LNF-91/068 (R)
15 ottobre 1991

E. Etim:

NON-BAYESAN DIAGNOSIS

Contribution to the
Fourth Advanced Course of Radiation Protection
Bressanone 3 – 7 September 1990 - Italy

INFN - Laboratori Nazionali di Frascati
Servizio Documentazione

LNF-91/068 (R)
15 Ottobre 1991

NON - BAYESIAN DIAGNOSIS

E. Etim
INFN - Laboratori Nazionali di Frascati, P.O. Box 13, I-00044 Frascati ROMA (Italy)

ABSTRACT

Diagnosis is usually understood to be, basically, a strategy, that is decision making in the face of uncertainties. In other words, it is all a matter of probabilities and no dynamics. This is not completely true. The crucial assumption which determines the structure of a diagnosis bears mostly on dynamics rather than on probabilities. The assumption is that of microscopic reversibility. It is this assumption which gives rise to the equation of detailed balance known, in probability theory, under the name of Bayes's theorem. Microscopic reversibility does not, however, imply, necessarily, time reversal invariance. By their very nature cause-effect correlations involve dynamics with a prefered direction of time. The time reversal invariant part of this dynamics can be projected out and shown to give rise to a detailed balance equation in two pairs of variables and thence to a generalisation of Bayes's theorem. The operator which realises this projection is the diagnosis.

1. INTRODUCTION

To the layman a diagnosis is the determination of a disease given its symptoms. Technically, it is neither quite so direct nor so simple. A diagnosis consists of two distinct processes: first, the determination of the probabilities for a disease given the symptoms and, secondly a decision or option for a particular disease given this probability distribution. Any rulefor reaching such a decision is called a strategy. It is in this sense that diagnosis is a strategy. It is clinical strategy.

On what basis are such rules formulated? The most obvious and common basis is that of utility. It consists in the optimization of benefits. Already, one notes that the utility assumption is outside of the strict confines of probability theory. In fact the probability for the particular disease, given the symptoms, which optimises the benefits may not necessarily be the largest in the distribution. There is a more primitive basis for decision-making. It operates most of the time and with surprisingly good results. It is decision making based on knowledge accummulated in the long process of learning from experience. Decision reached on this basis is known as induction. Induction is the form of decision- making that has and will continue to serve mankind best in his multiple activities.

The principal aim of induction is to distinguish one hypothesis (or guess) from another, in particular a more "reasonable" hypothesis from a less "reasonable" one. The theory of probability serves inductive decision-making by supplying a mathematically consistent quantification of "reasonableness" or "degree of belief". From this perspective, one considers a disease and a symptom to be special instances of the generic concepts of cause and effect, respectively. By definition, a cause is a hypothesis, an input, a disease etc. An effect, on the other hand, is an observation, an output, a symptom. etc. The conditional probability for an effect, given a cause, is known as the likelihood. The conditional probability for a cause, given an effect, is known as the posterior probability. A diagnosis is any linear relationship between these conditional probabilities.

The oldest and most commonly used form of diagnosis is that provided by Bayes's theorem. This theorem is usually derived from what, in probability theory, is a trivial identity. This is the relationship between the conditional probabilities $P(C \mid E)$, $P(E \mid C)$ of two events C and E and involving the absolute probabilities $\mu(C)$, $\mu(E)$, of these events. This relationship is a trivial consequence of the symmetry $C \cap E = E \cap C$ and hence of the equality $\mu(C \cap E) = \mu$ 15$(E \cap C)$ together with the definition of conditional probability. $C \cap E$ is the intersection of the event C with the event E. The above symmetry, however, is only a probabilistic model for a more general symmetry property. Cause-effect correlations obviously involve much more than only probabilities. By their very nature, these correlations must have an inherent time asymmetric structure. Such structure is determined by dynamics. The symmetry which the probabilistic model exemplifies is that of microscopic reversibility. This later gives rise to an equation of detailed balance. This is precisely the kind of equation relating $P(C \mid E)$ to $P(E \mid C)$ with $\mu(C)$ and $\mu(E)$ featuring as constants of proportionality. It is proposed in this paper to show how this symmetry arises in general and how it may be projected out of a time asymmetric evolution in probability space. The operator which realises this projection is, by definition, the diagnosis. The time reversal invariant part of the evolution operator satisfies an equation of detailed balance in two pairs of variables. These variables are the generalisations of $\mu(C)$ and $\mu(E)$ together with their transforms under the diagnosis. This equation leads to a generalisation of Bayes's theorem and constitutes, at the same time, an important correction to Bayes's theorem.

2. - BAYESIAN DIAGNOSIS

Let D be a probability space with measure μ. Let $C_\alpha \subseteq D$ ($\alpha = 1, 2,...$) and $E_j \subseteq D$ ($j = 1, 2,....$) be two finite or denumerable systems of subsets of D. A subset of D is also referred to as an event, the points of D atomic events. A system of subsets, e.g. C_α ($\alpha = 1, 2, ...$) is said to consist of mutually exclusive or pair-wise disjoint events if, for any two distinct subsets C_α, C_β (α,β = 1, 2 ...) of the system, one has

$$C_\alpha \cap C_\beta = \varnothing;\ \alpha \neq \beta;\ \alpha,\ \beta = 1, 2, ... \tag{1}$$

where $\varnothing$ is the null subset. Two disjoint events are said to be independent. The system is said to be exhaustive if the union of all its subsets is the entire space D., i.e

$$D = \bigcup_{\alpha=1} C_\alpha \tag{2}$$

A system of mutually exclusive and exhaustive events is said to be complete and to form a basis for D. Let the system of events $C_\alpha(\alpha = 1,2,...)$ and E_j ($j = 1,2,...$) be complete. Then any event $A \subseteq D$ admits, on the basis of eq. (2), the following expansion in the C - basis

$$A = A \cap D = \bigcup_{\alpha=1} (A \cap C_\alpha) \tag{3.1}$$

Similarly, we have the following expansion

$$A = \bigcup_{j=1} (A \cap E_j) \tag{3.2}$$

in the E – basis. In particular, putting $A \equiv E_j$ in eq. (3.1), one gets

$$E_j = \bigcup_{\alpha=1} (E_j \cap C_\alpha) \tag{4.1}$$

while putting $A \equiv C_\alpha$ in eq. (3.2) yields

$$C_\alpha = \bigcup_{j=1} (C_\alpha \cap E_j) \tag{4.2}$$

Applying the measure μ to eqs. (4) one gets the expansions

$$\mu(E_j) = \sum_{\alpha=1} \mu(E_j \cap C_\alpha) \tag{5.1}$$

$$\mu(C_\alpha) = \sum_{j=1} \mu(C_\alpha \cap E_j) \tag{5.2}$$

Since the C – and E – bases of subsets are complete and since $\mu(D) = 1$, we have, from eq. (2) and its equivalent for $C_\alpha \to E_j$.

$$\sum_{\alpha=1} \mu(C_\alpha) = 1 \tag{6.1}$$

$$\sum_{j=1} \mu(E_j) = 1 \tag{6.2}$$

The concept of conditional probabilities arises in connection with the possibility of expressing the measure $\mu\ (C_\alpha \cap E_j) = \mu\ (E_j \cap C_\alpha)$ of the intersection of the events C_α and E_j in terms of their measures $\mu\ (C_\alpha)$ and $\mu\ (E_j)$ respectively.
Accordingly, one defines

$$P\ (E_j \mid C_\alpha) := \frac{\mu(E_j \cap C_\alpha)}{\mu(C_\alpha)} \tag{7.1}$$

$$P\ (C_\alpha \mid E_j) := \frac{\mu(C_\alpha \cap E_j)}{\mu(E_j)} \tag{7.2}$$

to be, respectively, the conditional probability of the event E_j given C_α and the conditional probability of the event C_α given E_j. Extremely important in eqs. (7), but so obvious that it is usually ignored, is the symmetry

$$\mu\ (C_\alpha \cap E_j) = \mu\ (E_j \cap C_\alpha) \tag{8}$$

of the measure $\mu\ (C_\alpha \cap E_j)$ in the events C_α and E_j. Substituting from eqs. (7) into (8) one gets the following constraint equation

$$P\ (E_j \mid C_\alpha)\ \mu\ (C_\alpha) = P\ (C_\alpha \mid E_j)\ \mu\ (E_j) \tag{9}$$

for the conditional probabilities. Substituting eqs. (7) into (5) one gets the expansions

$$\mu\ (E_j) = \sum_{\alpha=1} P\ (E_j \mid C_\alpha)\ \mu\ (C_\alpha) \tag{10.1}$$

$$\mu\ (C_\alpha) = \sum_{j=1} P\ (C_\alpha \mid E_j)\ \mu\ (E_j) \tag{10.2}$$

Next, solving eq. (9) for the conditional probability $P\ (C_\alpha \mid E_j)$ and making use of eq. (10.1) for $\mu(E_j)$ yields the celebrated Bayes's formula

$$P\ (C_\alpha \mid E_j) = \frac{P\ (E_j \mid C_\alpha)\ \mu\ (C_\alpha)}{\sum_\beta P\ (E_j \mid C_\beta)\ \mu(C_\beta)} \tag{11}$$

