Fishing for the Unknown Unknowns:
A Bayesian Perspective

Robert K. Niven [a]

[a]*School of Engineering and Information Technology, The University of New South Wales,
Canberra, ACT, 2600, Australia.*
Email: r.niven@adfa.edu.au

Abstract: In this study, we consider the mathematical framework for *inductive* or *plausible reasoning*, or in other words, mathematically rigorous reasoning with uncertainty [1-3], which extends the well-known framework for *logical* or *deductive reasoning* (often simply called *logic*), based on certainty. This framework makes use of the Bayesian interpretation of probabilities as *plausibilities*, or mathematical parameters assigned a value from 0 to 1, based on what is known. Importantly, these need not correspond to measurable frequencies, for example from the collection of an infinite number of samples from a fixed population. Probabilities identified as plausibilities have been shown to satisfy the same mathematical rules as probabilities defined as measurable frequencies, based on the product and sum rules of probability theory [1-3]. However, they are of much broader application. In particular, Bayes' rule [4-5] enables a user to calculate the "probability of an hypothesis, given the data", or in other words the degree of belief or disbelief in the hypothesis, given what is known. Since its mathematical foundation is formulated to be universal in application, without exception, this provides a common platform for hypothesis testing, model comparison, parameter estimation or any other form of inference, and can be used in subsequent decision frameworks such as utility / loss analysis or risk assessment. In consequence, no construct other than that of probability is necessary for the analysis of uncertainty.

We then examine the "Rumsfeld trifecta" of *known knowns, known unknowns* and *unknown unknowns* from the Bayesian perspective. By symmetry, it is shown that all such influences must be handled identically mathematically, regardless of whether they are known or unknown, or even if their existence is known. In consequence, if any unknown unknowns have an influence on the hypothesis or model space of a given problem, their influence can be calculated mathematically from their corresponding influence on the data. Using this insight, we derive a form of Bayes' rule which directly gives the probability of the unknown unknowns, based on the data. This enables a user to identify the existence of unknown unknowns directly from the data.

We then examine an extended set of influences composed of the Rumsfeld trifecta and also the *unknown knowns*, i.e., those influences which are known but are accidentally or deliberately ignored. For many prob-lems, including military problems, these can arguably be of greater importance than the unknown unknowns. We show that the above mathematical insights also apply within this enlarged four-fold classification.

[1] Polya, G. (1954) Mathematics and Plausible Reasoning, Vol II, Patterns of Plausible Inference, Princeton U.P., NJ; [2] Cox, R.T. (1961) The Algebra of Probable Inference, John Hopkins Press, NJ; [3] Jaynes, E.T. (Bretthorst, G.L., ed.) (2003) Probability Theory: The Logic of Science, Cambridge U.P., Cambridge UK; [4] Bayes, T. (presented by Price, R.) (1763) An essay towards solving a problem in the doctrine of chance, Phil. Trans. Royal Soc. London 53: 370-418; [5] Laplace, P. (1774) Mémoire sur la probabilité des causes par les èvènements, *l'Académie Royale des Sciences*, **6**, 621-656.

1 INTRODUCTION

In their 2017 call for a strategic research investment in Modelling Complex Warfighting, the Australian Defence Science and Technology (DST) group identified their interest in "approaches for eliciting, understanding, modelling and managing unknowns" (DST, 2017a). This culminated in the DST Strategic Response 7 on Modelling Unknowns (DST, 2017b). Philosophically, this theme harks back to the famous statement of former U.S. Defense Secretary Donald Rumsfeld:

> *There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. These are things we don't know we don't know* (Rumsfeld, 2002).

Although ridiculed by some, this sentiment provides a precise and accurate statement of the nature of uncertainties, with deep roots in the philosophy of personal awareness (Luft & Ingham, 1955). Such terminology has reportedly been in common usage at NASA and in project management circles for decades (Wikipedia, 2019), and also evokes and enlarges upon the old military concept of the "fog of war" (Hale, 1896).

To the three categories mentioned, some philosophers have added a fourth, the *unknown known*, or that which is known but which we intentionally refuse to acknowledge (e.g. Zizek, 2004; Daase & Kessler, 2007).

