varepsilon}(s)\, e^{\lambda G^{\eta}\int_s^0 E^{\varepsilon}(\tau) d\tau}\, ds \qquad (16)$$

where $\Gamma = \lambda\alpha G^{h+\eta}$ which in the limit goes to

$$I = \alpha E^{e} G^{h} \qquad (17)$$

If one day robots which learn from their environment are created, similar equations will be good first order approximations. Same probability techniques can be used on these equations, and statistics such as ‘heritability’ can be calculated. If the multiplication above is treated as some kind of a fuzzy intersection, then we can see quite clearly that the same kind of an equation can easily ‘explain’ the existence of natural language among living things. At the limits the equation must reduce the crisp logic, and we can see that it does. Only in the case when both genetic capability is there and when there is proper environmental stimulation, does language exist. If one or the other is missing there is no language. We can show how this equation explains what psychologists have said (in words) for a long time. Computing the virtual variation, we obtain for the special (and simpler case) of $e = h = \alpha = 1$

$$dI = E \cdot dH + H \cdot dE \qquad (18)$$

If the environment is enriched, the corresponding increase in intelligence depends on the genetic capability. Thus putting a dog in school cannot give it human level intelligence. Similarly, if there is a change in the genetic make-up (e.g. the difference between a chimp and a human) the change in the intelligence depends on the environment. A human brought up without human contact cannot walk or talk or dress up.

APPENDIX

Exact Differentials and Path Functions

The distinction between the related concepts *state* and *process* is an important one. There are mathematical definitions and consequences of these ideas. A *state* (or property) is a *point function*. The state of any system is the values of its *state vector* (a bundle of properties which characterizes a system). If we use these variables as coordinates then any state of the system is a point in this n-dimensional space of properties/characteristics. Conversely each state of the system can be represented by a single point on the diagram (of this space). For example for an ideal gas the state variables are temperature, pressure, volume, etc. Each color can be represented as a point in the 3-D space spanned by the R, G and B vectors. Intelligence is commonly accepted to be a state variable, i.e. a point. The scalar, Spearman's g, (single number, not a vector) can be obtained from this vector by using a distance metric. The argument that the values of the components cannot be obtained from the scalar, g, may be valid depending on the distance metric however, the distance metric may be devised in a way in which the components can be obtained from the scalar. Distance on a metric space is a function only of the end points i.e. between two states. However, the determination of some quantities requires more than the knowledge simply of the end states but requires a specification of a particular path between these points. These are called *path functions*. The commonest example of a path function is the length of a curve. Another example is the work done by an expanding gas. So is Q, the heat (transferred). In that sense work and heat are interactions between systems (i.e. processes), not characteristics of systems (i.e. state parameters/variables). Intuitively, when we talk about small changes or small quantities we use the differentials dx or δx. However the crucial difference is that although there may exist a function such that

$$\int_b^a dF = \int_b^a f(x)dx = F(x)\Big|_b^a = F(a) - F(b) \qquad \text{(A.1)}$$

there is no function Q, (heat) such that

$$\int_b^a \delta q = Q(x)\Big|_b^a = Q(a) - Q(b) \qquad \text{(A2)}$$

Instead we write

$$\int_a^b \delta q = Q_{ab} \qquad \text{(A3)}$$

meaning that Q_{ab} is the quantity of heat transferred during the process from point *a* to point *b*. Similarly because the infinitesimal length of a curve in the plane is given by

$$ds = \sqrt{dy^2 + dx^2} \qquad \text{(A4)}$$

we cannot integrate ds to obtain

$$S(b) - S(a) = \int_a^b \delta s = \int_a^b ds \qquad \text{(A5)}$$

but instead first the curve y=f(x) must be specified. Equivalently, if z is a function of two independent variables x and y, and this relationship is given by z=f(x,y) then z is a point function. The differential dz of a point function is an *exact differential* and given by

$$dz = \left(\frac{\partial z}{\partial x}\right)dx + \left(\frac{\partial z}{\partial y}\right)dy \qquad \text{(A6)}$$

Consequently if a differential of form $dz = Mdx + Ndy$ is given, it is an exact differential only if

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial z} \qquad \text{(A7)}$$

Therefore in the mathematical function used for the simple two-factor (nature-nurture) *Intelligence Function* the *environmental path taken* does make a difference in the final result which is assumed to be a state function (although computed from mental processes).

REFERENCES

1. Hubey, H.M. (1996) Topology of Thought, CC-AI: The Journal for the Integrated Study of Artificial Intelligence, Cognitive Science, and Applied Epistemology, vol 13, No.2-3, 225-292.
2. Eccles, J. (1989) Evolution of the Brain: Creation of the Self, Routledge, New York
3. Schmidt-Nielsen, K. (1984) Scaling: Why is Animal Size So Important?, University of Cambridge Press, Cambridge.
4. Hubey, H.M. (2000) A Complete Unified Method for Taming the Curse of Dimensionality in Datamining and Allowing Logical-ANDs in ANNs, submitted to Datamining and Knowledge Discovery.
5. Hubey, H.M. (1994) Mathematical and Computational Linguistics, Mir Domu Tvoemu, Moscow, Russia.
6. NIST (2000) White Paper
7. Hubey, H.M. (2001) Evolution of Intelligence:Direct Modeling of Temporal Effects of Environment on a Global Absolute Scale vs Statistics, accepted by Kybernetes: The International Journal of Systems and Cybernetics.
8. Hubey, H.M. (1996) Logic, physics, physiology, and topology of color, Behavioral and Brain Sciences, vol 20:2, pp.191-194.
9. Hubey, H.M. (1999) Mathematical Foundations of Linguistics, Lincom Europa, Muenchen, Germany.
10. Hubey, H.M. (2000b) Fuzzy Logic and Calculus of Beauty, Moderation, and Triage, The Proceedings of the 2000 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS2000), June 26-29, Las Vegas.
11. Hubey, H.M. (1999) The Diagonal Infinity, World Scientific, Singapore.