The usual interpretation of these equations is the following: the event C_α represents a cause (e.g. a hypothesis), E_j an effect (e.g. an experimental observation). The conditional probability $P\ (E_j \mid C_\alpha)$ is then the so–called likelihood, that is, the probability for the occurence of the effect E_j under the assumption C_α. $P\ (C_\alpha \mid E_j)$, on the other hand, is the so–called posterior probability, that is, the probability for the cause C_α to produce a given effect E_j. Read disease for the cause C_α and symptom for the effect E_j and there is a temptation to interprete P

$(C_\alpha | E_j)$ and, hence eq. (11) as a diagnosis. The temptation has, seemingly, proved too strong to resist. Accordingly, eq. (11) constitutes the so -called Bayesian diagnosis.

3. EQUATIONS OF DETAILED BALANCE

It is useful to introduce a new set of notations which will help to understand the purely probabilistic relations in Sect. 2 in terms of dynamics. It will allow also to generalise these relations. First, put

$$G_{\alpha j} := \mu\,(C_\alpha \cap E_j) \tag{12.1}$$

$$G_{j\alpha} := \mu\,(E_j \cap C_\alpha) \tag{12.2}$$

so that eq. (8) becomes the symmetry of the matrix $G_{\alpha j,}$ i.e.

$$G_{\alpha j} = G_{j\alpha} \tag{13}$$

Next, set

$$G_\alpha := \mu(C_\alpha) \tag{14.1}$$

$$G_j := \mu(E_j) \tag{14.2}$$

and

$$G\,(j \,|\, \alpha) := P\,(E_j \,|\, C_\alpha) \tag{15.1}$$

$$G\,(\alpha \,|\, j) := P\,(C_\alpha \,|\, E_j) \tag{15.2}$$

$G\,(\alpha | j)$ and $G\,(j | \alpha)$ are transition probabilities while the G_α (α=1, 2,...) and G_j (j = 1, 2,...) are equilibrium probability distributions for the set of states C_α (α = 1, 2, ..) and E_j (j = 1,2,...) respectively. Eqs. (7) now become

$$G\,(j \,|\, \alpha) = \frac{G_{j\alpha}}{G_\alpha} \tag{16.1}$$

$$G\,(\alpha \,|\, j) = \frac{G_{\alpha j}}{G_j} \tag{16.2}$$

while eqs. (5) become

$$G_j = \sum_\alpha G_{j\alpha} = \sum_\alpha G_{\alpha j} \tag{17.1}$$

$$G_\alpha = \sum_j G_{\alpha j} = \sum_j G_{j\alpha} \tag{17.2}$$

On account of eqs. (16) and (17), the matrices $G(j \mid \alpha)$ and $G(\alpha \mid j)$ will be referred to as the reduced matrices of the symmetric matrix $G_{\alpha j} = G_{j\alpha}$. Note that the reduced matrices themselves are not symmetric unless $G_\alpha = G_j$ for all α, j (α, j = 1, 2;...).

Their most important property is that, from eqs. (16) and (17), they are nomalised to unity, i.e.

$$\sum_\alpha G(\alpha \mid j) = 1 \tag{18.1}$$

$$\sum_j G(j \mid \alpha) = 1 \tag{18.2}$$

Eqs. (6) now read

$$\sum_\alpha G_\alpha = 1 \tag{19.1}$$

$$\sum_j G_j = 1 \tag{19.2}$$

while eqs.(9) and (10) become, respectively,

$$G(j \mid \alpha)\, G_\alpha - G(\alpha \mid j)\, G_j = 0 \tag{20}$$

$$G_j = \sum_{\alpha=1} G(j \mid \alpha)\, G_\alpha \tag{21.1}$$

$$G_\alpha = \sum_{j=1} G(\alpha \mid j)\, G_j \tag{21.2}$$

Eq. (20) is equivalent to eq. (13). It is the so-called equation of detailed balance. The condition of symmetry of any given matrix can always be expressed in this form making use of eqs. (13), (16) and (17).

The reduced matrices will, in this general case, continue to satisfy eqs. (18) but G_α and G_j need not, necessarily satisfy eqs.(19). They satisfy instead, on the basis of eqs. (17), the condition

$$\sum_\alpha G_\alpha = \sum_j G_j \tag{22}$$

Eqs. (19) constitute additional constraints. Eqs. (21) are a pair of reciprocally related equations. Substituting one equation into the other one obtains

$$G_j = F_{jk}\, G_k \tag{23.1}$$

$$G_\alpha = F_{\alpha\beta}\, G_\beta \tag{23.2}$$

where the matrices F_{jk} and $F_{\alpha\beta}$ are defined by

$$F_{jk} := \sum_{\alpha} G(j \mid \alpha)\, G(\alpha \mid k) \tag{24.1}$$

$$F_{\alpha\beta} := \sum_{j} G(\alpha \mid j)\, G(j \mid \beta) \tag{24.2}$$

F_{jk} and $F_{\alpha\beta}$ are normalised like $G(\alpha \mid j)$ and $G(j \mid \alpha)$, i.e.

$$\sum_{j} F_{jk} = \sum_{\alpha} G(\alpha \mid k) = 1 \tag{25.1}$$

$$\sum_{\alpha} F_{\alpha\beta} = \sum_{j} G(j \mid \beta) = 1 \tag{25.2}$$

If, from eqs. (23) - (25), one interpretes F_{jk} and $F_{\alpha\beta}$ as correlations between the states E_j, E_k and C_α, C_β, respectively, of the same species one then has, by definition,

$$F_{jk} := G(j \mid k) \tag{26.1}$$

$$F_{\alpha\beta} := G(\alpha \mid \beta) \tag{26.2}$$

In this case, eqs. (24) become the Chapman-Kolmogorov equations

$$\sum_{\alpha} G(j \mid \alpha)\, G(\alpha \mid k) = G(j \mid k) \tag{27.1}$$

$$\sum_{j} G(\alpha \mid j)\, G(j \mid \beta) = G(\alpha \mid \beta) \tag{27.2}$$

Instead of eqs. (26), it may happen that states of the same species are not correlated, that is

$$F_{jk} = \delta_{jk} \tag{28.1}$$

$$F_{\alpha\beta} = \delta_{\alpha\beta} \tag{28.2}$$

In this case, eqs (24) become,

$$\sum_{\alpha} G(j \mid \alpha)\, G(\alpha \mid k) = \delta_{jk} \tag{29.1}$$

$$\sum_{j} G(\alpha \mid j)\, G(j \mid \beta) = \delta_{\alpha\beta} \tag{29.2}$$

An operator satisfying eq. (29) is said to be involutory. In Statistical Mechanics, the matrices $G_{\alpha j} = G(\alpha \mid j)\, G_j$ and $G_{j\alpha} = G(j \mid \alpha)\, G_\alpha$ are proportional to the squared moduli of the matrix elements of the Hamiltonian H, i.e.

$$G_{\alpha j} := |<\alpha \mid H \mid j>|^2 = <\alpha \mid H \mid j><j \mid H^+ \mid \alpha> \qquad (30.1)$$

$$G_{j\alpha} := |<j \mid H \mid \alpha>|^2 = <j \mid H \mid \alpha><\alpha \mid H^+ \mid j> \qquad (30.2)$$

where H^+ is the Hermitian adjoint of H. The symmetry of the matrix $G_{\alpha j}$ i.e. eq. (13), is thus equivalent to the Hermiticity of H, i.e.

$$H = H^+ \qquad (31)$$

The process of reduction from the matrix $G_{\alpha j}$ to the matrix $G(\alpha \mid j)$, i.e. eqs. (16), corresponds to coarse-graining, that is, summation over groups of microscopic states to define macroscopic states: $G_{\alpha j}$ is microscopic while $G(\alpha \mid j)$ is macroscopic. Unlike $G_{\alpha j}$, $G(\alpha \mid j)$ is not symmetric.