The foregoing classifications of influences are also related to the analysis of *latent* or *hidden variables*, which are not directly observed, but can be inferred by their influence on other observables (Bishop, 2006). This includes analyses of model identifiability (Christensen et al., 2011), mixture models (Bishop, 2006, chap. 9), probabilistic principal component analysis (Bishop, 2006, chap. 12), and variational Bayesian methods (Barber, 2012).

In this study, we examine the above ideas using the mathematical framework provided by Bayesian inference.

2 BAYESIAN PROBABILITIES AND BAYESIAN INFERENCE

We start by adopting the broader Bayesian interpretation of probabilities as "plausibilities", or mathematical parameters assigned a value from 0 to 1, based on what is known (Bayes, 1763; Laplace, 1774). In this perspective, probabilities need not correspond to measurable frequencies, although this can be the case if intended for a given problem. Probabilities identified as plausibilities have been shown to satisfy the same mathematical framework as probabilities defined as measurable frequencies (Polya, 1954; Cox, 1961; Tribus, 1969; Jaynes, 2003), based on the product and sum rules of probability theory. However, they are of much broader application. In consequence, as formulated by Polya, Cox, Tribus and Jaynes, they can be used to underpin a mathematical framework for *inductive* or *plausible reasoning*, or in other words, mathematically rigorous reasoning with uncertainty, as distinct from the well-known framework of *logical* or *deductive reasoning*, which is based on certainty.

In the Bayesian perspective, it is eminently reasonable to consider the probability of a particular person being elected to public office, the probability of a stock price at a certain time, the probability of a calamity or war, or in general the probability of a proposition, given what is known. By approaching the topic in this way, the goal of its originators was to establish a theoretical framework which is *universal* in application: i.e., which applies to the analysis of ALL problems involving uncertainty, without exception, and so is interconvertible between truth value or believability or propensity or predictability or degree of knowledge or credibility or faith or confidence or conviction or trust or bettability or utility – or conversely, of unpredictability or indeterminacy or indefiniteness or ignorance or reservation or concern or distrustfulness or skepticism or suspicion or doubt or disbelief or incredulity or fuzziness – or any other terminology or purpose that a particular user might have in mind[1]. In consequence, no construct other than probability is needed for the analysis of any such tendencies. For a Bayesian, the fundamental question is not how to represent the degree of certainty or uncertainty, or any related tendency, but how to calculate its value under different conditions.

For further detailed discussions of the Bayesian viewpoint, the reader is directed to major works, especially by Tribus (1969) and Jaynes (2003). Tutorial notes presented by the author to a MODSIM 2017 workshop are also available (Niven, 2017).

In the following sections we examine several insights into this topic.

---

[1]My sincere thanks to John Skilling for prompting this phrasing!

2.1 Discrete and Continuous Probabilities

We quickly introduce some nomenclature from probability theory:

(a) For a discrete variable[2] M with discrete values m drawn from the domain $\Omega_m \subset \mathbb{Z}$, the probability (or probability mass function, pmf) is defined by:

$$p(m) = \text{Prob}(M = m) \tag{1}$$

(b) For a discrete multivariate variable $\boldsymbol{M} = [M_1, ..., M_s]$ with values $\boldsymbol{m} = [m_1, ..., m_s]$ drawn from the discrete multivariate domain $\Omega_{\boldsymbol{m}} \subset \mathbb{Z}^s$, the joint probability (joint pmf) becomes:

$$p(\boldsymbol{m}) = \text{Prob}(\boldsymbol{M} = \boldsymbol{m}) = \text{Prob}(M_1 = m_1, ..., M_s = m_s) = \text{Prob}(M_1 = m_1 \cap ... \cap M_s = m_s) \tag{2}$$

where the commas indicate a logical *and* operation, designated formally by $\cap$. Often, the commas representing logical *and* statements are dropped.

(c) For a continuous variable X with continuous values x drawn from the domain $\Omega_x \subseteq \mathbb{R}$, the probability density function (pdf) $p(x)$ can be defined by the probability over the interval (Feller, 1966):

$$\text{Prob}(a < X \leq b) = \int_a^b p(x) dx \tag{3}$$

where dx represents an infinitesimal interval in x.

