Communications in Information and Systems
Volume 14, Number 3, 135–159, 2014

Split-complex numbers and Dirac bra-kets

Steven Deckelman and Barry Robson

We describe the real tessarines or "split-complex numbers" and describe a novel instance where they arise in biomedical informatics. We use the split-complex numbers to give a mathematical definition of a *Hyperbolic Dirac Network* (HDN) — a hyperbolic analogue of a classical probabilistic graphical model that uses quantum physics as an underlying heuristic.

The methods of theoretical physics should be applicable to all those branches of thought in which the essential features are expressible in numbers.
— *P. A. M. Dirac, Nobel Prize Banquet Speech 1933*

1. Introduction

A student's first foray into complex numbers often begin with the introduction of i as the square root of -1. i is a *number* having the property that $i^2 = -1$. The complex number system $\mathbb{C}$ is then defined to consist of all *numbers* of the form $a + bi$ with a, b real. With addition defined by addition of the real and imaginary parts, students learn that complex numbers are multiplied by applying the usual distributive laws and by replacing each occurrence of i^2 by -1, and that $\mathbb{C}$ is a field. More generally $(a + bi)(c + di) = (ac - bd) + i(ad + bc)$. Suppose we were to change this in the following way. In place of i we introduce a new quantity h. Consider now two "h-complex numbers" $a + bh$, $c + dh$. Then, in multiplying out (a+bh)(c+dh), rather than replacing each occurrence of i^2 by -1, we adhere to the following rule: replace each occurrence of h^2 by 1[1]. The ensuing *product* will yield a complex product that is different from ordinary complex

[1] In physics this is known as the Lorentz rotation $i \to h$. This rotation is in may respects the same in its consequences as the Wick rotation of time t to imaginary time $-it$, widely used to render quantum mechanics classical. It makes simple i-complex functions real. For example, Feynman early used it as an interim representation to facilitate solution to the path integral. However, it captures only some of the picture given by h, which can therefore be considered as a generalization of the Wick notation. Notably, exponentials of functions multiplied by h are particular

multiplication. h is acting like it thinks it's a hidden square root of 1. The resulting mathematical system is no longer a field but rather a commutative ring. h is a kind of hyperbolic analogue of i. For example

$$e^{h\theta} = \cosh\theta + h\sinh\theta$$

for θ real in a sense that will be made precise below. Neither is this idea new. In 1848 James Cockle [2], an associate of Cayley, dubbed this system of numbers the *(real) tessarines*. The tessarines are a kind of poor, less luminary cousin of the complex numbers. Not quite a field, the tessarines lack the power and versatility of $\mathbb{C}$. Yet the tessarines or split-complex numbers have a peculiar history of reappearing at critical times in the history of mathematics. In this article we describe a novel instance where these numbers arise in biomedical informatics.

2. Tessarines or split-complex numbers

What Cockle called the tessarines is actually the complex algebra generated by objects $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\ \}$ where $\mathbf{ij} = \mathbf{ji} = \mathbf{k}$, $\mathbf{i}^2 = -1$, $\mathbf{j^2} = 1$. That is, "numbers" of the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, or given by the matrix representation

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mapsto \begin{pmatrix} a+bi & c+di \\ c+di & a+bi \end{pmatrix}$$

where a, b, c, d are real numbers. When both $b = 0$ and $d = 0$, the resulting real tessarines were called *split-complex numbers* by later authors. The complex plane of split-complex numbers has a certain hyperbolic structure and is sometimes called the hyperbolic plane [24]. In order to emphasize the "hyperbolic" character of these split-complex numbers, we shall use the symbol h in place of $\mathbf{j}$ in what follows. The term h-complex algebra has also been used [14]. Split-complex numbers are a type of *generalized complex number* or *hypercomplex number*. In [25] I. M. Yaglom termed them *double numbers*. Even today, their terminology has not stabilized. In [15], Robson cites a litany of names that have been applied to these numbers including Cockle, Lorentz, dual, perplex, algebraic motor and anormal-complex amongst others. The quaternions are another, possibly more familiar, hypercomplex number. See also [11].

---

hyperbolic geometric functions, as multiplication by i would give periodic trigonometric functions (and multiplication by hi gives a combination of both), whereas having rendered the functions real by the Wick rotation, this subtlety is lost and would not be seen.

3. A modern algebra view of split-complex numbers

An algebra is a vector space in which, quoting Halmos [7], there is a decent notion of multiplication. A ring, if you will, in which we can scalar multiply elements using scalars in some field. One way they arise is to take a field $\mathbb{F}$ and let it act on a ring $\mathcal{R}$. The resulting $\mathbb{F}$-algebra will be a vector space over the field $\mathbb{F}$ in which vector multiplication is defined by multiplication in $\mathcal{R}$.

The complex numbers can be identified in the obvious way with points (vectors) in the plane $\mathbb{R}^2$. They form a real algebra using ordinary vector addition and the product $(a,b)\cdot(c,d) = (ac-bd, ad+bc)$ which extends in a natural way to a complex algebra. We usually don't refer to $\mathbb{C}$ as a complex algebra because it has the richer structure of being a field. It is a natural question to ask what sort of mathematical system ensues if we instead consider the ostensibly simpler product $(a,b)\circledast(c,d) = (ac, bd)$. What we get is no longer a field but can be viewed as a real or complex algebra. In fact it is the split-complex numbers in disguise.

Suppose we take the real numbers $\mathbb{R}$. We may view $\mathbb{R}$ as a field, but also as a vector space over $\mathbb{R}$ itself or even as an algebra. As a field, it is, a fortiori, a commutative ring. Algebraists recognize the mathematical system defined above by $(a,b)+(c,d) = (a+b, c+d)$, $(a,b)\circledast(c,d) = (ac,bd)$ as the direct product of rings $\mathbb{R}\times\mathbb{R}$. Notice that the direct product is not an integral domain. It's multiplicative identity is $(1,1)$. So elements $(a,0)$ or $(0,b)$ are not units. Thus it is not a field. It is however a commutative ring, in fact an algebra over $\mathbb{R}$.

Suppose now we take the complex numbers $\mathbb{C}$. We may view $\mathbb{C}$ as a field. But we may also view it as a vector space over the real numbers. It is the two dimensional real vector space spanned by the basis $\{1,i\}$. Suppose we replace this usual basis by the basis $\{\iota,\iota^*\}$ where $\iota = \frac{1+i}{2}$ and $\iota^* = \frac{1-i}{2}$, its complex conjugate. That is, rather than represent complex numbers in the form $a+bi$, we will represent them in the form $A\iota + B\iota^*$, with A, B real numbers. Now the complex number system, described in this way is a real vector space and can be given the structure of a real algebra by defining a product on the basis vectors ι,ι^* and then extending it to all of $\mathbb{C}$. Motivated by the h as a hidden root of 1 idea above, we define, using the symbol h in place of i,

$$\iota\circledast\iota = \iota, \quad \iota^*\circledast\iota^* = \iota^*, \quad \iota\circledast\iota^* = 0$$

and denote the resulting real algebra by $\mathbb{C}_\iota$. We note that this implies that

$$\iota = \frac{1+h}{2} \quad \text{and} \quad \iota^* = \frac{1-h}{2}$$

where $h^2 = 1$.