The reduction procedure is usually argued to be a model of the descent from reversible to irreversible dynamics. The model is, of course, not to be taken too seriously. It is a mathematical mimicry of a fundamental physical phenomenon but not a theoretical description of this phenomenon. Zermelo's paradox is no more than the observation that $G(\alpha \mid j)$ does not have the symmetry of the matrix $G_{\alpha j}$.

4. CAUSAL DYNAMICS IN CONFIGURATION SPACE

We introduce in this section configuration spaces of causes and effects and describe how they are dynamically related. Unlike in Sect. 2, causes and effects are not represented as subsets but as points in these spaces. They are, in the terminology of probability theory, atomic events.They are the physical states. By c_j (j = 1, 2...) we shall indicate a cause or an in-state. The set of all in-states c_j (j = 2, 2....) will be indicated by Γ. It is the configuration space. Γ is equipped with a measure μ so that it is also a probability space. A state $c_j \in \Gamma$ is then a random variable.

Analogously, by e_α (α = 1, 2) we indicate an effect or an out-state. The set of these state is denoted by E. It is the configuration space of out-states and is equipped with a measure ν so that it too is a probability space. An out-state $e_\alpha \in$ is a random variable.

The spaces G and E are dynamically related. To see this let

$$\Gamma_\alpha := e_\alpha(\Gamma) \subseteq \Gamma; \; \alpha = 1, 2,.... \qquad (32)$$

be the subset of all causes $c_j \in \Gamma$ capable of producing the effect $e_\alpha \in E$. For two distinct out–states e_α and e_β of E ($e_\alpha \neq e_\beta$), it does not follow that the corresponding subsets $\Gamma_\alpha, \Gamma_\beta \subseteq \Gamma$ are disjoint. The class of subsets in eq. (32) does not therefore, consist of mutually exclusive events.

We shall assume, however, that it is exhaustive, i.e;

$$\Gamma = \bigcup_{\alpha=1} \Gamma_\alpha \tag{33.1}$$

$$\Gamma_\alpha \cap \Gamma_\beta \neq \varnothing; \ \alpha \neq \beta; \ \alpha,\beta = 1, 2, \ldots\ldots \tag{33.2}$$

This means that the configuration space of out-states E is an indexing set for the class of subsets $\Gamma_\alpha \subseteq \Gamma$ ($\alpha = 1, 2..$). We denote this class by

$$B\,(\Gamma \,|\, E) := (\Gamma_\alpha \subseteq \Gamma, \ \Gamma_\alpha := e_\alpha(\Gamma), \ e_\alpha \in E) \tag{34}$$

In the same way, let

$$E_j := c_j\,(E) \subseteq E; \ j = 1, 2 \ldots\ldots \tag{35}$$

stand for the subset of all effects $e_\alpha \in E$ which may be produced by the cause $c_j \in \Gamma$. For two distinct in-states c_j and c_k in Γ ($c_j \neq c_k$), the corresponding subsets E_j, $E_k \subseteq E$ are not necessarily disjoint. The subsets in (35) are therefore not mutually exclusive. We assume however that they are exhaustive so that one has

$$E = \bigcup_j E_j \tag{36.1}$$

$$E_j \cap E_k \neq \varnothing; \ j \neq k; j, k = 1, 2 \tag{36.2}$$

The configuration space Γ of in-states is thus an indexing set for the class of subsets in eq. (35). We define this class by

$$B\,(E \,|\, \Gamma) := (E_j \subseteq E, \ E_j = c_j\,(E), \ c_j \in \Gamma) \tag{37}$$

From the construction of $B\,(\Gamma \,|\, E)$ and $B\,(E \,|\, \Gamma)$, we shall say that the out-state $e_\alpha \in E$ is dual to the event $\Gamma_\alpha \in B\,(\Gamma \,|\, E)$ and that the configuration space E itself is dual to the class $B\,(\Gamma \,|\, E)$ of subsets of Γ which it indexes.

Similarly $c_j \in \Gamma$ is dual to $E_j \in B\,(E \,|\, \Gamma)$ and the space Γ is dual to the class of subsets $B\,(E \,|\, \Gamma)$ of E which it indexes. Since $B\,(\Gamma \,|\, E)$ and $B\,(E \,|\, \Gamma)$ consists of exhaustive but not mutually exclusive subsets, they are over-complete.

Expansions of arbitrary subsets of Γ and E and of functions over the bases $B\,(\Gamma \,|\, E)$ and $B\,(E \,|\, \Gamma)$, similarly to eqs. (3) and (5), become extremely cumbersome. Fortunately we are not interested in this purely probabilistic approach. We are interested in the dynamical or causal functions $e_\alpha \rightarrow \Gamma_\alpha := e_\alpha(\Gamma)$ and $c_j \rightarrow E_j := c_j\,(E)$ and the relationship between them.

The properties of these functions emerge much more clearly in terms of their representations as transformations between the spaces of functions (more precisely, measures) defined on the configuration spaces Γ and E. We consider this problem in the next section.

5. CAUSAL DYNAMICS IN FUNCTION SPACE

In the last section we introduced the following structures: a configuration space D ($\equiv \Gamma$, E), an indexing set D' ($\equiv \Gamma$, E) for an over-complete basis B (D I D') of subsets of D $\neq$ D' in one-to-one correspondence with the points of D'.

The points in D' do not only label the over-complete system of subsets of D, they are also dynamically related to them. This relationship is represented by a function: $q \in D' \rightarrow q(D) \subseteq D$ (i.e. $q(D) \in B(D \mid D')$).

Since we are dealing with probability spaces, this relationship may be, equivalently, specified as an induced relationship between the spaces of real, finite, positive functions (i.e. measures) on D and D'. This allows to avoid the inconvenience of B (D I D') being over-complete.

To begin with, set $D \equiv \Gamma$, $D' \equiv E$ to get B (D I D') = B (Γ I E). Recall that Γ is a probability space with measure μ. The measure μ is a point function on B (Γ I E). We use it to define a point function $\hat{\mu}$ in its index space E by means of the equality.

$$\hat{\mu}\,(e_\alpha) := \mu\,(\Gamma_\alpha) \tag{38}$$

for $e_\alpha \in E$ dual to $\Gamma_\alpha \in B\,(\Gamma \mid E)$.
Using $\hat{\mu}\,(e_\alpha)$, define a measure μ' on E by means of the formula

$$\mu'\,(E_j) := \sum_{e_\alpha \in E_j} \hat{\mu}\,(e_\alpha) \tag{39}$$

Applying the analog of (38) to μ' one defines a point function

$$\bar{\mu} := \hat{\mu}' \tag{40}$$

in Γ by means of the equality

$$\bar{\mu}\,(c_j) := \mu'\,(E_j) \tag{41}$$

for $c_j \in \Gamma$ dual to $E_j \in B\,(E \mid \Gamma)$. Appling the anolog of (39) to $\bar{\mu}$ one defines a new measure μ'' on Γ by

$$\mu''\,(\Gamma_\alpha) := \sum_{c_j \in \Gamma} \bar{\mu}\,(c_j) \tag{42}$$

and a new point function $\hat{\mu}''$ in E (c.f. eq. (38) by

$$\hat{\mu}''\,(e_\alpha) := \mu''\,(\Gamma_\alpha) \tag{43}$$

for $e_\alpha \in E$ dual to $\Gamma_\alpha \in B\,(\Gamma \mid E)$. This chain of ever new functions terminates if

$$\mu = \mu'' \tag{44}$$

equivalently if

$$\hat{\mu} = \hat{\mu}'' \tag{45}$$

Eq. (44) says that the prime operation

$$R : \mu \rightarrow \mu' \tag{46}$$

transforming a measure μ on Γ into a measure μ' on E is an involution, that is,

$$R : \mu' \rightarrow \mu'' = \mu \tag{47}$$

In operator form, one writes this as

$$R^2 = 1 \tag{48}$$

which brings to mind eqs. (29).
Eq. (45) is a representation of the same operation R on the point function $\hat{\mu}$; that is,

$$\begin{aligned} R:&\ \hat{\mu} \rightarrow (\hat{\mu})' \\ R:&\ (\hat{\mu})' \rightarrow (\hat{\mu})'' = \hat{\mu} \end{aligned} \tag{49}$$

