

A mathematical approach without assumptions!

Indeed, mathematics can be seen as a representation of logical relationships without any formal intentions and is, in a sense, a collection of mindless mathematical laws. As such it becomes a tool not actually related to anyone's scientific aims or intentions.

Part 1: The relevant issues underlying comprehension itself.

The real underlying issue is, can there exist anything which can not be described in terms of a goal-free mathematical representation? The relevant issue here is the fact that explanations can not be expressed without a language. One needs to take the trouble to analyze that specific problem prior to making any attempt to define a relationship between mathematics and general knowledge.

Any individual's knowledge of reality is entirely built on their personal perceptions (note that explanations by others are an important component of those perceptions). That the actual perceptions arise from interpretations of earlier experiences is an issue seldom considered by the scientific community. I hold that there are a number of subtly different concepts which need to be clearly understood in order to discuss the implied issues. First I would like to avoid the word "perception" as it can be seen as implying an actual "interpretation" of those experiences. It is my position that the word "experiences" provides a much more objective reference to such interactions. Clearly we must first identify what it is we "think" we perceive before we can build any mental explanations and/or representations of the supposed source.

Every human (including the most brilliant scientist who has ever lived) can be seen as beginning life as a child born without a language. During his life he will experience many interactions with what he supposes to be reality. It is the need to reference those experiences which stand behind the language he will eventually learn to use to express his understanding of those experiences. That includes his interpretation of the meanings attached to the elements of that language.

Of significance here is the fact that language is an arbitrary construct. A secret code can represent all the information required to communicate any collection of ideas without allowing translation so long as one has no information about the concepts being represented by the elements of that secret code. The definitions of the code elements must be learned. Clearly any language can be essentially seen as a secret code until the underlying definitions are understood. (Bit codes on computers are an excellent example of this issue.) Finally, comprehension itself is the very essence of learning.

Communication is actually the central purpose of comprehending any language. For that reason, I would like to define “understanding” to be recognition of the truth to be assigned to a specific thought represented by a specific collection of defined concepts: i.e., those concepts are the basic language elements.

Part 2: A universal mathematical representation of any language.

The collection of "concepts" expressible via any language can be listed by what is commonly referred to as a dictionary. This dictionary must be finite as otherwise it can not be constructed. Given such a construct, each and every entry could be given a specific numerical index usable to refer to that specific concept (that index set being a secret means of referring to any specific concept). Using that collection of numerical indices, any experience can then be specified via (x_1, x_2, \dots, x_n) where each x_i is the specific numerical index of a required concept. One may think in terms of the english concepts such as words, spaces, punctuation, etc.

It should also be understood that the meaning of any specific element may very well be altered by the other specified indices within the set (x_1, x_2, \dots, x_n) . That is the issue of “context”, a phenomena present in most all languages (note that the actual meaning of any language element often depends upon the context of the useage). The required dictionary is no more than another set of such entries (x_1, x_2, \dots, x_n) .

It should be clear to the reader that (x_1, x_2, \dots, x_n) thus becomes an abstract representation of a thought in the scientist's personal language. If you wish, you can see it as a secret code for those thoughts understood only by one who understands that "list" of concepts: i.e., understands the language. What is important here is that there can exist no thought conceivable by that scientist which cannot be expressed by the notation (x_1, x_2, \dots, x_n) . This rather trivial expression could be a single comment, a sentence, a book, an entire library or all the information on earth as "n", the number of referenced elements used, has not been specified and is thus an entirely open issue.

Part 3: Opening up a possible universal mathematical representation.

Given the above notation, the scientist's understanding of his experiences (essentially his explanation of any or all aspects of reality) can be represented by $P(x_1, x_2, \dots, x_n)$ where P stands for the probability he holds the specific represented thought to be true.

Note that the constraint imposed by "internal consistency" is a very simple issue under such a representation. Under representation given, the truth of the specified thought is a function of the explanation and cannot change except by changing either the actual "thought" which is being represented by the specified collection of indices, or by changing the "explanation" itself (this is essentially the definition of internal consistency).

That implies a very profound aspect of such a representation. Suppose that, given some specific index set x_i , one creates a second index set (including a new dictionary) where every specific index x'_i is exactly the original x_i plus a given constant "c". It then must then be absolutely true that

$$P(x'_1, x'_2, \dots, x'_n) \equiv P(x_1, x_2, \dots, x_n)$$

as each probability specifies the presumed truth of exactly the same thought.

