Modelling the information seeking user by the decisions they make

Peter Bruza
Queensland University of Technology
Brisbane, Australia
p.bruza@qut.edu.au

Guido Zuccon
CSIRO Australian e-Health Research Centre
Brisbane, Australia
Guido.Zuccon@csiro.au

Laurianne Sitbon
Queensland University of Technology
Brisbane, Australia
laurianne.sitbon@qut.edu.au

ABSTRACT

The article focuses on how the information seeker makes decisions about relevance. It will employ a novel decision theory based on quantum probabilities. This direction derives from mounting research within the field of cognitive science showing that decision theory based on quantum probabilities is superior to modelling human judgements than standard probability models [2, 1]. By quantum probabilities, we mean decision event space is modelled as vector space rather than the usual Boolean algebra of sets. In this way, incompatible perspectives around a decision can be modelled leading to an interference term which modifies the law of total probability. The interference term is crucial in modifying the probability judgements made by current probabilistic systems so they align better with human judgement. The goal of this article is thus to model the information seeker user as a decision maker. For this purpose, signal detection models will be sketched which are in principle applicable in a wide variety of information seeking scenarios.

1. RELEVANCE JUDGEMENTS AS DECISIONS

Decades of research have uncovered a whole spectrum of human judgement that deviates substantially from what would be normatively correct according to logic and probability theory. Probability judgement errors have been found so consistently that they have names e.g., the "conjunction fallacy" in which subjects readily judge the conjunction of event A and B to be more likely that either of the individual events, e.g., $\Pr(A, B) > \Pr(A)$. These findings are barely known, let alone accounted for, in current systems for information retrieval and recommendation, which strictly adhere to the laws of probability. Therefore, they do not always account for *how humans make decisions*, rendering them potentially less useful.

From a cognitive point of view, the key to explaining the conjunction fallacy is the incompatibility of subspaces. Consider figure 1. The perspective around deciding document 1's relevance is represented as a two dimensional vector space where the basis vector R_1 corresponds to the decision "document 1 is relevant" and $\bar{R}_1$ corresponds to "document1 not being relevant". A similar two dimensional vector space corresponds to the decision perspective around document 2. Initially, the cognitive state of the decision maker is represented by the vector Ψ, which is suspended between both sets of basis vectors. This situation represents the subject being undecided about whether document 1 or document 2 is relevant. Suppose the subject now decides that document 1 is relevant. This decision is modelled by Ψ "collapsing" onto the basis vector labelled R_1. (The probability of the decision corresponds to the square of the length of the projection of the cognitive state Ψ onto the basis vector R_1, denoted $\|\mathbf{P_1}\psi\|^2$). Observe how the subject is now necessarily uncertain about document 2's relevance because the basis vector $R1$ is suspended between the two basis vectors R_2 and $\bar{R}_2$ by the angle θ. The hall mark of incompatibility is the state of indecision from one perspective (e.g., the relevance of document 2) when a decision is taken from another (e.g., document 1 is deemed relevant). Incompatibility means that deciding on the relevance of document 1 may mean uncertainty about the relevance of document 2, which necessarily implies the information seeker can't form the joint probability $\Pr(R_1, R_2)$ [1]. (This is crucially different to the situation in standard probability theory in which events are *always* compatible, and thus the joint probability is always defined).

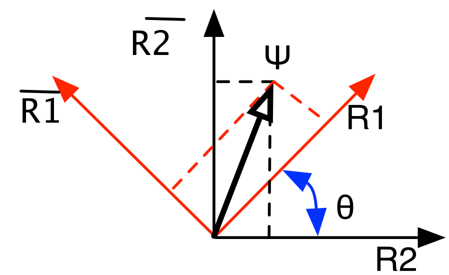

Figure 1: Incompatible subspaces in a decision on relevance

The consequence of incompatibility is an interference term $Intf$ which is central to the modelling human decision behaviour in the face of the conjunction fallacy. This term modifies the law of total probability allowing standard probabilities to be augmented by human judgement. [1] The partial derivation below shows that the interference term $Intf$ appears when the decision whether document 1 is relevant is made in relation to the incompatible subspace corresponding to relevance of document 2 (represented by projector $\mathbf{P_2}$ and its dual $\mathbf{P_2}^{\perp}$):

$$\Pr(R_1) = \|\mathbf{P_1}\psi\|^2 \quad (1)$$
$$= \|(\mathbf{P_1} \cdot \mathbf{I})\psi\|^2 \quad (2)$$
$$= \|(\mathbf{P_1} \cdot (\mathbf{P_2} + \mathbf{P_2^{\perp}})\psi\|^2 \quad (3)$$
$$= \|\mathbf{P_1}\mathbf{P_2}\psi\|^2 + \|\mathbf{P_1}\mathbf{P_2^{\perp}}\psi\|^2 + Intf \quad (4)$$

The intuition behind equation 4 is that the information seeker is undecided about relevance of document 1 due to perspective created by document 2. This indecision can be viewed as the subject "oscillating between two minds" which produces wave-like dampening or enhancement of the law of total probability via the interference term $Intf$ [1]: $\Pr(R_1) = \Pr(R_1, R_2) + \Pr(R1, \bar{R_2}) + Intf$. When events are compatible $Intf = 0$, and therefore the law of total probability holds, meaning that the probabilities adhere to standard probability theory. The interference term $Intf$ is a recent development in the literature in cognitive decision theory [1]. Models that employ it have been labeled "quantum cognition", as underlying probability theory is derived from quantum physics.