(d) For a continuous multivariate variable $\boldsymbol{X} = [X_1, ..., X_s]$ with values $\boldsymbol{x} = [x_1, ..., x_s]$ drawn from the multivariate domain $\Omega_{\boldsymbol{x}} \subseteq \mathbb{R}^s$, the joint pdf $p(\boldsymbol{x})$ can be defined by the probability over the hyperinterval:

$$\text{Prob}(\boldsymbol{a} < \boldsymbol{X} \leq \boldsymbol{b}) = \text{Prob}(a_1 < X_1 \leq b_1, ..., a_s < X_s \leq b_s) = \int_{a_1}^{b_1} ... \int_{a_s}^{b_s} p(x_1, ..., x_s) d\boldsymbol{x} \tag{4}$$

using $\boldsymbol{a} = [a_1, ..., a_s]$, $\boldsymbol{b} = [b_1, ..., b_s]$, $p(x_1, ..., x_s) = p(x_1 \cap ... \cap x_s)$ and $d\boldsymbol{x} = dx_1 dx_2 ... dx_s$.

However defined, a probability must satisfy normalisation, respectively for discrete or continuous parameters:

$$\sum_{\Omega_{m_1}} ... \sum_{\Omega_{m_N}} p(\boldsymbol{m}) = \sum_{\Omega_{\boldsymbol{m}}} p(\boldsymbol{m}) = 1 \quad \text{or} \quad \int_{\Omega_{x_1}} ... \int_{\Omega_{x_N}} p(\boldsymbol{x}) d\boldsymbol{x} = \int_{\Omega_{\boldsymbol{x}}} p(\boldsymbol{x}) d\boldsymbol{x} = 1 \tag{5}$$

While probabilities must take a value from 0 to 1, we see from (5) that a pdf $p(x)$ or $p(\boldsymbol{x})$ can take any non-negative real value, so long as normalisation is satisfied. The non-negativity ensures that no portion of the integral in (5) is negative. For a joint composite pmf-pdf, it is convenient to represent normalisation using the sum-integration operator (Rioul, 2008):

$$\oint_{\Omega_{x_1}} ... \oint_{\Omega_{x_N}} p(x_1, x_2, ..., x_N) = \oint_{\Omega_{\boldsymbol{x}}} p(\boldsymbol{x}) = 1 \tag{6}$$

in which all differentials, where they exist, are implied.

Furthermore, a joint pmf or pdf can be reduced in one or several of its arguments by the process of *marginalisation*, respectively by summation or integration:

$$p(m_1) = \sum_{\Omega_{m_2}} ... \sum_{\Omega_{m_N}} p(m_1, m_2, ..., m_N) \quad \text{or} \quad p(x_1) = \int_{\Omega_{x_2}} dx_2 ... \int_{\Omega_{x_N}} dx_N \, p(x_1, x_2, ..., x_N) \tag{7}$$

In the hybrid pmf-pdf notation, this gives:

$$p(x_1) = \oint_{\Omega_{x_N}} ... \oint_{\Omega_{x_N}} p(x_1, x_2, ..., x_N) \tag{8}$$

Finally, a conditional pmf or pdf can be expressed by a probabilistic division operation, in summary form:

$$p(x_1 | x_2, ..., x_N) = \frac{p(x_1, x_2, ..., x_N)}{p(x_2, ..., x_N)} = \frac{p(x_1, x_2, ..., x_N)}{\oint_{\Omega_{x_1}} p(x_1, x_2, ..., x_N)} \tag{9}$$

---

[2]In traditional probability theory, the term "random variable" is used to describe a variable in which there is uncertainty. A Bayesian does not need such terminology: every variable is uncertain.