The validity of that result would appear to lead to vanishing of another rather common mathematical expression.

$$\lim_{\Delta c \rightarrow 0} \frac{P(x'_1 + c + \Delta c, x'_2 + c + \Delta c, \dots, x'_n + c + \Delta c) - P(x'_1 + c, x'_2 + c, \dots, x'_n + c)}{\Delta c}$$

Anyone familiar with calculus will recognize that explicit expression as exactly the definition of the derivative of P with respect to c. Note that, in the specific case being examined here, the numerator must always exactly vanish while the denominator never vanishes. In classical calculus, the "limit" is the issue examined, not actual division by zero which classically is undefined.

Since "P" has been defined to be probability that the fact being represented is true, the above suggests P could be seen as an abstract mathematical function of the defined indices plotted as points on an x axis. If that were true the above would imply that

$$\frac{dP}{dc} = \sum_{i=1}^n \frac{\partial P(x_1, x_2, \dots, x_n)}{\partial x_i} \frac{dx_i}{dc} = 0.$$

As "c" is defined to be a simple shift in the origin of the representation, which requires the derivative of x_i with respect to c to be unity, the above would seem to imply

$$\sum_{i=1}^n \frac{\partial P(x_1, x_2, \dots, x_n)}{\partial x_i} \equiv 0.$$

That interpretation cannot possibly be true as there are three very serious mathematical problems with such an interpretation.

First, in a standard mathematical function, the number of arguments does not vary. In the approach presented here, the number of arguments “n” clearly varies from thought to thought.

Second, if one is to interpret the index as representing a point on an “x” axis, the order of the elements in (x_1, x_2, \dots, x_n) will certainly be lost and the order of elements is a significant issue in all languages of which I am aware.

Third, the subscript on x indicates the specific element index from the dictionary to be used. It should be clear that any specific index could be used more than once in a given expression. In such a case, the element of interest would plot to exactly the same point: i.e. the existence of such repetitions would be totally lost.

Part 4: Actual conversion into commonly understood valid mathematical notation.

Considering the third problem, the loss of information is easily fixed by adding another coordinate to the representation. For reasons which will become obvious later, I will call that axis the “tau” axis. The existence of that axis allows any repeated elements to be plotted to different τ positions. Note that the introduction of τ has added ignorance to the representation. Including this τ axis requires a vector notation $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$.

Addition of ignorance can also solve the first problem specified above. All one need do is to find the specific representation $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ with the largest number of entries and add “unknown” entries to every known experience sufficient to yield a specific value to n sufficient to cover all experiences. To evaluate the represented probability, the ignorance introduced is handled by integrating the mathematical representation over all possibilities for these added arguments. This adds the net impact on the result.

That leaves the second problem, the order of the elements. That problem can be solved by adding another hypothetical axis orthogonal to both x and τ . I will call that axis t because it indeed corresponds to what is commonly called time. All languages I am aware of have a temporal order given to their elements. Each and every expression $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n)$ can now be replaced by a “collection” of expressions of the form $(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$ within which order of the actual elements is of no significance.

Note that once again ignorance has been added. In evaluating $P(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$, only the elements specified by a specific value of “t” are actually defined. The others are to be seen as added “unknown elements”: i.e., their possible values are essentially defined to be unknown and one must integrate over the possibilities.

Under the elements and extensions as defined, we have what could be a mathematical expression required to be valid for all internally consistent explanations of any phenomena. It should be quite obvious that the representation is beginning to resemble the common physics representation of a collection of points moving in time. However, note that t is not defined as a continuous variable. Essentially t specifies the element existed when it was added.

At this point, the indices used to indicate specific concepts have become two dimensional vector entities. A resemblance to modern physics can be increased by adding two more orthogonal axes. It should be clear to the reader that creating a three dimensional vector representation of the original dictionary elements adds no real complexity to the opening proposition. Any thought specified in the original definition of x_i can simply be specified via (x_i, y_i, z_i) where \vec{x}_i is the specific vector index of a required concept. Every step discussed above goes through exactly as before and \vec{x}_i now becomes a four dimensional vector consisting of the components (x_i, y_i, z_i, τ_i) .

Uncertainty, (values for the undefined elements) has become the single most prevalent feature of this representation. However, the scientific perception of actual values in such a functional representation of the whole universe is chock full of such uncertainty.