Comparing (49) with (45), one gets

$$\hat{\mu} = \hat{\mu}'' = (\hat{\mu})'' \tag{50}$$

Eq. (50) is identically satisfied if the transformation $R : \mu \rightarrow \mu'$ commutes with the transformation

$$K : \mu \rightarrow \hat{\mu} \tag{51}$$

We describe eq. (51) by saying that the measure μ on Γ is conditioned by the point function $\hat{\mu}$ in E. K is the conditioning operator (cf. eq. 38). We express the fact that K and R commute in formulae, as

$$[K, R] = 0 \tag{52}$$

where [K, R] is the commutator of K and R.
Making use of (52) in (40) yields the relations

$$\bar{\mu} = \hat{\mu}' = (\hat{\mu})' \tag{53.1}$$

$$(\bar{\mu})' = \bar{\mu}' = \hat{\mu} \tag{53.2}$$

Carrying out the same procedure as above for the measure ν on E one finds the chain of relations

$$\hat{\nu}(c_j) := \nu(E_j) \tag{54}$$

for $c_j \in \Gamma$ dual to $E_j \in B(E|\Gamma)$.

$$\nu'(\Gamma_\alpha) := \sum_{c_j \in \Gamma_\alpha} \nu(c_j) \tag{55}$$

$$\bar{\nu}(e_\alpha) := \hat{\nu}'(\Gamma_\alpha) \tag{56}$$

where

$$\bar{\nu} := \hat{\nu}' = (\hat{\nu})' \tag{57.1}$$

$$(\bar{\nu})' = \bar{\nu}' = \hat{\nu} \tag{57.2}$$

Finally

$$\nu'' = \nu$$
$$(\hat{\nu})'' = \hat{\nu}'' = \hat{\nu} \tag{58}$$

At this point there is still no relationship between the measures μ on Γ and ν on E. We know only that a given measure μ on Γ is transformed by an involutory operator into a measure μ' on E and that this transformation is subject to eq. (52). Eq. (52) says that the assignment of probabilities to atomic events according to eq. (51) is compatible with the definition of the probabilities for these events induced by the trasformation R (cf. eqs. (38),(41), (54), (56)). Thus if one sets

$$\mu' = \nu \tag{59.1}$$

and by eq.(48)

$$\nu' = \mu \tag{59.2}$$

eqs. (53) and (57) yield

$$\bar{\mu} = \hat{\nu} \tag{60.1}$$
$$\bar{\nu} = \hat{\mu} \tag{60.2}$$

Applying (60.1) to $c_j \in \Gamma$ and (60.2) to $e_\alpha \in E$, one defines the a priori probabilities of these states as

$$p_j := \bar{\mu}(c_j) = \hat{\nu}(c_j) = \nu(E_j) = \mu'(E_j) \tag{61.1}$$

$$p_\alpha := \bar{\nu}(e_\alpha) = \hat{\mu}(e_\alpha) = \mu(\Gamma_\alpha) = \nu'(\Gamma_\alpha) \tag{61.2}$$

Since the systems of subsets B $(\Gamma \mid E)$ and B $(E \mid \Gamma)$ are over–complete the probability distributions p_j (j = 1, 2,...), and p_α (α= 1,2, ...) are not normalised to unity. We can now specify the causal functions $e_\alpha \rightarrow \Gamma_\alpha$ and $c_j \rightarrow E_j$.

We first postulate that each state $c_j \in \Gamma$ (j = 1, 2, ...) and $e_\alpha \in E$ (α = 1,2) is associated with a probability measure φ_j (j = 1,2, ...) and ψ_α (α = 1,2,....), respectively, acting in its configuration space. We know that to each measure φ on Γ is associated a triplet of functions $(\hat{\varphi}_j, \varphi'_j, \bar{\varphi}_j)$ satisfying

$$\hat{\varphi}_j(e_\alpha) = \varphi_j(\Gamma_\alpha) \tag{62.1}$$

$$\varphi'_j(E_k) := \sum_{e_\alpha \in E_k} \hat{\varphi}_j(e_\alpha) = \bar{\varphi}_j(c_k) \tag{62.2}$$

Similarly for each ψ_α on E is associated the triplet $(\hat{\psi}_\alpha, \psi'_\alpha, \bar{\psi}_\alpha)$ satisfying

$$\hat{\psi}_\alpha(c_j) = \psi_\alpha(E_j) \tag{63.1}$$

$$\psi'_\alpha(\Gamma_\beta) := \sum_{c_j \in \Gamma_\beta} \hat{\psi}_\alpha(c_j) = \bar{\psi}_\alpha(e_\beta) \tag{63.2}$$

The postulate, therefore, introduces correlations between the states c_j and c_k in Γ and between e_α and e_β in E as can be seen from eqs. (62.2) and (63.2).

The same correlations exist between the events E_j and E_k and between Γ_α and Γ_β. Using the fact that

$$\hat{\varphi}_j = \overline{\varphi'_j} \tag{64.1}$$

$$\hat{\psi}_\alpha = \overline{\psi'_j} \tag{64.2}$$

(cf. eqs. (53) and (57)), one rewrites eqs. (62.2) and (63.2) as

$$\varphi'_j(E_k) = \sum_{e_\alpha \in E_k} \overline{\varphi'_j}(e_\alpha) \tag{65.1}$$

$$\psi'_\alpha(\Gamma_\beta) = \sum_{C_j \in \Gamma_\beta} \overline{\psi'_\alpha}(c_j) \tag{65.2}$$

Eqs. (65) are expansions of the set functions $\varphi'_j(E_\kappa)$ and $\psi'_\alpha(\Gamma_\beta)$ restricted to the supports of the subsets E_k and Γ_β. We specify the correlations between the states $c_j, c_k \in \Gamma$ (equivalently, between the subsets $E_j, E_k \in$ B $(E \mid \Gamma)$) and between the states $e_\alpha, e_\beta \in E$ (equivalently, between the subsets $\Gamma_\alpha, \Gamma_\beta \in$ B $(\Gamma \mid E)$) by means of the conditions

$$\varphi'_j\,(E_k) = \overline{\varphi}_j\,(c_k) = \delta_{jk} \tag{66.1}$$

$$\psi'_\alpha(\Gamma_\beta) = \overline{\psi}_\alpha\,(e_\beta) = \delta_{\alpha\beta} \tag{66.2}$$

According to eqs. (66), the measures φ'_j and ψ'_α are the indicator functions of the subsets E_j and Γ_α respectively.
Eqs (66) may therefore be rewritten as

$$\varphi'_j\,(E) = \varphi'_j\,(E_j) = 1 \tag{67.1}$$

$$\varphi'_\alpha\,(\Gamma) = \psi'_\alpha\,(\Gamma_\alpha) = 1 \tag{67.2}$$

Now the set of indicator functions of an exhaustive system of subsets is like a complete system of projection operators. This observation motivates then the completeness assumption: the sets of measures φ_j (j=1, 2,....) on Γ and ψ_α (α = 1, 2,....) on E are complete. This means that any given measure f on Γ and u on E can be expanded in terms of the basis functions,φ_j (j= 1, 2,...) and ψ_α (α = 1, 2,.....), respectively, as follows

$$f(\Gamma_\alpha) = \sum_j f_j\,\varphi_j\,(\Gamma_\alpha) \tag{68.1}$$

$$u(E_j) = \sum_\alpha u_\alpha\,\psi_\alpha\,(E_j) \tag{68.2}$$

From eqs. (65), the expansion coefficients f_j (j = 1, 2) and u_α (α = 1, 2,...) are given by

$$f_j := \overline{f}\,(c_j) = f'(E_j) \tag{69.1}$$
$$u_\alpha := \overline{u}\,(e_\alpha) = u'(\Gamma_\alpha) \tag{69.2}$$