$\mathbb{R} \times \mathbb{R}$ and $\mathbb{C}_\iota$ are isomorphic real algebras. The isomorphism is of course $\Phi(a,b) = a\iota + b\iota^*$. It may also be described as the quotient ring

$$\mathbb{R}[h]/\langle h^2 - 1\rangle.$$

In this case the isomorphism $\Phi : \mathbb{R}[h]/\langle h^2 - 1\rangle \to \mathbb{C}_\iota$ may be described by

$$\Phi(x + yh) = (x+y)\iota + (x-y)\iota^*$$

or by $\Phi^{-1} : \mathbb{C}_\iota \to \Phi : \mathbb{R}[h]/\langle h^2 - 1\rangle$ where

$$\Phi^{-1} : (A\iota + B\iota^*) = \left(\frac{A+B}{2}\right) + \left(\frac{A-B}{2}\right) h.$$

$\mathbb{C}_\iota$ may of course be extended to a complex algebra by simply enlarging the field. That is,

$$\mathbb{C}[h]/\langle h^2 - 1\rangle.$$

4. The hyperbolic character of split-complex numbers

In addition to the algebraic description of the split-complex numbers, there are also geometric and function theoretic ways of thinking about them. They are in some sense *hyperbolic* analogues of ordinary complex numbers. For example, consider $\mathbb{C}$ equipped with the usual addition and $\circledast$ multiplication. For $z = x + iy$, $z \circledast \overline{z} = x^2 - y^2$. The set of "unit vectors" in this plane will be the points on the unit hyperbola $x^2 - y^2 = 1$. Students of of physics will recognize the split-complex number "norm" as the non-Euclidean Minkowski norm of special relativity.

Suppose now that $\Omega \subset \mathbb{C}$ is a domain (open connected set) and consider the class of analytic (or holomorphic) functions on $f : \Omega \to \mathbb{C}$, $\mathcal{O}(\Omega)$. Let H (for hyperbolic) denote the operator

$$H(f(\sigma + it)) = f(\sigma + ht).$$

$H : \mathcal{O}(\Omega) \to \mathcal{H}(\Omega)$ will be a linear operator where $\mathcal{H}(\Omega)$ is the class of smooth (e.g. C^2) functions $f = u + hv : \Omega \to \mathbb{C}$ satisfying the "hyperbolic Cauchy-Riemann equations"

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x}.$$

That is, H maps analytic functions of a complex variable in the classical plane into "hyperbolically analytic" functions in the hyperbolic plane. We may define a complex function to be hyperbolically analytic at z in the hyperbolic plane if

$$\lim_{\eta \to 0} \frac{f(z+\eta) - f(z)}{\eta} \quad \text{exists}$$

where we define the limit only for units η which approach 0 in such as way as to avoid points on the "null cone" $\{z : z = c\iota \text{ or } z = c\iota^*\}$ (for example by approach within sectors) and where the operations are understood to be in the hyperbolic plane, e.g. the η in the denominator signifies η^{-1} in the split-complex number system. See [8] for more details. As an example consider $f(z) = z^2$ defined by $f(x+iy) = x^2 - y^2 + 2xyi$. Then $Hf(x+ih) = x^2 + y^2 + 2xyh$ which is different. For example $f(i) = -1$ but $Hf(i) = 1$. We justify the use of the term hyperbolic since for $f(z) = e^z$, we see

$$H(e^{x+iy}) = e^{x+hy} = e^x e^{hy} = e^x(\cosh y + h \sinh y).$$

In particular,

$$H(\cos x) = \cosh x \quad \text{and} \quad H(\sin x) = \sinh x, \qquad \text{for real } x.$$

The AMS subject classification category 30Gxx, "Generalized Function Theory", comprises this area of mathematical research.

5. Split-complex numbers are not natural

Perhaps one reason split-complex numbers are not better known is that lacking the structure of a field, they are not a particularly natural mathematical system for applications. In particular the split-complex numbers do not form a division ring. Although a commutative ring, real scalar multiples of the quantities $1 + h$ and $1 - h$ fail to be units(e.g. have multiplicative inverses). Of course this does not detract from their intrinsic interest as a purely mathematical object. But it can be a delightful surprise when the unexpected utility of an idea in pure mathematics arises in applied mathematics. For remainder of this paper we work exclusively in the split-complex number system and so will drop the $\circledast$. For example we will write $(a\iota + b\iota^*) \circledast (c\iota + d\iota^*)(= ac\iota + bd\iota^*)$ simply as $(a\iota + b\iota^*)(c\iota + d\iota^*)$.

6. Hyperbolic Dirac nets

Physical analogy and intuition has a long and distinguished tradition as a source of inspiration and deep mathematical insights. Take, for example, Jean Bernoulli's ingenious solution to the brachistochrone problem, based on the path light takes through an inhomogeneous stratified medium as described in [12] or the original solution of the Dirichlet problem based on physical reasoning for a physical electrical potential being determined by the laws of electrostatics given a charge distribution on the boundary. Or recall P.A.M. Dirac's "function" from the early days of quantum mechanics $\delta(x)$,[13], which is supposed to be zero for $x \neq 0$ and ∞ at $x = 0$ together with the property that $\int_{-\infty}^{\infty} \delta(x)dx = 1$ which later gave rise to the theory of distributions or generalized functions. Or consider Wiener's development of stochastic processes based on trying to model physical Brownian motion. We could, of course, go on and on. Recently Barry Robson [14] has proposed using quantum mechanics as a basis for heuristics with the design and implementation of inference nets. The resulting net, called a Hyperbolic Dirac net, is based on split-complex numbers. Inference nets, a topic in artificial intelligence, are very important in bioinformatics, data mining and biomedical analytics as well as having many other applications. An example from biomedical informatics would be a patient record database. The science of designing and implementing such nets in a computationally tractable way is a nontrivial problem in computer science. In Section 10 below we give a more detailed example.

For our purposes an *inference net* may be thought of as a probabilistic graphical model. That is a graph $\mathcal{G} = (\chi, \mathcal{E})$ with nodes $\chi = \{X_1, X_2, \ldots, X_n\}$ and edges $\mathcal{E}$. $\mathcal{G}$ may be directed or undirected or partially directed. The presence of an edge between two nodes indicates a direct probabilistic interaction. Each node represents a (discrete) random variable and it is assumed there corresponds a joint probability distribution $P(X_1, X_2, \ldots, X_n)$ and underlying sample space. A *sine qua non* for such a net is that the joint be encoded in a computationally tractable way in a set of *conditional probability tables*. For more details we refer the readers to the excellent sources [9], [23], [10]. In a biomedical context, we may think of the random variables as representing states, events, or measurements that are subject to clinical and biomedical rules. We want to be able to build the net, modify the net and perform queries on the net. The canonical example of a directed inference net is called a *Bayes net* while the canonical example of an undirected net is called a *Markov net.* For example we might be interested in the probability that a patient has sleep apnea given the symptoms of snoring and fatigue. In principle, every probabilistic query can be answered if the joint $P(X_1, X_2, \dots, X_n)$ is available. The process of evaluating queries involves computations with sums of real number products of probabilities involving the joint $P(X_1, X_2, \dots, X_n)$. By a query we mean a conditional probability $P(X|Y)$ where X and Y can be lists of random variables. In a *hyperbolic Dirac net* we replace $P(X|Y)$ by $\langle X|Y\rangle$ where the X and Y can be lists of random variables. In [14] Barry Robson describes the implementation of such a net. There is good reason to believe that the physical analogies between quantum mechanics and inference nets will be fruitful because both of these fields are concerned with questions about *identical character*, *distinguishability* of states, events or measurements, *correlation*, *mutual information* and *orthogonal character.* For example the net might represent a database consisting of records in which we want to be able to analyze relationships between attributes or parameters of patients. Think of a stack of records to analyze, one row being a record and one patient, and what is on each same record is a "co-event", a joint occurrence because it comes together in one patient. The random variables above would correspond to the values of various fields and may "recur", that is appear more than once as for example what might happen if a person had a broken leg twice.