2.2 Bayes' Rule

We now consider Bayes' rule, one of the most profound developments in probability theory (Bayes, 1763; Laplace, 1774). If we consider the joint probability $p(AB)$ of two propositions A and B, which need not be independent, this can be expanded using the product rule of probability theory in two different ways:

$$p(AB) = p(A)\,p(B|A) = p(B)\,p(A|B) \tag{10}$$

This can be rearranged as:

$$p(B|A) = \frac{p(B)\,p(A|B)}{p(A)} \tag{11}$$

Now if we assign $A = D$ to represent observable data, and $B = H$ to represent an hypothesis or model, we obtain the following relation for the *probability of an hypothesis H, given the data D*, generally referred to as "Bayes' rule" (Bayes, 1763; Laplace, 1774):

$$p(H|D) = \frac{p(H)\,p(D|H)}{p(D)} \tag{12}$$

The standard terminology for each term in this expression is:

$p(H|D)$ = *posterior probability* of the hypothesis H given the data D
$p(H)$ = *prior probability* of the hypothesis H
$p(D|H)$ = probability (*likelihood*) of the data D given the hypothesis H
$p(D)$ = probability (*evidence*) for the data D

Of these probabilities, the prior $p(H)$ contains all that is known about the hypothesis H, including any theoretical foundation, as well as any other information such as the symmetries or geometrical relations inherent in the problem examined (Jaynes, 2003). Commonly, the prior is allocated by the user based on this physical insight, whereas the other terms on the right are extracted from the data. Bayesian inference can thus be viewed as a method to update the prior using the data, expressed by the likelihood-evidence ratio, to give the posterior.

Formally, all probabilities in (12) should in fact be conditioned on the background information I:

$$p(H|DI) = \frac{p(H|I)\,p(D|HI)}{p(D|I)} \tag{13}$$

The background information represents all that is known to a given problem, and must be the same for all probabilities in (13).

Philosophically, the posterior $p(H|D)$ (or formally $p(H|DI)$) is not a measurable frequency, but expresses the observer's *degree of belief* in H, given D. The profound meaning of Bayes' rule (12) is that it provides a rigorous mathematical equation with which to make this assignment. Further discussions of these philosophical ideas are given by many authors including Tribus (1969) and Jaynes (2003). In contrast to orthodox statistical methods such as hypothesis testing, goodness-of-fit tests, maximum likelihood methods or estimation theory, Bayes' rule (12) provides a universal and mathematically rigorous method for *all* forms of inference or estimation in *all* situations, since it is based on the fundamental and universal concept of probabilities. For this reason, it is rapidly becoming the dominant paradigm for inference and estimation across many disciplines, including in astronomy, genetics, medicine, image analysis, political polling and finance.

Furthermore, if different sets of data $\{D_1, D_2, ...\}$ are received in succession, Bayesian updating by Bayes' rule (12) can be conducted iteratively, using the posterior from each iteration as the prior for the next, to calculate $p(H|D_1, D_2, ...)$. For a countable set Ω_H of mutually exclusive hypotheses H_i, the evidence can be further expanded to give $p(D) = \sum_{\Omega_H} p(H_i)\,p(D|H_i)$, giving the extended Bayes rule:

$$p(H_i|D) = \frac{p(H_i)\,p(D|H_i)}{p(D)} = \frac{p(H_i)\,p(D|H_i)}{\sum\limits_{\Omega_H} p(H_i)\,p(D|H_i)} \tag{14}$$

For continuous multivariate data $\boldsymbol{x}$ and a continuum of models or hypotheses, represented by the multivariate parameter $\boldsymbol{\theta}$, Bayes' rule can be rewritten in the continuous form:

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})}{\int\limits_{\Omega_{\boldsymbol{\theta}}} d\boldsymbol{\theta}\,p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})} \tag{15}$$

where $\Omega_{\boldsymbol{\theta}}$ is the domain of $\boldsymbol{\theta}$. Eqs (14)-(15) can be united into the summary form:

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})}{p(\boldsymbol{x})} = \frac{p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})}{\oiint_{\Omega_{\boldsymbol{\theta}}} p(\boldsymbol{\theta})\,p(\boldsymbol{x}|\boldsymbol{\theta})} \tag{16}$$