Substituting (69) in (68), one finds

$$f(\Gamma_\alpha) = \sum_j f'\,(E_j)\,\varphi_j\,(\Gamma_\alpha) \tag{70.1}$$

$$u(E_j) = \sum_\alpha u'\,(\Gamma_\alpha)\,\psi_\alpha\,(E_j) \tag{70.2}$$

Now setting $f \equiv \psi'_\beta$ and $u \equiv \varphi'_k$ in eqs. (70) and recalling eqs. (66) and the fact that $\psi''_\beta = \psi'_\beta$ and $\varphi''_k = \varphi_k$, one obtains the conditions

$$\sum_j \varphi_j\,(\Gamma_\alpha)\,\psi_\beta\,(E_j) = \delta_{\alpha\beta} \tag{71.1}$$

$$\sum_{\alpha} \psi_\alpha (E_j)\, \varphi_\alpha (\Gamma_\alpha) = \delta_{jk} \tag{71.2}$$

Next put $f \equiv \mu$ in eq. (70.1) and $u \equiv \nu$ in (70.2) and make use of eqs. (59) to get

$$\mu (\Gamma_\alpha) = \sum_{j} \nu (E_j)\, \varphi_j (\Gamma_\alpha) \tag{72.1}$$

$$\nu (E_j) = \sum_{\alpha} \mu (\Gamma_\alpha)\, \psi_\alpha (E_j) \tag{72.2}$$

On account of eqs. (66), eqs. (72.1) and (72.2) are true inverses of each other. The measures μ on Γ and ν on E so related will be referred to as the background measures. According to eqs. (61) they determine the a priori probability distributions p_j (j = 1, 2,....) and p_α (α = 1, 2,....) of the in -and out–states c_j and e_α respectively. Introducing the probabilities defined by the matrices

$$P_{\alpha j} := \varphi_j (\Gamma_\alpha) \tag{73.1}$$

$$\hat{P}_{j\alpha} := \psi_\alpha (E_j) \tag{73.2}$$

and making use of (61), one rewrites eqs. (72) as

$$p_\alpha = \sum_{j} P_{\alpha j}\, p_j \tag{74.1}$$

$$p_j = \sum_{\alpha} \hat{P}_{j\alpha}\, p_\alpha \tag{74.2}$$

and eqs. (71) as

$$\sum_{j} P_{\alpha j}\, \hat{P}_{j\beta} = \delta_{\alpha\beta} \tag{75.1}$$

$$\sum_{\alpha} \hat{P}_{j\alpha}\, P_{\alpha k} = \delta_{jk} \tag{75.2}$$

$P_{j\alpha}$ and $\hat{P}_{j\alpha}$ are the cause - effect correlation functions. They are similar to the conditional probabilities but should not be confused with them.

Eqs. (74) are the dynamical or causal equations of motion. The indices α, j, may include time dependence. In this case, eqs. (74) are evolution equations, with P the forward time (i.e; cause $\rightarrow$ effect) and $\hat{P}$ the backward time (i.e. effect $\rightarrow$ cause) evolution operators. In both stationary and non-stationary regimes, eqs. (74) may be shown to be transformations resulting from changes of basis from one set of states to another. They are the generalisations of eqs.

(10). Unlike the P $(E_j \mid C_\alpha)$ and P $(C_\alpha \mid E_j)$ in eqs. (10), $P_{\alpha j}$ and $\hat{P}_{j\alpha}$ are not the reduced matrices of a symmetric matrix. Secondly, $P_{\alpha j}$ and $\hat{P}_{j\alpha}$ are not normalised as the conditional probabilities.
In eqs. (70), put u = f ' to get the inverse pair

$$f(\Gamma_\alpha) = \sum_j f'(E_j)\, \varphi_j(\Gamma_\alpha) \tag{76.1}$$

$$f'(E_j) = \sum_\alpha f(\Gamma_\alpha)\, \psi_\alpha(E_j) \tag{76.2}$$

Recall that the functions f, f ' are related by

$$R: f \rightarrow f' \tag{77.1}$$

$$R: f' \rightarrow f \tag{77.2}$$

$$f(\Gamma_\alpha) := \langle \psi'_\alpha \mid f \rangle \tag{78.1}$$

$$f'(E_j) := \langle \varphi'_j \mid f' \rangle \tag{78.2}$$

Now write eqs. (77) as

$$\mid f' \rangle = R \mid f \rangle \tag{79.1}$$

$$\mid f \rangle = R \mid f' \rangle \tag{79.2}$$

Substituting these into eqs. (78) gives

$$f(\Gamma_\alpha) = \langle \psi'_\alpha \mid f \rangle = \langle \psi_\alpha \mid R^+ \mid f \rangle \tag{80.1}$$

$$f'(E_j) = \langle \varphi'_j \mid f' \rangle = \langle \varphi_j \mid R^+ R \mid f \rangle \tag{80.2}$$

where R+ is the adjoint of R with respect to the scalar product in eqs. (78). Substituting (80) into (76.1), one gets the completeness relations of the $\mid \varphi_j \rangle$ and $\mid \varphi'_j \rangle$

$$\sum_j \mid \varphi_j \rangle \langle \varphi_j \mid = RR^+ \tag{81.1}$$

$$\sum_j \mid \varphi'_j \rangle \langle \varphi'_j \mid = 1 \tag{81.2}$$

Substituting, on the other hand, (80) into (76.2), one gets the completeness relations of the $\mid$ $>\psi_\alpha>$ and $\mid \psi'_\alpha \rangle$

$$\sum_{\alpha} | \psi_\alpha > < \psi_\alpha | = RR^+ \tag{82.1}$$

$$\sum_{\alpha} | \psi'_\alpha > < \psi'_\alpha | = 1 \tag{82.2}$$

Take the scalar product of $| \psi'_\alpha >$ and $| \psi'_\beta >$ using eqs. (81.1) and of $| \psi_\alpha >$, $| \psi_\beta >$ using eq. (81.2). Making use of eqs. (78) and (71.1), one gets

$$< \psi'_\alpha | \psi'_\beta > = \delta_{\alpha\beta} \tag{83.1}$$

$$< \psi_\alpha | \psi_\beta > = C_{\alpha\beta} \tag{83.2}$$

where $C_{\alpha\beta}$ are the matrix elements

$$C_{\alpha\beta} := \sum_j R_{\alpha j} R_{\beta j} \tag{84.1}$$

$$R_{\alpha j} = R^+_{j\alpha} := < \psi_\alpha | R | \varphi_j > = < \varphi_j | R^+ | \psi_\alpha >. \tag{84.2}$$

Similarly, take the scalar product of $| \varphi'_j >$ with $| \varphi'_k >$ using (82.1) and of $| \varphi_j >$ with $| \varphi_k >$ using (82.2). One finds again

$$< \varphi'_j | \varphi'_k > = \delta_{jk} \tag{85.1}$$

$$< \varphi_j | \varphi_k > = C_{jk} \tag{85.2}$$

where

$$C_{jk} := \sum_\alpha R_{j\alpha} R_{k\alpha} \tag{86.1}$$

$$R_{j\alpha} = R^+_{\alpha j} := < \varphi_j | R | \psi_\alpha > = < \psi_\alpha | R^+ | \varphi_j > \tag{86.2}$$

Note from eqs. (73) and (78) that

$$P_{\alpha j} = \varphi_j (\Gamma_\alpha) = < \psi_\alpha | R^+ | \varphi_j > = < \varphi_j | R | \psi_\alpha > = R_{j\alpha} \tag{87.1}$$

$$\hat{P}_{j\alpha} = \psi_\alpha (E_j) = < \varphi_j | R^+ | \psi_\alpha > = < \psi_\alpha | R | \varphi_j > = R_{\alpha j} \tag{87.2}$$

Making use of these in eqs. (75) one sees that eqs. (75) express the condition $R^2 = 1$.
We next split $R_{\alpha j}$ into its symmetric (G) and anti-symmetric (L) parts, i.e.

$$R = G + L \tag{88.1}$$

$$R^+ = G - L \tag{88.2}$$

equivalently

$$G = \frac{1}{2}(R + R^{+}) \tag{89.1}$$

$$L = \frac{1}{2}(R - R^{+}) \tag{89.2}$$

As discussed in Sect. 3, the symmetric part G is associated with an equation of detailed balance, i.e.