7. Dirac's bra-ket formalism

Quantum mechanics uses what can be thought of as a special notation to handle quantum uncertainty. This is called the Dirac bra-ket notation. P.A.M. Dirac introduced the bra-ket notation in [3] to the study of quantum mechanics[2]. Much more than a mere *notation*, the Dirac notation has been compared to the transition from Roman numerals to Arabic numerals in its power to express and conceptualize the ideas of quantum mechanics.

---

[2]In quantum mechanics Dirac's system is a system of inference. That is, it is a system for drawing conclusions about probabilities of events. If a particle that is prepared with momentum p on a circular track of circumference L and is observed at position x, $\langle x|p\rangle$ denotes the wave function for the pure momentum state given by

$$\langle x|p\rangle = \frac{1}{\sqrt{L}} exp\left(\frac{ixp}{\hbar}\right),$$

[1] gives a very readable introduction to the Dirac formalism in quantum mechanics for nonspecialists. From the perspective of linear algebra, given a vector space V, a ket is a vector, a bra is a dual vector. In the case of a finite dimensional vector space, we may identify kets $|y\rangle$ with column vectors and bras $\langle x|$ with row vectors. $\langle x|y\rangle$ is then a product of a row vector against a column vector-an inner product in effect. $|y\rangle\langle x|$ is called the outer product and corresponds to a matrix product of a column vector against a row vector-a matrix, or operator in effect. Much more can be said. One can discuss forms like $\langle x|A|y\rangle$ where A is a matrix representing some operator.

8. Hyperbolic Dirac nets: using Dirac's bra-ket notation with inference nets

In [14],[20], [22], Barry Robson proposes a general approach to inference nets based on the Dirac notation by building both $P(X|Y)$ and $P(Y|X)$ into a single inference engine. Because our purpose is expository our descriptions will be informal. Robson's approach suggests packaging both probabilities together by ordered pair

$$(P(X|Y), P(Y|X))\,.$$

One of the things this will enable is the ability to model relations involving reversed causality or conditionality as in "Disease X causes symptom Y" vs. "Symptom Y causes (a diagnosis of) Disease X" . To wit,

$$\langle X|Y\rangle = \sqrt{P(X|Y)}\iota + \sqrt{P(Y|X)}\iota^* \tag{1}$$

or equivalently

$$\langle X|Y\rangle = (\sqrt{P(X|Y)}, \sqrt{P(Y|X)}).$$

---

where $\hbar$ is the reduced Planck's constant. $\langle x|p\rangle$ describes the "de Broglie wave". The $\sqrt{L}$ is included as an L^2 normalization factor to ensure

$$\int_0^L |\langle x|p\rangle|^2 dx = 1.$$

$|\langle x|p\rangle|^2$ in effect gives the probability density that the particle prepared with momentum p will be observed at position x. This is quantum uncertainty: the act of observing the bead is an interaction with it that changes its momentum (destroys the momentum measurement.) The analogy that $\langle x|p\rangle$ represents the square root of a probability (density) will be mentioned again below. It is also a system of expectations(averages) of values when in $\langle X|\mathcal{R}|Y\rangle$, $\mathcal{R}$ is a Hermitian operator of observation.

Actually, the square root is not strictly necessary, and is sometimes left off, i.e.

$$\langle X|Y\rangle = P(X|Y)\iota + P(Y|X)\iota^*.$$

Ignoring the square root seems to work perfectly well[3] Either way, we can project out the ι or ι^* components: the projection functions R and R^*, defined by

$$R(a\iota + b\iota^*) = a, \quad R^*(a\iota + b\iota^*) = b, \tag{2}$$

are analogues of the Penrose reduction operator in quantum mechanics, although we should perhaps think of that here as including a Lorenz rotation $i \to h$ along with considerations discussed in [1] on page 135.

Note further that $\langle X|Y\rangle$ will have hermitian symmetry, i.e. $\langle Y|X\rangle = \langle X|Y\rangle^*$.

More generally, forms of the type

$$\langle X|\mathcal{R}|Y\rangle$$

---

[3]This seems more a matter of the choice of units by which we chose to represent distinguishable information [22], but this is also a matter of normalization and method of computation of component probabilities. Here the procedures of physicists, but now applied to purely h-complex amplitudes, are surprisingly insightful, suggesting that we should drop the square root. Of course, if our empirical probabilities are replaced by exponentials of i-complex quantities, it can be argued that the square root is physical in that periodic wave function case, but consider the following. Whilst extractable weights of vectors $\langle A|$ and $|B\rangle$ are seen as related to square roots of probabilities, $\langle A|B\rangle$ is their product, and if we allow A to recur independently such that $P(A, A) = P(A)^2$ as if A is countable like "male" in a sample of a population, $\langle A|A\rangle = P(A)$, not $\sqrt{P(A)}$. Application of Dirac's recipe to extract an observable probability [3] means ket normalization to prepare B to measure A conditional upon it, i.e. $P(B) = P(B|A) = 1$ so that $\langle A|B\rangle' = P(A|B)\iota + \iota^*$, and the Born rule as $\langle A|B\rangle'(\langle A|B\rangle')^*$then gives $P(A|B)$. Similarly, if physical, $P(A) = P(A|B) = 1$ means $'\langle A|B\rangle(\langle A|B\rangle')^* = P(B|A)$(though Robson feels that in the special case of conjugate variables A and B like momentum and position ι and ι^*should be switched to permit general normalization of $P(A|B) = P(B|A)$ without breaching the uncertainty principle). There is a simple and persuasive eigensolution interpretation: Dirac noted that h (his σ such that $\sigma\sigma = 1$) can be seen as a linear operator which, unlike i with eigenvalues $+i$ and $-i$, has real eigenvalues $+1$ and -1, and replacing h by each of these in turn gives eigensolutions $\langle A|B\rangle = P(A|B)$ and $\langle A|B\rangle = P(B|A)$ respectively.

are also permitted where $\mathcal{R}$ is interpreted as a *relator*, i.e. a relationship between X and Y. $\langle X|Y\rangle$ represents the default case where the relator $\mathcal{R}$ indicates "is directly influenced by".