Part 5: Clarifying the resemblance to valid mathematics to an exact match.

Anyone familiar with modern physics will tend to see $P(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$ as essentially equivalent to the the expression for the probability of a specific distribution of points which is changing in time. That idea presumes each and every element exists between the specified known times. In essence this is no more than an extension of the uncertainty introduced by the creation of elements tau t and fixed n.

The definition of a probability requires that P be positive definite and furthermore, that the integral over all possibilities must be unity. This suggests that P should be set equal to $\Psi^\dagger \Psi$ where $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$ is a complex function and Ψ^\dagger is the complex conjugate of that function.

If the integral of $\Psi^\dagger \Psi$ over all arguments is finite, one may merely divide Ψ by the square root of that number and $P = \Psi^\dagger \Psi$ will then be bounded by zero and one for absolutely all valid functions $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$.

The simplicity of that result is actually somewhat surprising. In modern physics, they begin with mathematical equations (based on their understanding of reality) which must be solved. For assorted reasons, they end up defining the solution to be a complex function Ψ . The product $\Psi^\dagger \Psi$ is then “interpreted” to be a probability and much is made of the problems associated with the issue above (referred to as normalization).

In the attack given here, the issue is attacked from exactly the opposite direction. P was defined to be a probability and, if “normalization” is not possible, the function $\Psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n, t)$ can not be a valid solution. A somewhat different but much more satisfying resolution of the underlying issue.

Setting $P = \Psi^\dagger \Psi$, the original algebraic constraint that the sum of the differentials of P with respect to all \vec{x}_i must vanish requires

$$\sum_{i=1}^n \vec{\nabla}_i P = \sum_{i=1}^n \vec{\nabla}_i \Psi^\dagger \Psi = \left\{ \sum_{i=1}^n \vec{\nabla}_i \Psi^\dagger \right\} \Psi + \Psi^\dagger \left\{ \sum_{i=1}^n \vec{\nabla}_i \Psi \right\} = 0.$$

Direct substitution will confirm that $\vec{\nabla}_i \Psi = i\vec{k}_i \Psi$ together with its complex conjugate constitutes a solution to the above equation.

This representation brings up a rather interesting issue the reader should be aware of. In the physics community, the letter “i” is commonly used with two rather different meanings (either the square root of -1) or the commonest index in a sum.

Both P and Ψ are no more than functions defined over a four dimensional Euclidian space. This implies that, if one has found a solution Ψ_0 where the sum over k_i vanishes, one can use a Fourier transform to convert Ψ_0 into a solution Ψ where that self same sum can have any value desired. Analytically, this is essentially transforming the solution to a different frame of reference. This mechanism is generally used in physics to transform a solution to what is referred to as a specified rest frame.

In addition, “t” can be seen as no more than another argument of P logically equivalent to any component of \vec{x}_i , thus one can also assert that direct substitution also confirms the solution

$$\frac{\partial}{\partial t}\Psi = 0.$$

At this point, of the language being used to represent our experiences has begun to strongly resemble our current mathematical representations of a quantum mechanical reality.

If we give an alternate name to what we have defined to be “elements” and call them “particles”, the identity becomes close to complete. This should not be thought of as a serious alteration of the original hypothesis. Clearly this proposed view of “supposed reality” must include particles of graphite on a sheet of paper, particles of chalk on a blackboard or even ink patterns within a published book. These and all the other existing elements together must constitute our understanding of reality implying that the mindless mathematical representations are themselves a significant language element.

There exists one other very important issue I have failed to bring up. At this point we have added so many elements to the mathematical representation that actually constructing a proper tau assignment such that no two points are identical becomes essentially impossible.

To solve that problem, I will add to the above constraints one additional mathematical constraint: The function $\delta(x)$ (called a delta function in modern physics) is defined to be zero for all $x \neq 0$ and infinite for $x = 0$. Furthermore, the integral of $\delta(x)$ is defined to be “unity” if the integration limits cover the origin.