$$G(j \mid \alpha)\, G_\alpha = G(\alpha \mid j)\, G_j \tag{90}$$

where

$$G_\alpha := \sum_j G_{\alpha j} = \sum_j G_{j\alpha} \tag{91.1}$$

$$G_j := \sum_\alpha G_{j\alpha} = \sum_\alpha G_{\alpha j} \tag{91.2}$$

$$G(\alpha \mid j) := \frac{G_{\alpha j}}{G_j} = \frac{G_{j\alpha}}{G_j} \tag{91.3}$$

$$G(j \mid \alpha) := \frac{G_{j\alpha}}{G_\alpha} = \frac{G_{\alpha j}}{G_\alpha} \tag{91.4}$$

From the definitions in eqs. (91) we have the normalisations

$$\sum_j G(j \mid \alpha) = \sum_\alpha G(\alpha \mid j) = 1 \tag{92}$$

and hence from eqs. (90) and (92)

$$G_\alpha = \sum_j G(\alpha \mid j)\, G_j \tag{93.1}$$

$$G_j = \sum_\alpha G(j \mid \alpha)\, G_\alpha \tag{93.2}$$

Solving eq. (90) for $G(j \mid \alpha)$ and making use of eq. (93.1) one gets, as in Sect. 3, Bayes's formula

$$G(j \mid \alpha) = \frac{G(\alpha \mid j)\, G_j}{\sum_k G(\alpha \mid k)\, G_k} \tag{94}$$

From eqs. (92) and (93) one gets the constraint

$$\sum_{\alpha} G_{\alpha} = \sum_{j} G_j \tag{95}$$

From eqs. (87) and (89) one finds that

$$G_{\alpha j} = \frac{1}{2}\left(\psi_{\alpha}(E_j) + \varphi_j(\Gamma_{\alpha})\right) \tag{96.1}$$

$$L_{\alpha j} = \frac{1}{2}\left(\psi_{\alpha}(E_j) - \varphi_j(\Gamma_{\alpha})\right) \tag{96.2}$$

Hence $G_{\alpha j}$ is a probability only if

$$R = R^{+} \tag{97}$$

for then eq. (97) is equivalent to the equality

$$\psi_{\alpha}(E_j) = \varphi_j(\Gamma_{\alpha}) \tag{98}$$

between probabilities. By the same token $L_{\alpha j}$ is not a probability since $R_{\alpha j}$ cannot be totally anti-symmetric.
Inserting (88.1) in the condition $R^2 = 1$ in eq. (48) one gets the conditions

$$G^2 + L^2 = 1 \tag{99.1}$$
$$\{G, L\} := GL + LG = 0 \tag{99.2}$$

Furthermore, one finds from eqs. (88) that

$$[G, L] = \frac{1}{2}[R, R^{+}] \tag{100}$$

where $[A, B] := AB - BA$ is the commutator of A and B. It follows from (99.2) and (100) that if R is, in addition, a normal operator that is, if R and R^{+} commute, then

$$GL = LG = 0 \tag{101}$$

Substituting (96) into (101) and making use of (71), one finds the relations

$$C_{jk} := \sum_{\alpha} \varphi_j(\Gamma_{\alpha})\, \varphi_k(\Gamma_{\alpha}) = \sum_{\alpha} \psi_{\alpha}(E_j)\, \psi_{\alpha}(E_k) \tag{102.1}$$

$$C_{\alpha\beta} := \sum_{j} \psi_{\alpha}(E_j)\, \psi_{\beta}(E_j) = \sum_{j} \varphi_j(\Gamma_{\alpha})\, \varphi_j(\Gamma_{\beta}) \tag{102.2}$$

The matrices C_{jk} and $C_{\alpha\beta}$ are those introduced in eqs. (83) - (86), i.e.

$$C_{jk} = \langle \varphi_j | \varphi_k \rangle \tag{103.1}$$

$$C_{\alpha\beta} = \quad < \psi_\alpha | \psi_\beta > \qquad (103.2)$$

Eq. (101) is satisfied trivially in the particular case in which R is symmetric, i.e. $R = R^+$, for then L = O by eq. (89.2). In this case

$$C = G^2 = 1 \qquad (104)$$

equivalently

$$< \varphi_j | \varphi_k > = < \varphi'_j | \varphi'_k > = \delta_{jk} \qquad (105.1)$$

$$<\psi_\alpha | \psi_\beta > = < \psi'_\alpha | \psi'_\beta > = \delta_{\alpha\beta} \qquad (105.2)$$

Making use of (96) in (102) allows to express C in the case in which eq. (101) holds as

$$C = G^2 - L^2 \qquad (106)$$

Combining this with eq. (99.1) gives

$$G^2 = \frac{1}{2}(1+C) \qquad (107.1)$$

$$L^2 = \frac{1}{2}(1-C) \qquad (107.2)$$

In the symmetric case ($R=R^+$), write eq. (104) with the help of eqs (91) as

$$\sum_j G(\alpha | j)\, G(j | \beta)\, G_j = \frac{1}{G_\alpha} \delta_{\alpha\beta} \qquad (108.1)$$

$$\sum_\alpha G(j | \alpha)\, G(\alpha | k)\, G_\alpha = \frac{1}{G_j} \delta_{jk} \qquad (108.2)$$

One then finds, upon making use of (92) the new relations

$$\sum_j G_j \,.\, G(j | \alpha) = \frac{1}{G_\alpha} \qquad (109.1)$$

$$\sum_\alpha G_\alpha \,.\, G(\alpha | j) = \frac{1}{G_j} \qquad (109.2)$$

Eqs. (109) are to be compared with eqs. (93). Combining (109) and (103) yields

$$\sum_{j,k} G_j \,.\, [G(j | \alpha) \,.\, G(\alpha | k)] \,.\, G_k = 1 \qquad (110.1)$$

$$\sum_{\alpha,\beta} G_\alpha \,.\, [G(\alpha | j) \,.\, G(j | \beta)] \,.\, G_\beta = 1 \qquad (110.2)$$

From eqs. (109) one finds, by making use of eqs. (91.1) and (91.2), the equality

$$\sum_j G_j^2 = \sum_\alpha G_\alpha^2 \tag{111}$$

If the spaces spanned by $|\psi_\alpha>$ (α=1, 2...) and $|\varphi_j>$ (j=1, 2,...) are not finite, then the quadratic expressions in (110) may diverge. Eq. (110) is to be compared with eq. (95). Recall that in the symmetric case G_α and G_j are probabilities (cf eqs. (91.1) and (91.2)).

However they are not the a priori probabilities of causes and effects, respectively. The latter are p_α and p_j defined in eqs. (61). In the symmetric case, the relationships between (G_α, G_j) and (p_α, p_j) are easily found from eqs. (91.1), (91.2) and (74). They are

$$\sum_\alpha p_\alpha G_\alpha = \sum_j p_j \tag{112.1}$$

$$\sum_j p_j G_j = \sum_\alpha p_\alpha \tag{112.2.}$$

whence

$$\sum_\alpha p_\alpha = \sum_j p_j \tag{113.1}$$

$$\sum_\alpha p_\alpha G_\alpha = \sum_j p_j G_j \tag{113.2}$$

If

$$p_\alpha = \lambda\, G_\alpha \tag{114.1}$$

$$p_j = \lambda\, Gj \tag{114.2}$$

then eqs. (112.1) and (112.2) reduce to eqs. (95) and (110), respectively.
Returning to eqs. (110.1) let, for fixed α, G_j (j=1, 2..) and G (j | α) (j=1, 2..) be given data from which to solve (110.1) for G (α | k). Similarly, let, for fixed j, G_α (α=1, 2..) and G (α |j) (α=1, 2...) be given, from which to solve (110.2) for G (j | β).