To understand this better, we describe how to convert a Bayes net to an HDN. By an HDN we mean, as a provisional definition, a Dirac bra-ket factorization of the joint. Let's first consider the very simplest case of only two nodes X and Y with corresponding joint distribution defined by

$$P(X,Y) = P(Y|X)P(X). \tag{3}$$

Since by Bayes Rule in probability, $P(X|Y)P(Y) = P(Y|X)P(X)$ and $\iota + \iota^* = 1$ we may write[4]

$$\begin{aligned} P(X,Y) &= P(X,Y)\iota + P(X,Y)\iota^* = P(X|Y)P(Y)\iota + P(Y|X)P(X)\iota^* \\ &= [P(X|Y)\iota + P(Y|X)\iota^*][P(Y)\iota + P(X)\iota^*] = \langle X|Y\rangle\langle X|?\rangle\ \langle ?|Y\rangle \end{aligned}$$

where ? is understood to be a dummy[5] node for which $P(?) = 1$. In effect $\langle Y|?\rangle = P(Y)\iota + \iota^*$ and $\langle ?|X\rangle = \iota + P(X)\iota^*$. In this way one may think of constructing the corresponding HDN for (3) by the analogous product

$$P(X,Y) = \langle Y|?\rangle\langle X|Y\rangle\langle ?|X\rangle$$

For our next example consider the Bayes net defined by the factorization

$$P(X_1,X_2,X_3,X_4,X_5) = P(X_1|X_2,X_3)P(X_2|X_4)P(X_4)P(X_3|X_5)P(X_5).$$

By its dual P^* we mean the measure (pre-normalized joint) defined by

$$P^*(X_1,X_2,X_3,X_4,X_5) = P(X_1)P(X_2,X_3|X_1)P(X_4|X_2)P(X_5|X_3).$$

As before we write down the analogous split number factorization

$$H(X_1,X_2,X_3,X_4,X_5) = \langle ?|X_1\rangle\langle X_1|X_2,X_3\rangle\langle X_2|X_4\rangle\langle X_4|?\rangle\langle X_3|X_5\rangle\langle X_5|?\rangle$$

while noting that $H = P\iota + P^*\iota^*$.

[4]Recall products are now in the split-complex number system.

[5]Physicists would demur from using this language. "?" would be viewed the act of preparation or observation of the state, which we usually can assume occurs with certainty even if the probability set or measured doesn't. You can't of course distinguish in an HDN really between preparation and observation, which is sort of true in physics too, keeping in mind entanglement as "spooky action as a distance" is what would be meant.

Note that by (2) $P = RH$ and $P^* = R^*H$.

More generally by a Hyperbolic Dirac net we mean a data structure obtained by a dualization procedure performed on a given net that consists of a pair of split-complex number factorizations of a joint and it's dual. By the dual of a net we mean the new net we obtain by reversing the direction of the edges such as (8). That is, a a factorization into split-complex numbers of bra-kets of the form $\langle A|B\rangle$ of the quantity

$$H = P\iota + P^*\iota^*.$$

In computational practice this could be implemented by a dualization operation performed on the individual conditional probability tables used to implement the original net.

An HDN is a dualization[6] procedure performed on a given inference net that consists of a pair of split-complex number factorizations of the joint probability and its dual (adjoint, reverse direction of conditionality).

9. HDN queries and further analogies with quantum mechanics

As with classical probabilistic graphical models a fundamental operation on an HDN is that of a query $\langle A|B\rangle$. These often lead to formulas reminiscent of

[6]The inference net on which this dualization is performed is defined as an estimate of a probability as an expression comprising simpler probabilities and or association measures, i.e. each with fewer attributes (i.e. arguments, events, states, observations or measurements) that the joint probability estimated, where each attribute corresponds to nodes of a general graph and the probabilities or association measures represent their interdependencies as edges. It is not required that the inference net be an acyclic directed graph, but the widely used BN that satisfies that description by definition is a useful starting point for making use of the given probabilities to address the same or similar problems. Specifically for the estimation of a joint probability, and HDN properly constructed with prior probabilities, and whether or not it contains cyclic paths, is purely real valued and its construction principles represent a generalization of Bayes Theorem. Any imaginary part indicates the degree of departure from Bayes Theorem over the net as a whole, and the direction of conditionality in which the degree of departure occurs, and thus the HDN provides an excellent book-keeping tool that Bayes Theorem is satisfied overall. Specifically for the estimation of a conditional probability, it follows conversely from the above that any expression for a joint probability validated by the above means can serve as the generator of an HDN for the estimation of a conditional probability simply by dividing it by the HDN counterparts of prior probabilities, whence the resulting net is not purely real save by coincidence of probability values.

formulae from quantum mechanics. We will illustrate this in the very special but simple case of a three node Bayes net:

$$A \longleftarrow \chi \longleftarrow B\,. \tag{4}$$

with corresponding joint $P(A, \chi, B) = P(A|\chi)P(\chi|B)P(B)$ and HDN structure $P\iota + P^*\iota^*$.

Proposition. *Let $P\iota + P^*\iota^*$ be the HDN constructed from the BN (4). Then*

$$\langle A|B\rangle = \sum_{\chi} \langle A|\chi\rangle\langle\chi|B\rangle \tag{5}$$

The identity (5) is known as the *Law of Composition of Probability Amplitudes*[7] in quantum mechanics. The sum is understood to be taken over all states of an intermediate node χ. (5) can be viewed as an example of a *Feynman path integral*(see [5]). A physical example of the equation would be the behavior of electrons in a the classical experiment in quantum mechanics in which stream of electrons are fired toward a backstop with a mounted detector obstructed by a wall with multiple holes through which the electrons may pass through. (See ([6]) volume III chapter 3.)

---

[7] The quantum mechanical composition law holds because quantum mechanics is defined in a vector space. Indeed, the composition law is the same as the dot product of two vectors $[\langle A|X1\rangle, \langle A|X2, \ldots]$ and $[\langle X1|B\rangle, \langle X2|B\rangle, \ldots]$ (although we should not use the term "dot product" when multiplying a row and column vector). That a vector space is the right thing to use in quantum mechanics is a matter of agreement with experiment. It can be shown that classically it is not a matter of "coherence" (probability bookkeeping to satisfy Bayes rule and marginal summation) as its appearance might suggest, but rather of the quality of the estimate of the implied joint probability, since it works classically under the assumption of certain independencies. The implications are interesting because these interdependencies must therefore hold in quantum mechanics, in the vector space, in stark contrast to the strong interdependencies between quantum mechanical conjugate variables such as p and x, which is not seen classically except in certain special cases (e.g. the gas law $PV = RT$ for constant temperature T, Ohm's electrical law $V = IR$, and so on), and certainly not for p and x. It is almost as if degrees of interdependency are much more discrete, all-or-nothing and almost "quantized", for fundamental particles but not for things and events involving vast numbers of them. However, it is not perhaps surprising because we see this discreteness already in the degree of distinguishability of fundamental particles, i.e. as in fermion and boson statistics.