3   Known Knowns, Known Unknowns and Unknown Unknowns

We now examine the "Rumsfeld trifecta" from the perspective of Bayesian probabilities. Using the following assignments of multivariate parameters or influences[3]:

$$
\begin{aligned}
\boldsymbol{x} &= [x_1, ..., x_s] &&= \text{observable parameters of interest} \\
\boldsymbol{K} &= [K_1, ..., K_M] &&= \text{"known knowns"} \\
\boldsymbol{U} &= [U_1, ..., U_N] &&= \text{"known unknowns"} \\
\boldsymbol{W} &= [W_1, ..., W_O] &&= \text{"unknown unknowns"}
\end{aligned}
\tag{17}
$$

where the counts s, M, N and O are finite. Here the known knowns and known unknowns are those variables formally considered by the observer, while the unknown unknowns consist of those which are omitted from their analysis. We then consider the conditional probability $p(\boldsymbol{x}|\boldsymbol{KUW})$, i.e., the probability of $\boldsymbol{x}$ given all influences, comprising the Rumsfeld trifecta of $\boldsymbol{K}, \boldsymbol{U}$ and $\boldsymbol{W}$. By definition (9):

$$p(\boldsymbol{x}|\boldsymbol{KUW}) = \frac{p(\boldsymbol{xKUW})}{p(\boldsymbol{KUW})} \tag{18}$$

However, a different observer might make a different partition of the known knowns and known unknowns, for example:

$$
\begin{aligned}
\hat{\boldsymbol{K}} &= [K_1, ..., K_{M-1}] &&= \text{"known knowns"} \\
\hat{\boldsymbol{U}} &= [K_M, U_1, ..., U_N, W_1] &&= \text{"known unknowns"}
\end{aligned}
\tag{19}
$$

In consequence the unknown unknowns, or those variables omitted by the second observer, consist of the set:

$$
\hat{\boldsymbol{W}} = [W_2, ..., W_O] = \text{"unknown unknowns"}
\tag{20}
$$

Provided the choices of $\boldsymbol{K}, \boldsymbol{U}$ and $\boldsymbol{W}$ span all possible influences on the data, and are countably finite, we see that $\hat{\boldsymbol{K}}\hat{\boldsymbol{U}}\hat{\boldsymbol{W}} = \boldsymbol{KUW}$ regardless of the partition chosen. Hence by definition:

$$p(\boldsymbol{x}|\hat{\boldsymbol{K}}\hat{\boldsymbol{U}}\hat{\boldsymbol{W}}) = \frac{p(\boldsymbol{x}\hat{\boldsymbol{K}}\hat{\boldsymbol{U}}\hat{\boldsymbol{W}})}{p(\hat{\boldsymbol{K}}\hat{\boldsymbol{U}}\hat{\boldsymbol{W}})} = \frac{p(\boldsymbol{xKUW})}{p(\boldsymbol{KUW})} = p(\boldsymbol{x}|\boldsymbol{KUW}) \tag{21}$$

In consequence, we find that in the probabilistic representation of uncertainty, the mathematical handling of the known knowns, known unknowns and unknown unknowns *must be identical by their symmetry*, regardless of the category to which any particular parameter is assigned (or even whether its existence is known).

Secondly, if we make the distinction between data $\boldsymbol{x}$ and models $\boldsymbol{\theta}$ (which could be discrete or continuous), we can rewrite Bayes' rule subject to the Rumsfeld trifecta:

$$p(\boldsymbol{\theta}|\boldsymbol{xKUW}) = \frac{p(\boldsymbol{\theta}|\boldsymbol{KUW})\,p(\boldsymbol{x}|\boldsymbol{\theta KUW})}{p(\boldsymbol{x}|\boldsymbol{KUW})} = \frac{p(\boldsymbol{\theta}|\boldsymbol{KUW})\,p(\boldsymbol{x}|\boldsymbol{\theta KUW})}{\oiint_{\Omega_{\boldsymbol{\theta}}} p(\boldsymbol{\theta}|\boldsymbol{KUW})\,p(\boldsymbol{x}|\boldsymbol{\theta KUW})} \tag{22}$$

We see from (22) that all influences – including the unknown unknowns $\boldsymbol{W}$ – are automatically included in Bayesian inference, even if one does not know that they exist! In contrast, there could be additional variables which have no influence on the data being considered. Formally:

$$p(\boldsymbol{\theta}|\boldsymbol{xKUWN}) = p(\boldsymbol{\theta}|\boldsymbol{xKUW}) \tag{23}$$

for all non-influencing parameters $\boldsymbol{N}$. From eqs. (22)-(23), if any variable has an influence on the model or workings $\boldsymbol{\theta}$ of the problem, as displayed by the posterior $p(\boldsymbol{\theta}|\boldsymbol{xKUW})$, then it must have some corresponding influence on the data $\boldsymbol{x}$, as expressed in the likelihood $p(\boldsymbol{x}|\boldsymbol{\theta KUW})$ and the evidence $p(\boldsymbol{x}|\boldsymbol{KUW})$.