It follows that the delta function provides exactly the required mathematical constraint on any valid Ψ . In essence, Ψ must vanish if the delta function does not vanish, and

$$\sum_{i \neq j} \delta(\vec{x}_i - \vec{x}_j) \Psi \equiv \sum_{i \neq j} \delta(y_i - y_j) \delta(z_i - z_j) \delta(x_i - x_j) \delta(\tau_i - \tau_j) \Psi = 0.$$

This constraint guarantees that Ψ must exactly vanish if any two elements exist at exactly the same point. This constraint exactly handles the issue for which the tau axis was introduced. It is interesting to note that this issue never even occurs in the common view of reality as a physical collection of entities but must nevertheless be true here. It has an number of serious consequences a few of which will be brought up later.

At this point, the final mathematical constraint on Ψ can be explicitly written. First I will introduce some anti-commuting operators (elements which change sign when commuted) plus some related matrix elements. The resultant mathematics is as follows,

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

and, $[\alpha_{ix}, \alpha_{jx}] = \alpha_{ix}\alpha_{jx} + \alpha_{jx}\alpha_{ix} = \delta_{ij}$, $[\alpha_{iy}, \alpha_{jy}] = [\alpha_{iz}, \alpha_{jz}] = [\alpha_{i\tau}, \alpha_j] = \delta_{ij}$

$$[\beta_{ij}, \beta_{kl}] = \delta_{ik}\delta_{jl} \quad \text{and} \quad [\alpha_{ix}, \beta_{kl}] = [\alpha_{iy}, \beta_{kl}] = [\alpha_{iz}, \beta_{kl}] = [\alpha_{i\tau}, \beta_{kl}] = 0.$$

One can then assert a properly Fourier transformed Ψ exists which must obey

$$\left\{ \sum_i \vec{\alpha}_i \cdot \vec{\nabla}_i + \sum_{i \neq j} \beta_{ij} \delta(\vec{x}_i - \vec{x}_j) \right\} \Psi = \frac{\partial}{\partial t} \Psi = 0$$

as it amounts to little more than a Fourier transformed assertion that $0+0=0$.

Part 6: The differences between this result and modern physics.

Clearly the mental picture here is quite different from the standard mental model presented in modern physics. Probably the single most disturbing factor is the existence of that tau axis. Under the picture presented here, the value assigned to τ_i can not be known. That brings up an interesting thought.

Heisenberg's uncertainty principle (an issue central to modern quantum mechanics) asserts that exact measurement of both position and momentum for any physical particle can not be made. That clearly implies that, if the position in the tau direction can not be known at all, the momentum in the tau direction must be quantized.

That idea suggests that every particle has a quantized value associated with its momentum. This brings to mind the relativistic expression of energy.

$$E = mc^2 = \sqrt{p^2c^2 + m_0^2c^4}$$

The zero subscript on m denotes that it is the rest mass, m being the relativistic mass.

That brings up another possible interpretation of this tau axis. There still exists one aspect of the tau dimension that is not yet defined or specified. That would be the velocity of a particle in that direction. If it is presumed that all particles move at exactly the same speed, the velocity in the tau direction becomes a defined variable and time ends up defining the physical position of events at any contact interactions.

That perspective yields three rather stunning results. Any particle proceeding in a direction orthogonal to tau must have a zero momentum in the tau direction. That implies that the quantized value of its momentum must vanish zero. Secondly any particle at rest in (x, y, z) space must be traveling at that specified speed in the tau direction. Its momentum must have a non zero quantized component. Clearly this quantized momentum plays the role of rest mass in the modern physics model.

Finally, the idea of a fixed velocity in the (x, y, z, τ) space implies that interactions between particles occur at specific times: i.e., if interactions occur between a number of elements the interactions define the same time for those specific interactions.

This brings us right back to the original definition of time: “in order for two objects to interact, they must exist at the same place and time”. In modern physics, time is defined by readings on clocks not physical interactions, quite a different issue. It is not particularly difficult to show that, interactions in (x, y, z, τ) require cyclic phenomena within a defined sub structure to actually measure changes in tau.

In essence, it is not at all difficult to show that special relativistic effects are embedded in the mathematical model presented here. General relativity is more difficult to demonstrate though it can also be shown that the relationships are embedded in the model I have presented (since the equations essentially embody the entire universe, defining an accelerating frame of reference is not a trivial issue).

Part 7: The general purpose of this essay.

If there were more room available, the above can be shown to yield all of modern physics. If anyone reading this presentation can find a single constraint I have placed on the information to be explained or an error in the logic of my presentation, I would very much appreciate being informed.

Richard D. Stafford, Ph.D. in Theoretical Physics, Vanderbilt University, 1971