Proof. Note first that

$$P(A|B) = \frac{P(A,B)}{P(B)} = \frac{1}{P(B)} \sum_{\chi} P(A,\chi,B)$$
$$= \frac{1}{P(B)} \sum_{\chi} P(A|\chi)P(\chi|B)P(B)$$
$$= \sum_{\chi} P(A|\chi)P(\chi|B).$$

At the same time

$$P(B|A) = P(A|B)\frac{P(B)}{P(A)} = \frac{P(B)}{P(A)} \sum_{\chi} P(A|\chi)P(\chi|B)$$
$$= \frac{P(B)}{P(A)} \sum_{\chi} P(\chi|A)\frac{P(A)}{P(\chi)}\frac{P(B|\chi)P(\chi)}{P(B)}$$
$$= \sum_{\chi} P(\chi|A)P(B|\chi).$$

Thus

$$\langle A, B\rangle = P(A|B)\iota + P(B|A)\iota^*$$
$$= \left(\sum_{\chi} P(A|\chi)P(\chi|B)\right)\iota + \left(\sum_{\chi} P(\chi|A)P(B|\chi)\right)\iota^*$$
$$= \sum_{\chi} \langle A|\chi\rangle\langle\chi|B\rangle.$$

□

Actually, we mean here not "$P(A|B) =$" and "$P(B|A) =$" but estimates of them, recalling that an inference network is an estimate of a conditional or joint probability which is therefore of direct interest here. As proof of an exact law of composition of probabilities, it is easy to show that we must have $\sum_{\chi}[P(A|B,\chi) - P(A|\chi)] = 0$. Since any component $P(A|B,\chi) - P(A|\chi)$ can be positive or negative leading to overall cancelation, this is possible, and less restrictive than various assumptions about A, B, and independencies.

Something may be said about information theoretic and thermodynamic models and analogies that facilitate conceptual and operational aspects of these HDNs, and they are insightful in many respects. In physics, $P(A|B)$

is essentially represented by $P(A)e^{I(A;B)}$, where P(A) represent the probability of A prior to the information obtained from measurement (when $I(A;B) = 0$), as determined by the structure and scale of the system of interest. Mutual information $I(A;B) = I(B;A) = \ln(P(A,B)/P(A)P(B))$ is in quantum mechanics A/$\hbar$ where A is the physical *action* such as energy E times time, or momentum times position, and $\hbar$ is reduced Planck's constant. For purposes below we note here the *atomic form* for more than two arguments: $I(A;B;C\dots) = \ln(P(A,B,C,\dots)/P(A)P(B)P(C)\dots)$ "from which other forms like from which other forms like $I(A,B;C,D) = I(A;B;C;D) - I(A;B) - I(C;D)$ are readily produced. These examples describe, however, pairs of physical conjugate variables in quantum mechanics, where one determines the other such that $P(A|B) = P(B|A)$, and indeed we are really dealing with $e^{hI(A;B)}$ and for our analogy focusing on $e^{hI(A;B)} = e^{+I(A;B)}\iota + e^{-I(A;B)}\iota^*$ where, as in Dirac's system, a special form of normalization is required to generate $P(A|B)$ (Dirac's normalization followed by the Born rule) as the observable probability. To break conjugate symmetry we have to introduce an external interaction, of which an act of observation of the wave as a particle is one example, or simply think classically from the outset. Classically, the above examples are no longer conjugate variables, and in any event for many physical systems it is sufficient for our purposes, and well known, that we can write $I(A;B) = E/kT$ where E is the energy, T absolute temperature, and k is Boltzmann's constant. In both quantum and classical cases the divisor ($\hbar$ or kT) represents the uncertainty or noise which dictate how accurately the information can actually be measured, and ensure that the information $I(A;B)$ is expressed in dimensionless units (though we speak of "bits", binary log units, or "nats", natural log units, of information) Note that consideration of the normalization for conjugate variables for the purely h-complex case does not suggest a need for square roots of probabilities, but that nonetheless we are free to use any root or power to express the units based on the accuracy with which we can measure. One of several consistent reasons for appearance of the square root in quantum mechanics is that the lower limit of fineness of measuring action A is actually $A = \Delta 2\hbar$. Importantly note also that the self probabilities and mutual information or association constants can be used to express an HDN, because we can also equivalently write

$$\langle A|B\rangle = [P(A)\iota + P(B)\iota^*]e^{I(A;B)} = [P(A)\iota + P(B)\iota^*]K(A;B)$$

As a consequence of the algebra of h, we can also write

$$kT \ln\langle A|B\rangle = E(A)\iota + E(B)\iota^* + E(A;B)$$

which provides the energy model, since E(A) and E(B) are the energies internal to a particle, atom, or molecule A or B respectively, and E(A; B) = E(B; A) is the *interaction energy* between A and B. Perhaps the most important conceptual aid provided is when considering that an HDN is a bidirectional general graph, in contrast to a BN as a unidirectional acyclic graph. It is is the cyclic aspect that is controversial, but it is evident that all of A, B, C, D, etc. in a system can have some degree of energy of interaction involving two or more of them, implying multiple cycles. To make use of this concept, there are rules to follow. Objects such as A and B can clearly appear more than once in all the interaction terms added up, say E(A; B) + E(A; C; D) + E(B; F), but what we cannot do is count an interaction more than once. For example, E(A; B; C) + E(A; C; D; F) would count the A with C interaction twice. Similarly, but perhaps less obviously as it is somehow less visual, we cannot count mutual information between entities twice. The recipe for an HDN is obvious, although in HDNs we apply it to the product of all association constants such as $K(A;B;C) = e^{E(A;B;C)}/kT$ that we consider as departing significantly from 1 (i.e. $I(A;B;C)$ and $E(A;B;C)$ departing significantly from 0), multiplied by all the self probabilities P(A), P(B), P(C) etc. That gives a valid estimate joint probability for which conditionals can be readily formed, considerably simpler than the way in which we normally think about constructing a BN. It could of course be argued that this is a defect of BNs arising from their acyclic philosophy, rather than an advantage of HDNs. However, it does take two directions of conditionally and an h-complex algebraic treatment to see it. That is, to see and prove that cyclic paths (without branches or with branches corrected for symmetry) are purely real valued and in effect valid joint probabilities. For further background on these matters we refer the reader to [4],[16],[17],[18].

10. An example

10.1. A homeostasis problem starting from a Bayes net

By way of example, we should consider the computation of joint probabilities, which is one of the main functions of a BN. Compared with a conditional probability where we have an h-complex value, a properly computed joint probability based on coherent probabilities must be real, so that appears to be the boring case. The issue is, of course, is it purely real, i.e. is it properly computed and based on coherent probabilities? This also brings to light other issues of probabilistic modeling that are, perhaps, not at all obvious. To illustrate this, we will use a problem posed in the Wikipedia entry of BNs [26]. Although about keeping a lawn well-watered, it has a form common to probabilistic models of homeostasis. Suppose that there are two events which could cause grass to be wet [G]: either the sprinkler is on [S] or it's raining [R]. Also, suppose that the rain has a direct effect on the use of the sprinkler (namely that when it rains, the sprinkler is usually not turned on). The joint probability function is: $P(G, S, R) = P(G|S, R)P(S|R)P(R)$. As regards estimating joint probabilities by means of a BN, Rebonato has noticed that human experts are rather bad at assigning probabilities that are coherent, i.e. satisfying both marginal summation and Bayes rule. Robson has given several worked HDN examples [19] that highlight this and show how an HDN can repair the difficulty, as well as providing a broader set of examples that highlight issues in relating HDNs to probabilities of statements in natural language [21]. The above sprinkler example is, however, of interest in illustrating key differences between a BN and an HDN because it is already coherent. The issues here therefore relate to alternative representations and new insight, not repair. We know that it is coherent because(assuming no trivial erros in probability assignments) the above expansion is exact: $P(G|S, R)P(S|R)P(R) = P(G, S, R)P(S, R)P(R)/P(S, R)P(R) = P(G, S, R)$. The solution for a joint probability, using the probabilities given in this source, is