---

[3]The choice of the last symbol becomes obvious from its name "double U". There is an interesting paradox here: we do not know the unknown unknowns, yet we can still acknowledge and account mathematically for their existence.

We can indeed take this insight further: while Bayes' rule (12) is commonly written to connect hypotheses and data, its forebear (11) can be used to connect the probabilities of any two propositions. We can therefore directly connect the data x and unknown unknowns W by the relation:

$$p(W|xKU) = \frac{p(W|KU)p(x|KUW)}{p(x|KU)}, \quad \text{or in ratio form} \quad \frac{p(W|xKU)}{p(W|KU)} = \frac{p(x|KUW)}{p(x|KU)} \quad (24)$$

Scrutinising this carefully, we first note that the prior $p(W|KU)$ should strictly be written as $p(W|KUI)$, i.e., it gives the prior probability of the unknown unknown(s) given all that is known of the problem, including the known knowns, known unknowns and background information. Even if W is independent of K and U, this is unlikely to be zero.

Eq. (24) is therefore an extremely important equation, since it provides a means to mathematically evaluate the probability of all unknown unknowns, based on the data. To evaluate (24), the user needs the prior $p(W|KU)$, the probability (likelihood) of the data given all possible influences $p(x|KUW)$, for example from recent data, and also the probability of the data given only the known influences $p(x|KU)$, for example from a controlled experiment or comparable historical data set. The ratio form of (24) can be interpreted as the "*unknown unknown amplification factor*", i.e., the extent to which the data reveal the influence of one or more unknown unknowns, beyond the multiplicative threshold defined by the prior $p(W|KU)$.

4   KNOWN KNOWNS, KNOWN UNKNOWNS, UNKNOWN UNKNOWNS AND UNKNOWN KNOWNS

To the "Rumsfeld trifecta", we now add the "unknown knowns", or those influences which are known but are accidentally or deliberately ignored. As commented by several authors (e.g. Zizek, 2004; Daase & Kessler, 2007), these may surpass the unknown unknowns in their importance.

Denoting the unknown knowns as G, we find that the above relations are replicated in this enlarged set of influences. Firstly, by symmetry, for different partitions made by two different observers:

$$p(x|\hat{K}\hat{U}\hat{W}\hat{G}) = \frac{p(x\hat{K}\hat{U}\hat{W}\hat{G})}{p(\hat{K}\hat{U}\hat{W}\hat{G})} = \frac{p(xKUWG)}{p(KUWG)} = p(x|KUWG) \quad (25)$$

Secondly, Bayes' rule becomes:

$$p(\theta|xKUWG) = \frac{p(\theta|KUWG)\,p(x|\theta KUWG)}{p(x|KUWG)} = \frac{p(\theta|KUWG)\,p(x|\theta KUWG)}{\oint_{\Omega_{\theta}} p(\theta|KUWG)\,p(x|\theta KUWG)} \quad (26)$$

Thirdly, we can write down a combined amplification factor:

$$\frac{p(WG|xKU)}{p(WG|KU)} = \frac{p(x|KUWG)}{p(x|KU)} \quad (27)$$

The latter suggests that by careful analysis of data, a user can identify the collective effect of the unknown unknowns and unknown knowns. By successive isolation or recursive experiments, for example using:

$$\frac{p(\omega|xKU)}{p(\omega|KU)} = \frac{p(x|KU\omega)}{p(x|KU)}, \quad \text{for } \omega \in WG \quad (28)$$

the user may then be able to calculate the effect of individual unknown influences within the set WG.

5   CONCLUSIONS

In this study, we consider the mathematical framework for *inductive* or *plausible reasoning*, which extends the well-known framework for *logical* or *deductive reasoning*, based on certainty. This framework adopts the Bayesian interpretation of probabilities as *plausibilities*, or mathematical parameters assigned a value from 0 to 1, based on what is known. These can be shown to satisfy the same rules as probabilities defined as measurable frequencies, and so can be calculated or manipulated mathematically (Polya, 1954; Cox, 1961; Tribus, 1969; Jaynes, 2003). In particular, Bayes' rule (12)-(16) enables a user to calculate the "probability of an hypothesis, given the data", or in other words the degree of belief in the hypothesis, given what is known.