Let us seek these solutions in the form

$$G(\alpha \mid j) = \sum_{\beta, k} D_{\alpha\beta}\ \bar{D}_{jk}\, G(k \mid \beta) \tag{115.1}$$

$$G(j \mid \alpha) = \sum_{k, \beta} D_{jk}\, \bar{D}_{\alpha\beta}\, G(\beta \mid k) \tag{115.2.}$$

Eqs. (115.1) and (115.2) are consistent if $\bar{D}$ is the inverse of D. In particular, eq. (94) is a special case of eqs. (115) for

$$D_{\alpha\beta} = G_\alpha \delta_{\alpha\beta} \tag{116.1}$$

$$\bar{D}_{jk} = \frac{1}{G_j} \delta_{jk} \tag{116.2}$$

$$D_{jk} = G_j \delta_{jk} \tag{116.3}$$

$$\bar{D}_{\alpha\beta} = \frac{1}{G_\alpha} \delta_{\alpha\beta} \tag{116.4}$$

Eq. (110) is however, much more than eq. (94) since it incorporates the condition $G^2 = 1$ for $G_{\alpha j}=R_{\alpha j}$. Eqs. (115) are solutions of eq. (110) for the given data if

$$\sum_\alpha G_\alpha D_{\alpha\beta} = G_\beta \tag{117.1}$$

$$\sum_\alpha D_{\beta\alpha} \frac{1}{G_\alpha} = G_\beta \tag{117.2}$$

$$\sum_\alpha G_\alpha \bar{D}_{\alpha\beta} = G_\beta \tag{118.1}$$

$$\sum_\alpha \bar{D}_{\beta\alpha} G_\alpha = \frac{1}{G_\beta} \tag{118.2}$$

$$\sum_j G_j D_{jk} = G_k \tag{119.1}$$

$$\sum_j D_{kj} \frac{1}{G_j} = Gk \tag{119.2}$$

and

$$\sum_j G_j \bar{D}_{jk} = G_k \tag{120.1}$$

$$\sum_j \bar{D}_{kj} G_j = \frac{1}{G_k} \tag{120.2}$$

Inserting eqs. (116) in (117) - (120) one finds that for all α and j

$$G_\alpha = G_j = 1 \tag{121}$$

These are consistent with the conditions

$$\sum_j G_{\alpha j}\, G_{j\beta} = \delta_{\alpha\beta} \tag{122.1}$$

$$\sum_\alpha G_{j\alpha}\, G_{\alpha k} = \delta_{jk} \tag{122.2}$$

and the definitions of G_α and G_j in eqs. (91.1) and (91.2)

An operator D, with inverse $\bar{D}$, which transforms G (α | j) into G (j | α) according to eqs. (115) and satisfying eqs. (117) - (120) will be called a diagnosis. The transformation in eqs. (115) will, by extension, be called the diagnosis. Bayes's formula (i.e. eq. (94)) is, as we have pointed out, a special kind of diagnosis obtained from (115) by the position in eqs. (116).

Let G' stand for the operator with matrix elements G (α | j) and G (j | α). Let D^+ and $\bar{D}^+$ be the Hermitian adjoints of D and $\bar{D}$ respectively. Eqs. (115.1) and (115.2) can be combined into the operator equation

$$G' = D.G'^{+}.\bar{D}^{+} \tag{123}$$

or equivalently

$$\bar{D}G' = (\bar{D}G')^{+} \tag{124}$$

thus, the effect of the diagnosis is to transform G' into the symmetric operator

$$G := \bar{D}.G' \tag{125}$$

Eq. (125) is consistent with eqs. (91.3) and (91.4) since $G_\alpha = G_j = 1$ for $G^2 = 1$.
For the general case in which the operator R is not symmetric (i.e. $R_{\alpha j} \neq R_{j\alpha}$), one considers in place of (115) the transformations

$$R_{\alpha j} = \sum_{\beta \cdot k} D_{\alpha\beta}\, \bar{D}_{jk}\, R_{k\beta} \tag{126.1}$$

$$R_{j\alpha} = \sum_{k \cdot \beta} D_{jk}\, \bar{D}_{\alpha\beta}\, R_{\beta k} \tag{126.2}$$

where $\bar{D}$ is the inverse of D. From the condition R^2=1, one finds from eqs. (126) that D and $\bar{D}$ satisfy the constraints

$$D_{\alpha\gamma}\, \bar{D}_{\beta\gamma} = \quad D_{\gamma\alpha}\, \bar{D}_{\gamma\beta} = \delta_{\alpha\beta} \tag{127.1}$$

$$D_{jl}\, \bar{D}_{kl} = D_{lj}\, \bar{D}_{lk} = \delta_{jk} \tag{127.2}$$

Eqs. (127) imply that D is symmetric i.e.,

$$D = D^{+} \tag{128}$$

In terms of $\hat{P}_{j\alpha} = R_{\alpha j}$ and $P_{\alpha j} = R_{j\alpha}$ eqs. (126) become

$$\hat{P}_{j\alpha} = \sum_{\beta,k} D_{\alpha\beta}\ \bar{D}_{jk}\, P_{\beta k} \tag{129.1}$$

$$P_{\alpha j} = \sum_{k,\beta} D_{jk}\, \bar{D}_{\alpha\beta}\, \hat{P}_{k\beta} \tag{129.2}$$

Combining (126.1) and (126.2) we have the operator transformation

$$R = D\, R^{+}\, \bar{D} = D\, (\, \bar{D} R)^{+} \tag{130}$$

or equivalently

$$G = G^{+} \tag{131}$$

where

$$G := \bar{D} R = (\bar{D} R)^{+} \tag{132}$$

Again the effect of a dignosis is to transform R into the symmetric matrix G. Eqs. (129) satisfy eqs. (74) if

$$\sum_{j} p_j\, D_{jk} := q_k \tag{133.1}$$

$$\sum_{\alpha} p_\alpha\, D_{\alpha\beta} := q_\beta \tag{133.2}$$

and

$$\sum_{j} q_j\ \ \hat{P}_{j\alpha} = q_\alpha \tag{134.1}$$

$$\sum_{\alpha} q_\alpha\, P_{\alpha j} = q_j \tag{134.2}$$

According to eqs. (133), a diagnosis D operates a change of basis in the spaces spanned by $|\,\varphi_j>$ and $|\,\psi_\alpha>$, respectively, changing the pair of a priori probabilities (p_j, p_α) into a new pair (q_j, q_α).

Let u be the operator which diagonalises D, i.e.

$$\sum_{l,m} u_{jl}\, D_{lm}\, \bar{u}_{mk} = D_j\, \delta_{jk} \tag{135.1}$$

$$\sum_{\gamma,\delta} u_{\alpha\gamma}\, D_{\gamma\delta}\ \ \bar{u}_{\delta\beta} = D_\alpha\, \delta_{\alpha\beta} \tag{135.2}$$

where D_j, D_α are the eigenualues of D_{jk} and $D_{\alpha\beta}$, respectively. Making use of (135) in (126) one gets the relation

$$G(\alpha \mid j)\, D_j = G(j \mid \alpha)\, D_\alpha \tag{136}$$

where

$$G(\alpha \mid j) := u_{\alpha\beta}\, R_{\beta k}\, \bar{u}_{kj} \tag{137.1}$$

$$G(j \mid \alpha):\ (\bar{u}^{+}_{jk}\, R_{k\beta}\, u^{+}_{\beta\alpha})^{+} \tag{137.2}$$

Eq. (136) is equivalent to eq. (131). It is a different way of expressing the fact that G is symmetric. In fact from eq. (132) one finds

$$u_{\alpha\beta}\, G_{\beta k}\ \ \bar{u}_{kj} = u_{\alpha\beta}\, \bar{D}_{\beta\gamma}\, R_{\gamma k}\, \bar{u}_{kj} = \frac{1}{D_\alpha} G(\alpha \mid j) \tag{138.1}$$

$$u_{\alpha\beta}\, G_{\beta k}\ \ \bar{u}_{kj} = u_{\alpha\beta}\, R^{+}_{\beta\gamma}\, \bar{D}_{\gamma k}\, \bar{u}_{kj} = \frac{1}{D_j} G(j \mid \alpha) \tag{138.2}$$

whence eq. (136). Note from eqs (132) and (133) that

$$\sum_\alpha G_{j\alpha}\, q_\alpha = p_j \tag{139.1}$$

$$\sum_j G_{\alpha j}\, q_j = p_\alpha \tag{139.2}$$

Applying the operator u to both sides of (139) one obtains

$$\sum_\alpha G(j \mid \alpha)\, q'_\alpha = D_j\, p'_j \tag{140.1}$$

$$\sum_j G(\alpha \mid j)\, q'_j = D_\alpha\, p'_\alpha \tag{140.2}$$

where

$$t'_j = \sum_k u_{jk}\, t_k \tag{141.1}$$

$$t'_j = \sum_\beta u_{\alpha\beta}\, t_\beta \tag{141.2}$$

and $t \equiv p, q$. On the other hand, applying u to both sides of (133) yields

$$D_j p'_j = q'_j \quad (142.1)$$

$$D_\alpha p'_\alpha = q'_\alpha \quad (142.2)$$

Substituting from (142) into (140) one gets

$$\sum_\alpha G(j \mid \alpha) q'_\alpha = q'_j \quad (143.1)$$

$$\sum_j G(\alpha \mid j) q'_j = q'_\alpha \quad (143.2)$$

Similarly applying u to both sides of the equations (cf. eqs. (74) and (87))