$$P(G, S, R) = P(G|S, R)P(S|R)P(R) = 0.99 \times 0.01 \times 0.2 = 0.00198.$$

That a BN represents an exact solution is less common (except for introductory and educational purposes). That is because if we have enough data to count and obtain $P(G|S, R)$, then we have enough to obtain joint probability $P(G, S, R)$ directly, and therefore no BN is required to provide an estimate per se. Note also that, in the above P(G,S,R) expansion example, the following rule for a joint probability is satisfied. We write every probability as self or joint probability by removing the conditional bar if present, and divide by the probabilities of single or joint events on which they are conditional. This will apply to brakets like $\langle A|B, C\rangle$, although there we have two directions of conditionality to worry about, relating to $P(A|B, C)$ and $P(B, C|A)$. When written in a numerator-divisor form using only joint and self-probabilities, each state or event A, B, etc. must occur once more in the numerator than denominator. Re-expressing probabilities as information $I(x) = -\ln P(x)$ where x is a self or joint probability, shows that this makes sense. We can represent information contributions as we wish, e.g. isolated or $I(A)$ or jointly as in $I(A, B)$ and in $I(A, B, C)$ which progressively recognize more interdependencies, but we add or remove information so that it is not counted twice. Incidentally, this is not an absolute rule for an HDN. $\langle A, A|A, A, A\rangle$ can make sense just as $P(A, A|A, A, A)$ can and counting does: what is the probability of seeing A twice after seeing it three times? It might be noted that this breaks the rule "each interaction once" by analogy with thermodynamics. Strictly speaking, that is so, but only in the sense that some particles, molecules or other entries appear with the same name and we chose not to distinguish them, as in counting black balls from an urn or in counting white balls from an urn, or males in a room. Dependent on the nature of the sampling and size of sample, we may still need to be careful not to count a same thing twice and thereby to avoid including more than once interactions involving it. See the very end of our paper for comment on this!

10.2. HDN solutions

In one sense the obvious solution is that $P(G|S, R)P(S|R)P(R)$ directly suggests $\langle G|S, R\rangle\langle S|R\rangle\langle R|?\rangle$. That it does so is an example of a first step in forming a HDN when one wants to start from a BN, which is obviously a highly automatable approach. However, it has certain asymmetries contravening coherence that will need repairing as discussed later below. That is, the conditional probabilities will be implied in at least one direction and we not have the purely real value implied by the dual $(x, y) = x = y$. In another sense, when entering the problem with a basic awareness of probability theory and need for coherence, there are two other obvious direct solutions. They might be also considered as obvious basic repairs to $\langle G|S, R\rangle\langle S|R\rangle\langle R|?\rangle$. The first and simplest preserves the interdependency between S and R from the outset. That it is coherent with respect to Bayes rule can be seen because the dual balances and implies a purely real number:

$$\begin{aligned}
&\langle ?|G\rangle\langle G|S, R\rangle\langle S, R|?\rangle \\
&= \big(P(G)P(G, S, R)P(S, R)/P(S, R)P(G), \\
&\quad P(G)P(G, S, R)P(S, R)/P(G)P(S, R)\big) \\
&= \big(P(G, S, R), P(G, S, R)\big) = P(G, S, R) = 0.00198.
\end{aligned}$$

In practice, this approach is reasonable. If we have data to evaluate $P(G|S, R)$ and $P(G, S, R)$ we certainly have data to evaluate $P(S, R)$ and hence$\langle S, R|?\rangle$. The second solution follows the BN example more closely by having a dependency on rain with the self-probability P(R) as adjustable input. If we wish we can write $\langle ?|G\rangle\langle G|S, R\rangle\langle S|?\rangle\langle R|?\rangle$, however, the HDN sees explicitly that there is an asymmetry. S and R are interdependent in one direction of conditionality, and conditional in another. We are in effect repairing the branch $(\dots R, S) \to (R, \dots), (S, \dots)$ so that to the right of the arrow, R and S are not seen as independent. We introduce (multiply by)

$$\langle S; R|?\rangle = (K(S; R), 1)$$

to achieve this repair. Note that association constant $K(S; R) = P(S, R)/P(S)P(R)$.

$$\begin{aligned}
&\langle ?|G\rangle\langle G|S, R\rangle\langle S, R|?\rangle\langle S|?\rangle\langle R|?\rangle \\
&= \big(P(G)P(G, S, R)P(S)P(R)/P(S, R)P(G), \\
&\quad P(G)P(G, S, R)P(S)P(R)/P(G)P(S)P(R)\big) \times (P(S, R)/P(S)P(R), 1) \\
&= \big(P(G, S, R), P(G, S, R)\big) = P(G, S, R) = 0.00198.
\end{aligned}$$

10.3. A step by step example

We should illustrate the above by a step by example, as we would reach $P(G, S, R) = 0.00198$ in practice. Unfortunately, while the BN example allows, in effect, the calculation of $P(G, S, R)$ given $P(R)$, effectively using $P(G, S, R) = P(G, S|R)P(R)$, the BN describes only one direction of conditionality. Therefore, there are some required probabilities unspecified and which are not totally determined by the above. For example we cannot use $P(G, S, R) = P(S, R|G)P(G)$ without assigning values to $P(S, R|G)$ and to $P(G)$. In the absence of such information, one might use logic as a guideline. Using wet grass for G etc. explicitly for clarity, we should not expect P(wet grass) to be less or significantly more than (using inclusive OR) P(rain OR sprinkler) = P(rain) + P(sprinkler)P(sprinkler, rain) = 0.2 + 0.322 × 0.022 = 0.5. However, we here try to make it a little more interesting, since it would seem that there should be sufficient degrees of freedom to express say P(wet grass) = 0.8, indicating that there could be another reason for wet grass such as dew, flooding from a stream, a water mains leak, and so on. For the model

$$\langle ?|G\rangle\langle G|S, R\rangle\langle S; R|?\rangle\langle S|?\rangle\langle R|?\rangle =$$

the following assignments are then consistent with the joint probability of 0.00198 obtained by Ref. [26] and the probabilities given there or deducible from them.

$$\begin{aligned}
\langle \text{rain}|?\rangle &= (0.2, 1) \times \langle \text{sprinkler}|?\rangle \\
&= (0.322, 1) \times \langle \text{wet grass} \,|\, \text{sprinkler, rain}\rangle \\
&= (0.99, 0.002475) \\
&= (0.063756, 0.002475) \times (K(\text{sprinkler; rain}), 1) \\
&= (P(\text{sprinkler, rain})/P(\text{sprinkler})P(\text{rain}), 1) \\
&= (0.002/0.3220.2, 1) = (0.03106, 1) \\
&= (0.00198, 0.002475) \times \langle ?|\text{wet grass}\rangle \\
&= (1, 0.8) = (0.00198, 0.00198) = 0.00198
\end{aligned}$$

To ask the what is the probability that there is wet grass and the sprinkler is on given that it is raining are involved in the grass being wet, divide by $\langle \text{rain}|?\rangle = (0.2, 1)$, i.e.

$$\langle \text{rain, sprinkler} \,|\, \text{wet grass}\rangle = 0.00198/(0.2, 1) = (0.0099, 0.00198).$$