We then examine the "Rumsfeld trifecta" of *known knowns, known unknowns* and *unknown unknowns* from the Bayesian perspective. By symmetry, it is shown that all such influences must be handled identically mathematically, regardless of whether they are known or unknown, or even if their existence is known. In consequence,

if any unknown unknowns have an influence on the hypothesis or model space of a given problem, their influence can be calculated mathematically from their corresponding influence on the data. Using this insight, we derive a form of Bayes' rule which directly gives the probability of the unknown unknowns, based on the data. This enables a user to identify the existence of unknown unknowns directly from the data.

We then examine an extended set of influences composed of the Rumsfeld trifecta and also the *unknown knowns*, i.e., those influences which are known but are accidentally or deliberately ignored. For many problems, including military problems, these can be of greater importance than the unknown unknowns. We show that the above mathematical insights also apply within this enlarged four-fold classification.

We conclude with the following quote from Confucius (c.551-479 BC):

> *The Master said: "Yu, shall I teach thee what is wisdom? To know what we know, and know what we do not know, is wisdom."*

REFERENCES

Barber, D. (2012), Bayesian Reasoning and Machine Learning, Cambridge U.P., UK.

Bayes, T. (presented by Price, R.) (1763). An essay towards solving a problem in the doctrine of chance, Phil. Trans. Royal Soc. London 53: 370-418.

Bishop, C.M. (2006), Pattern Recognition and Machine Learning, Springer, NY.

Christensen, R., Johnson, W., Branscum, A., Hanson, T.E. (2011), Bayesian Ideas and Data Analysis, CRC Press, Boca Raton, pp 94-97.

Confucius, c.551-479 BC, The Analects, II: 17.

Cox, R.T. (1961). The Algebra of Probable Inference, John Hopkins Press, NJ.

Daase, C. & Kessler, O. (2007). Knowns and Unknowns in the War on Terror: Uncertainty and the Political Construction of Danger, Security Dialogue, 38 (4): 411-434.

DST (2017a). Modelling Complex Warfighting Strategic Research Investment, Defence Science and Technology, Department of Defence, Canberra, Australia, 30 October 2017.

DST (2017b). Outcomes from MCW SRI Symposium, Defence Science and Technology, Department of Defence, Canberra, Australia, 14-15 December 2017.

Feller, W. (1966), An Introduction to Probability Theory and Its Applications, Volume II, John Wiley and Sons, Inc., NY.

Hale, L. (1896). The Fog of War, Aldershot Military Academy, UK.

Jaynes, E.T. (Bretthorst, G.L., ed.) (2003). Probability Theory: The Logic of Science, Cambridge U.P., Cambridge UK.

Laplace, P. (1774). Mémoire sur la probabilité des causes par les évènements, *l'Académie Royale des Sciences*, **6**, 621-656.

Luft, J. & Ingham, H. (1955). The Johari window, a graphic model of interpersonal awareness, Proc. Western Training Laboratory in Group Development, University of California, Los Angeles, USA.

Niven, R.K. (2017). Workshop, Probabilistic Inference using Maximum Entropy Analysis and Bayesian Inference (Theory and Applications), 22nd International Congress on Modelling and Simulation (MODSIM2017), Hobart, Tasmania, Australia, 6 December 2017, available on ResearchGate.

Polya, G. (1954). Mathematics and Plausible Reasoning, Vol II, Patterns of Plausible Inference, Princeton U.P., NJ.

Rioul, O. (2008). Théorie des Probabilités, Lavoisier, Paris (French).

Rumsfeld, D. (2002). News briefing, Washington DC, USA, 12 February 2002.

Tribus, M. (1969). Rational Descriptions, Decisions and Designs, Permagon Press, NY.

Wikipedia (2019). There are known knowns, https://en.wikipedia.org/wiki/There_are_known_knowns.

Zizek, S. (2004). What Rumsfeld doesn't know that he knows about Abu Ghraib, https://www.lacan.com/zizekrumsfeld.htm