$$\sum_j p_j R_{j\alpha} = p_\alpha \quad (144.1)$$

$$\sum_\alpha p_\alpha R_{\alpha j} = p_j \quad (144.2)$$

yields

$$\sum_j p'_j G(j \mid \alpha) = p'_\alpha \quad (145.1)$$

$$\sum_\alpha p_\alpha G(\alpha \mid j) = p'_j \quad (145.2)$$

We have made use in eqs. (145) of the fact that

$$uu^+ = u^+u = 1 \quad (146)$$

This follows from eqs. (135) and the symmetry (i.e. $D = D^+$) of D. Making use of (142) in (136) yields

$$p'_\alpha G(\alpha \mid j) q'_j = p'_j G(j \mid \alpha) q'_\alpha \quad (147)$$

whence

$$G(\alpha \mid j) = \frac{p'_j G(j \mid \alpha) q'_\alpha}{q'_j \; p'_\alpha} \quad (148)$$

Inserting for q'_j and p'_α from eqs. (143) and (145) one finally gets

$$G(\alpha \mid j) = \frac{p'_j \, G(j \mid \alpha) \, q'_\alpha}{\sum_k p'_k \, G(k \mid \alpha) \sum_\beta G(j \mid \beta) \, q'_\beta} \tag{149}$$

The structure of eq. (149) is similar to that of Bayes's theorem, however, applied twice: once to q'_j and then to p'_α.

6. SUMMARY

Diagnosis can be described as follows: the a priori probability distributions p_j ($j = 1, 2,...$) and p_α ($\alpha = 1, 2,...$) for effects and causes, respectively, are assumed given together with the correlations (i.e. likelihood) $R_{j\alpha}$ of the effects $j = 1, 2,.....$ given the cause α. The diagnostic problem consists in finding the correlations $R_{\alpha j}$ (i.e. the a posterior probabilities) for the cause α given the data on the a priori probabilities and the likelihood. The solution starts from the positions:

$$\sum_j p_j R_{j\alpha} = p_\alpha \tag{150.1}$$

$$\sum_\alpha p_\alpha R_{\alpha j} = p_j \tag{150.2}$$

and

$$\sum_j R_{\alpha j} R_{j\beta} = \delta_{\alpha\beta} \tag{151.1}$$

$$\sum_\beta R_{j\alpha} R_{\alpha k} = \delta_{jk} \tag{151.2}$$

By definition, a diagnosis is the symmetric operator $D=D^+$ which transforms $R_{\alpha j}$ into $R_{j\alpha}$ according to the formulae

$$R_{\alpha j} = \sum_{\beta,k} D_{\alpha\beta} \bar{D}_{jk} R_{k\beta} \tag{152.1}$$

$$R_{j\alpha} = \sum_{k,\beta} D_{jk} \bar{D}_{\alpha\beta} R_{\beta k} \tag{152.2}$$

or equivalently

$$G: = \bar{D}R = R^+ \bar{D} = G^+ \tag{153}$$

Given D, the unitary operator U ($\bar{U} = U^+$) exists which diagnolises D i.e. such that

$$\sum_{\gamma,\sigma} U_{\alpha\gamma} D_{\gamma\sigma} \bar{U}_{\sigma\beta} = D_\alpha \delta_{\alpha\beta} \tag{154.1}$$

$$\sum_{l,m} U_{jl} D_{lm} \bar{U}_{mk} = D_j \delta_{jk} \tag{154.2}$$

where the D_α ($\alpha = 1, 2,..$) and D_j ($j = 1, 2,..$) are the eigenvalues of D in the spaces of causes and effects, respectively. D transforms the a priori probabilities p_α ($\alpha = 1, 2$ ..) and p_j ($j = 1,2...$) into new ones q_α ($\alpha = 1, 2,...$) and q_j ($j = 1, 2...$) according to the formulae

$$\sum_{k} D_{jk}\, p_k := q_j \tag{155.1}$$

$$\sum_{\beta} D_{\alpha\beta}\, p_\beta := q_\alpha \tag{155.2}$$

The q_j, q_α are a priori probabilities and are trasformed into each other by the matrices $R_{\alpha j}$ and $R_{j\alpha}$ as follows:

$$\sum_{j} R_{\alpha j}\, q_j = q_\alpha \tag{156.1}$$

$$\sum_{\alpha} R_{j\alpha}\, q_\alpha = q_j \tag{156.2}$$

The unitary operator U trasforms $t \equiv p, q$ into a new pair $t' \equiv p', q'$ given by

$$\sum_{k} U_{jk}\, t_k = \sum_{k} t_k\, \bar{U}_{kj} = t'_j \tag{157.1}$$

$$\sum_{\beta} U_{\alpha\beta}\, t_\beta = \sum_{\beta} t_\beta\, \bar{U}_{\beta\alpha} = t'_\alpha \tag{157.2}$$

The new pair $t' \equiv p', q'$ are related to the eigenvalues D_j, D_α of D by

$$D_j = \frac{q'_j}{p'_j} \tag{158.1}$$

$$D_\alpha = \frac{q'_\alpha}{p'_\alpha} \tag{158.2}$$

and are related to each other by

$$\sum_{j} p'_j\, G(j \mid \alpha) = p'_\alpha \tag{159.1}$$

$$\sum_{\alpha} p'_{\alpha} G(\alpha \mid j) = p'_{j} \tag{159.2}$$

and

$$\sum_{j} G(j \mid \alpha)\, q'_{\alpha} = q'_{j} \tag{160.1}$$

y

$$\sum_{\alpha} G(\alpha \mid j)\, q'_{j} = q'_{\alpha} \tag{160.1}$$

where $G(j \mid \alpha)$ is the trasform of G in eq. (153) under U i.e.

$$\frac{1}{D_j} G(j \mid \alpha) := U_{jk} G_{k\beta} \bar{U}_{\beta\alpha} \tag{161.1}$$

$$\frac{1}{D_\alpha} G(\alpha \mid j) := U_{\alpha\beta} G_{\beta k} \bar{U}_{kj} \tag{161.2}$$

whence the equality

$$G(\alpha \mid j)\, D_j = G(j \mid \alpha)\, D_\alpha \tag{162}$$

or equivalently

$$p'_{\alpha} G(\alpha \mid j)\, q'_{j} = p'_{j}\, G(j \mid \alpha)\, q'_{j} \tag{163}$$

upon making use of eqs. (158).

Solving (163) for $G(\alpha \mid j)$ yields

$$G(\alpha \mid j) = \frac{p'_{j} G(j \mid \alpha)\, q'_{\alpha}}{q'_{j}\; p'_{\alpha}} \tag{164}$$

This becomes, on making use of eqs. (159.1) for p'_{α} and (160.1) for q'_{j},

$$G(\alpha \mid j) = \frac{p'_{j} G(j \mid \alpha)\, q'_{\alpha}}{\left(\sum_{k} p'_{k} G(k \mid \alpha)\right)\cdot\left(\sum_{\beta} G(j \mid \beta) q'_{\beta}\right)} \tag{165}$$

Eq. (165) looks like Bayes's formula of second order since two pairs of a priori probabilities (p'_j, p'_α) and (q'_j, q'_α) are involved. Eq. (165) is the solution of the diagnostic problem. It is not in the canonical form of Bayes's theorem. Bayes's theorem is, however, a special case of this solution. To see this, rewrite eq. (153) as

$$G = \frac{1}{2}(\bar{D}R + R^{+}\bar{D}) \tag{166}$$

hence for $D = \bar{D} = 1$, eq. (166) becomes

$$G = \frac{1}{2}(R+R^{+}) \tag{167}$$

G is in this case the symmetric part of R. But for $D = \bar{D} = 1$, eqs. (152) say that $R = R^{+}$ so that this case corresponds to Bayes's theorem.