The probability that there is rain and wet grass if the sprinkler is on is

$$\langle \text{rain, wet grass} \,|\, \text{sprinkler}\rangle = 0.00198/(0.322, 1) = (0.00615, 0.00198).$$

The following requires the HDN, at least in the sense that the original BN cannot answer it without acknowledging P(wet grass). To ask the new question, what is the probability that is raining and the sprinkler is on if the grass is wet, divide by $\langle ?|\text{wet-grass}\rangle = (1, 0.8)$, i.e.

$$\langle \text{rain, sprinkler} \,|\, \text{wet grass}\rangle = 0.00198/(1, 0.8) = (0.00198, 0.002475).$$

How generally does this hold for the sprinkler example? To compute all possible 8 combinations of rain/not rain, sprinkler/not sprinkler/ wet grass/ not wet grass we initially had to guess at some probabilities not given or deducible from the source reference. These include P(not wet grass | not sprinkler, not rain) and P(not sprinkler, not rain | not wet grass). To proceed more objectively, an automatic marginalization procedure was developed to give consistent marginalization results. This gave a balanced, purely real dual $(x, y) = x = y$ up to at least three significant digits d in $d.dd \times 10^{-e}$ exponent format for all 8 combinations of negation, marginalization involved not only deducing implied probability values but also adjustment to the given probabilities including the above guesses if they were no fully coherent with respect to known probabilities overall. One solution for a full set of coherent and realistic probabilities is $P(R) = 0.2008$, $P(S) = 0.32$, $P(G) = 0.5295$, $P(R, S) = 0.0007161$, $P(R, \text{not } S) = 0.2003$, $P(\text{not } R, S) = 0.35105$, $P(\text{not } R, \text{not } S) = 0.4490$, $P(G, S, R) = 0.0007160$, $P(R, S, \text{ not } G) = 6.11510^{-008}$. Note that $P(\text{wet grass}) = 0.5295$, and recall that we should not expect P(wet grass) to be less or significantly more than (using inclusive OR) P(rain OR sprinkler) = P(rain) +P(sprinkler) P(sprinkler, rain) = 0.2 + 0.322 × 0.022 = 0.5.

10.4. An alternative model solution

If one is seeking to reproduce the full sprit of the original BN example, however, there is a less obvious solution or family of solutions based on the above mentioned fact that $P(G|S, R)P(S|R)P(R)$ directly suggests $\langle G|S, R\rangle\langle S|R\rangle$ $\langle R|?\rangle$. We might then seek to repair by thinking about the following. (a) The self-probability of wet grass, i.e. input in the other direction already discussed above, so that we need to introduce (multiply by) $\langle ?|G\rangle$. (b) Sprinkler and rain are again interdependent in one direction and not another, so that we require introducing (and multiplying by) $\langle R; S|?\rangle$ again. (c) Most controversially, the fact that the sprinkler BN is an exact expansion makes the input of the BN is somewhat unusual, and we need two conditionalities on rain, otherwise the R in $\langle G|S, R\rangle$ has a prior input via $\langle S|R\rangle\langle R|?\rangle$ but is left hanging without an implied input prior probability P(R) represented by $\langle R|?\rangle$. The latter is a common structure and intuitive interpretation, so it is useful to consider the consequences. Writing the using round brackets in such a way where we can see the implied branch to which (c) relates, the repaired HDN is $\langle ?|G\rangle\langle G|S, R\rangle\langle R; S|?\rangle(\langle R|?\rangle, \langle S|R\rangle\langle R|?\rangle)$. The bracket contents are not a dual, we simply mean $\langle ?|G\rangle\langle G|S, R\rangle\langle R; S|?\rangle\langle R|?\rangle\langle S|R\rangle\langle R|?\rangle$. Again, to make it more readable, we will use wet grass for G, R for rain, and S for sprinkler. For the rain, sprinkler, and wet grass case, the joint probability is no longer 0.00198, $\langle \text{rain}|?\rangle = (0.2, 1)$, $\langle \text{sprinkler} \,|\, \text{rain}\rangle = (0.01, 0.00621)$,

$$
\begin{aligned}
&\langle \text{wet grass} \,|\, \text{sprinkler, rain}\rangle = (0.99, 0.002475)\\
&= (0.00198, 0.00001537) \times (K(\text{sprinkler; rain}), 1)\\
&= (P(\text{sprinkler, rain})/P(\text{sprinkler})P(\text{rain}, 1))\\
&= (0.002/0.3220.2, 1) = (0.03106, 1) \times \langle ?|\text{wet grass}\rangle = (1, 0.8)\\
&= (0.0000614988, 0.000012296) \times \langle \text{rain}|?\rangle = (0.2, 1)\\
&= (0.00001229976, 0.000012296) = 0.0000123.
\end{aligned}
$$

Briefly stated, such a symmetric dual above holds reasonably over all 8 combinations of negation. Using the above probability assignments obtained by marginal summation and notably with P(wet grass) = 0.5295, we again obtained the more precise result $(x, y) = x = y$ up to at least three significant digits d in d.dd x 10-e format. More important here, however, is the question of how such a result can possibly be correct, even if it does exhibit reasonable choices of probability that are coherent. We obtained earlier above the solution $\langle ?|G\rangle\langle G|S, R\rangle\langle S; R|?\rangle\langle S|?\rangle\langle R|?\rangle = 0.00198$. It is even easier to show also that $\langle ?|G\rangle\langle G|S, R\rangle\langle S, R|?\rangle = 0.00198$. These agree with the original BN calculation. For $\langle ?|G\rangle\langle G|S, R\rangle\langle R; S|?\rangle(\langle R|?\rangle, \langle S|R\rangle\langle R|?\rangle)$, however, we obtained 0.0000123. Evidently, we cannot of course be calculating the same thing. Indeed, it is readily shown that

$$\begin{aligned}&\langle ?|G\rangle\langle G|S, R\rangle\langle R; S|?\rangle(\langle R|?\rangle, \langle S|R\rangle\langle R|?\rangle)\\ =&P(G, S, R)P(R|S).\end{aligned}$$

Consequently, one might see the correction required as simply

$$\begin{aligned}P(\text{rain} \,|\, \text{sprinkler}) &= P(\text{sprinkler, rain})/P(\text{sprinkler})\\ &= 0.002/0.322 = 0.00621\end{aligned}$$

$P(G, S, R)P(R|S) = 0.0000123$, so $P(G, S, R) = P(G, S, R)P(R|S)/0.00621$ $= 0.00198$. From one perspective, therefore, this is not so much wrong as a miscalculation of P(G,S,R) but a correct estimate of something other than P(G,S,R). It may be considered as one particular kind of estimate of P(G,S,R, R), recalling the validity of $\langle A, A|A, A, A\rangle$ and $P(A, A|A, A, A)$ discussed above. Perhaps more reasonably in this case, it might be considered the impact of two manifestations or interpretation of rain, R as rain that wets the grass, but also R_S that may or not identify with the same rain but is rain as perceived as relative to the sprinkler, giving us $P(G, S, R, R_S)$. At the same time, noting the above correction of diving by $P(R|S)$, it may be said that we were initially estimating some kind of joint probability and then considering what it becomes given that there is rain relevant to the action of the sprinkler. After all, it may be that the owner or automatic device shuts down the sprinkler if it detects dull light and light rain in the air as humidity. We have in effect split R into two subclasses R (perhaps conceptually closer to the original R) and R_S for which we can write $\langle R_S|R\rangle = P(R_S|R)$ without necessarily saying or finding that $\langle R_S|R\rangle = P(R_S|R) = 1$. The situation is not qualitatively different from separating female patients into pregnant and not-pregnant. Certainly the following where R_S is held distinct from R exemplifies a graph structure that, with other symbols than G, S, and R of course, is not uncommon, and is valid:

$$\langle ?|G\rangle\langle G|S,R\rangle\langle R;S|?\rangle(\langle R|?\rangle,\langle S|R_S\rangle\langle R_S|?\rangle).$$