2. MODELLING THE DYNAMICS OF RELEVANCE DECISIONS

We propose to view judgments of relevance as a signal-detection decision task, and thereby open the door to applying decision models developed in cognitive science. The "signal" is a relevant fragment of information (say in a document), which is present amongst irrelevant information, i.e., "noise". We sketch two signal noise models: A Markov model and a quantum model.

The subject's cognitive state will be modelled as a 7-dimensional vector spanning the grades of relevance $\{-3, \dots, +3\}$, where -3 denotes highly irrelevant and +3 denotes highly relevant judgement. It is usual to assume at t_0 that each grade as equally likely, but for different information seeking tasks this can be varied. For example, [3] showed that the rank of a document affects the mean relevance grade given. In this case, the values in the initial cognitive state state vector can be biased towards higher scores because when documents are presented in rank order, relevant documents are more likely to be seen first.

In the quantum signal detection model the initial state will be evolved using the Hamiltonian matrix H. The intuition about behind H is that it models how the user moves between different grades of relevance, e.g., the entry in cell h_{ij} of H represents amplitude diffusing from relevance at grade i to grade j. A parameter σ determines the diffusion rate out of a particular relevance grade, and the parameter μ determines the diffusion rate back into that grade. The matrix H is used to model the dynamics of the graded relevance decision over time as a unitary operator $U(t) = e^{itH}$[1]. In the Markov signal detection model, standard practice will be adopted: A seven-state intensity matrix K will be used instead of the Hamiltonian H and the dynamics provided by using the Kolmogorov forward equation $T(t) = e^{tK}$.

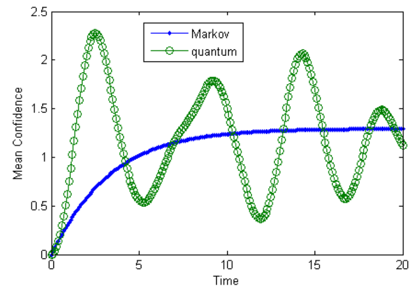

Figure 2: Quantum and Markov signal-noise relevance model simulations

3. EXPERIMENTAL DESIGN

Figure 2 illustrates the dynamics: At any time t, the probability of relevance at a particular grade j can be computed depending on the experimental condition (derived from [3]). The aim of the experimental conditions is to establish how incompatibility in the user's cognitive space influences relevance judgements. A between subjects design is proposed. In the first condition (A), subjects will be shown a single document d and asked to rate its relevance to the scenario on a 7 point scale. The subject is directed not to consider scores given to preceding documents as in [3], i.e., relevance judgements are assumed *compatible*. In the second condition (B), pairs of documents will be shown together one above the other, and the subject asked to rate first the top document and then the one below (which is the same d as in condition A). They *will not be* instructed to take a relevance decision in isolation, i.e., allowing for *incompatibilty*. The graded relevance judgement for each document will be recorded and time-stamped for each subject.

In the first condition (A), subjects where there is no order in the decision, the probability at time t can be computed by $||\mathbf{P^j}_A \cdot U(t)\psi||^2$, where $\mathbf{P^j}_A$ denotes the projector corresponding to relevance grade j in condition A, and the unitary operator U at time t. In the second condition where order is imposed, first B (a decision taken at time t'), then A (a decision taken at a later time t'), the probability of relevance at grade i is now given by $||\mathbf{P^j}_A \cdot U(t-t') \cdot \mathbf{P^i}_B \cdot U(t')\psi||^2$ The difference between these two probabilities allows the interference term $Intf$ to be computed [1]. An analogous approach can be taken with the Markov model but where the dynamics is driven by the operator $T(t)$.

4. REFERENCES

[1] J. Busemeyer and P. Bruza. *Quantum cognition and decision*. Cambridge University Press, 2012.

[2] J. Busemeyer, E. Pothos, R. Franco, and J. Trueblood. A quantum theoretical explanation for probability judgment errors. *Psychological Review*, 118(2):193–218, 2011.

[3] M. Eisenberg and C. Barry. Order effects: A study of the possible influence of presentation order on user judgments of document relevance. *Journal of the American Society for Information Science*, 39(5):293–300, 1988.