This is yet another appearance of a reference to counting and distinguishability, and how much we choose to distinguish things. We chose now to distinguish two kinds of rain, and the Inuit are reported to have a whole of words for distinguishing kinds of snow. It is quite a common issue that come up on the HDN and with the hyperbolic bracket, which seems to give greater power, if indeed imposing more responsibility and need for effort, to manage such issues. All real values of $\langle A|B\rangle$ lie on a continuum from 1 (the case of absolute identity between A and B, they are the same thing) to $\langle A|B\rangle = 0$ where A and B are absolutely distinguishable: classical probability theory would say mutually exclusive, while QM would speak of $\langle A|$ and $|B\rangle$ as orthogonal vectors. To ensure we do mean real values, we can perhaps better speak of $\langle A|A\rangle$, which with a value $\langle A|A\rangle = 1$ means absolute equivalence, cannot be meaningfully considered as occurring a second time and countable, while $\langle A|A\rangle = 0$, means that A cannot occur more than once, ...and so cannot be meaningfully considered as occurring a second time and countable. This is a curious full circle, operationally, at least, bringing 0 and 1 in some sense together. It is one of many interesting things that can emerge from considering number systems with an extra and imaginary dimension, and hence a richer information content and perspective.

Note added in press

On reflection, one of the things that we feel that we have learned from these studies is that the Bayes net, being traditionally confined to a directed acyclic graph, is not only unnecessarily restrictive, but also at risk of being severely unrealistic by neglecting so many interactions in order to satisfy that restriction. Most notably, it cannot traditionally have cyclic paths. Also by being unidirectional, it is easy to make mistakes: in many respects the matter of coherence regarding Bayes rule and marginal summation is very important, but in a Bayes Net it cannot be seen, because it is purely unidirectional in conditionality, and so ironically has little to do with Bayes rule. One could say that Bayes rule is really all about the duals $(P(A|B), P(B|A))$ and $(P(B), P(A))$, and their product.

References

[1] Marvin Chester, *Primer of Quantum Mechanics*, Dover Publications, 2003.

[2] James Cockle, *A New Imaginary in Algebra*, London-Edinburgh-Dublin Philosophical Magazine (3) 33, pp. 345–349, 1848.

[3] Paul A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, 1930.

[4] R. Fano, *Transmission of information; a statistical theory of communications*, MIT Press, Cambridge, MA, 1961.

[5] Richard P. Feynman and Albert R. Hibbs *Quantum Mechanics and Path Integrals*, McGraw-Hill New York, 1965.

[6] Richard P. Feynman, Robert B. Leighton and Matthew Sands, *The Feynman Lectures on Physics* (3 volumes), Addison Wesley, Reading Massachusetts, 1965.

[7] Paul R. Halmos, *Finite Dimensional Vector Spaces*, Springer-Verlag New York, 1987.

[8] Andrei Khrenikov and Gavriel Segre, *An Introduction to Hyperbolic Analysis*, `arXiv:math-ph/0507053v2`.

[9] Daphne Koller and Nir Friedman, *Probabilistic Graphical Models Principles and Techniques*, MIT Press, 2009.

[10] Lior Pachter and Bernd Sturmfels, *Algebraic Statistics For Computational Biology*, Cambridge University Press, New York, 2005.

[11] Roger Penrose, *The Road to Reality. A Complete Guide to the Laws of the Universe*, Vintage Books, New York, 2004.

[12] George Polya, *Mathematics and Plausible Reasoning Volume I, Induction and Analogy in Mathematics*, Princeton University Press, 1954.

[13] Robert D. Richtmyer, *Principles of Advanced Mathematical Physics, Vol. I*, Springer-Verlag 1985.

[14] Barry Robson, The New Physician as Unwitting Quantum Mechanic: Is Adapting Dirac's Inference System Best Practice for Personalized Medicine, Genomics and Proteomics?, J. Proteome Res. (Am. Chem. Soc.), **6**(8) 3114–3126, 2007.

[15] Barry Robson, *Quantum Mechanics in Split Complex Algebra Yields Classical Behavior and Inference*, preprint.

[16] Barry Robson, *Clinical and Pharmacogenomic Data Mining. 1. The generalized theory of expected information and application to the development of tools*, J. Proteome Res. (Am. Chem. Soc.) 2(3), 283-302, 2003.

[17] Barry Robson, *Clinical and Pharmacogenomic Data Mining: 3. Zeta Theory As a General Tactic for Clinical Bioinformatics*, J. Proteome Res. (Am. Chem. Soc.) **4**(2); 445–455, 2005.

[18] Barry Robson, *Clinical and Pharmacogenomic Data Mining: 4. The FANO Program and Command Set as an Example of Tools for Biomedical Discovery and Evidence Based Medicine*, J. Proteome Res., (A. Chem. Soc.), **7**(9), pp. 3922–3947, 2008.

[19] Barry Robson, *Hyperbolic Dirac Nets for Medical Decision Support. Theory, Methods, and Comparison with Bayes Nets*, Computers in Biology and Medicine, in 2014 Aug;51:183-97.

[20] Barry Robson, *Towards Automated Reasoning for Drug Discovery and Pharmaceutical Business Intelligence*, Pharmaceutical Technology & Drug Research 2012 1:3 (27 March 2012).

[21] Barry Robson, *POPPER, a Simple Programming Language for Probabilistic Semantic Inference in Medicine*, 2014.

[22] Barry Robson, Thomas P. Caruso and Ulysses G. J. Balis, *Suggestions for a Web Based Universal Exchange and Inference Language for Medicine*, Computers in Biology and Medicine, **43**(12):2297–2310, 2013.

[23] Stuart Russell and Peter Norvig, *Artificial Intelligence, A Modern Approach, Second Edition*, Prentice Hall, Upper Saddle River NJ, 2003.

[24] Garret Sobczyk, *The Hyperbolic Number Plane*, The College Mathematics Journal, Vol. 26, No. 4(Sep., 1995), pp. 268–280

[25] I. M. Yaglom, *Complex Numbers in Geometry*, Academic Press, New York and London, 1968.

[26] `http://en.wikipedia.org/wiki/Bayesian_network`

Department of Mathematics, Statistics and Computer Science
University of Wisconsin-Stout, Menomonie, WI 54751, USA
E-mail address: deckelmans@uwstout.edu

Quantal Semantics Inc., North Carolina, USA
St. Matthews University, Cayman Islands
and The Dirac Foundation, Oxfordshirei, UK
E-mail address: robsonb@aol.com

Accepted December 17, 2014