QM and IR: Another Perspective *

Extended Abstract†

Paul B. Kantor
DIMACS Rutgers
Piscataway NJ. 08854 United States
paul.kantor@rutgers.edu

ABSTRACT

Quantum Mechanics is a surprising conceptual framework, and a growing number of researchers know something about it through the notions of "entanglement" or "quantum weirdness." However, the relation between QM and the world that it describes includes some other puzzling characteristics, which distinguish it from both classical Physics and classical probability theory. This note explores some possible ways in which the QM theory of measurement might be of use in axiomatizing the processes by which a person gains information from the system and *vice versa*. The key innovation is to propose both a model, as seen by the system, of the discontinuous change of the state of the user, and a coupled representation of the states of the user *and* the system, which might ultimately be used to guide the selection of materials to be presented to the user.

CCS CONCEPTS

Information Retrieval; *Users and Information Retrieval; Retrieval models and ranking; Specialized information retrieval; Environment-specific retrieval - Enterprise search*

KEYWORDS

Axiomatics; Quantum Mechanics; Measurement Theory; Information Retrieval

1 INTRODUCTION

In Information Retreival (IR) there are two parallel traditions, which were called the "Systems View" and the :"User View." With the passage of time their meanings have evolved or decayed. Roughly, Systems View is primarily concerned with understanding and improving systems, and doing that by studying the behavior of algorithms. By contrast, the User View is primarily concerned with understanding more about the ways in which people behave, when seeking information, particularly in the presence of systems intended to be helpful. The two perspectives meet, unavoidably, in several ways. Pundits of the User View generally argue that better knowledge of the user is crucial to the development of better systems. While there is merit to this claim, one would hesitate to claim that to build a better roadway, we must know why people want to drive to Abilene. On the other hand, those toiling in the gardens of the Systems View find that we must present our systems to users, in order to measure how well they work.

A second relation between the two views has to do with whether a system wants to "understand a query" or "understand the user's need." This distinction, which was apparent even when search engines were library catalogs, has become sharper as search engines support themselves by their ability to present advertisements that attract a sufficient number of "clicks." For these commercial purposes the system does not need to understand any particular user, but needs only to propose items that attract a large enough fraction of the users who might be on the other side of the present user-system conversation[1].

The author's interest has been, for many years, in the problem faced by a system that serves a single user, or a quite small set of users, over a period of months or years. If there is a small set of users, they are presumed to share a common goal, although they may also be competitors in ways other than reaching that goal. [Examples are: bio-medical researchers seeking a cure for some specific disease; intelligence analysts seeking to understand and possibly predict the actions of particular groups or individuals; etc.] For problems of this kind, the system will "want to" build the best possible model of the user's interests, and use that model to select materials to be offered.

This note addresses a potentially new formulation of such a model, in the language of density matrices, and proposes an even more potentially[2] useful model, in the language of measurement theory, of how the conversation between the user and the system may be represented in this model. The model is not complete or usable, in its current form. It is presented in the hope that someone clever will find a way to fully operationalize the concepts and their implications.

2 SOME RELEVANT LITERATURE

There are two relevant literatures: the classical literature of QM, and some contemporary literature on IR, seeking to produce a useful marriage. To be useful, the marriage should produce ranking algorithms that are measurably different from those produced by standard reasoning about IR, such as the "probability ranking principle," or by recommender systems [13].

---

[1] Many names have been proposed for the extended dialogue that represents any search that extends beyond a single query. We use "conversation" here as an easy compromise. We remain agnostic about the intelligence on either end.

[2] A double negative is intended here.

2.1 Recent excursions in the IR field

There is a growing literature, sparked by some speculations of Keith van Rijsbergen [15] and there is not room to adequately review it here. Some key references are: [7,8,10,11,12,14,18,19]. Almost certainly the advances proposed here are anticipated by some among those authors, even if they are not expressed in precisely the same language. The author hopes that the present formulation will permit some further advances.

2.2 The situation in QM

Quantum Mechanics, having established its revolution somewhere between 110 and 90 years ago, has been remarkably stable since then. Advances in physical technique, which give us Moore's Law, have also produced a steady chain of experimental verification, including the observation of the Bose-Einstein condensate, some 70 years after it was predicted [1], and recent measurements confirming Bell's inequality (essentially, a formulation of the Einstein Rosen Podolsky "Paradox") over ever increasing distances [17].

Broadly, it can be said that QM introduces two important notions [3,16] in the description of physical states: vectors, and phases [6,9]. QM teaches that the state of a physical system can be described as a vector in an abstract (Hilbert) space. These vectors are described in terms of (physical) bases, which correspond to sets of orthogonal vectors in that same space, and which are eigenvectors of the operations associated to the corresponding physical measurements. Phases are complex numbers, and the vector space is a vector space over the complex numbers. These phases give rise to interference between vectors, because the (observable) probabilities of the results of measurements are given by the square of a particular matrix element. All of this is familiar to those following the literature on quantum computation or cryptography.

Perhaps less familiar are the basic processes used to describe measurement in QM. These are described using a mathematical entity called the "density matrix," which represents both the quantum-mechanical nature of the system, *and* the (classical) nature of the observer's uncertainty about that system.

The initial state is represented by some initial density matrix ρ_0. If the system is in a so-called "pure state" (which, although it can always be taken as a basis vector in some representation, is most often some linear superposition of basis vectors, with definite complex coefficients) it is represented as $|s \geq \sum_k c_k |e_k >$ where the $|e_k>$ are the basis states. The density matrix is given by the projection operator[3]: $\rho =| S >< S |$ Left to itself this will change according to some unknown time evolution operator *U(t,t')* and become $U(t,t')\rho_0 U(t,t')^\dagger$. But that is not what concerns us here. Rather, we want to draw upon the discontinuous measurement transformation.

"Orthodox" QM4 suggests that we think of the measurement as corresponding to some operator, having a (complete) set of eigenspaces corresponding to the possible eigenvalues of the operator. If we let P_k represent the projection onto the k-th eigenspace of the operator corresponding to the measurement, then the density matrix, when we know that a measurement has been made, but we do not know the outcome, becomes: $\rho_1 = \sum_k P_k \, \rho_0 P_k^\dagger$ In particular, if the initial state is a pure state, $|s >= \sum_k p_i \, |b_i >$ then the result of the measurement is the mixed state5 given by: $\rho_1 = \sum_i p_i \, P_i |b_i >$ To interpret this in the language of information retrieval we must ask what might correspond to "the operators and to the "eigenspaces."

3 PROPOSED FORMULATION

We propose here that in the conversation between the system and the user, the system makes "measurements" of user by presenting documents. These measurements change the state of the user. However, as in QM, the system (the "observer") does not necessarily know the result of the measurement. Instead, the best representation that the system can make, of the resulting state of the user, is a mixture representable by a density matrix, based on whatever the system knows about the user *before* making the measurement. Without mapping the mathematics to well-defined psychological states (or, as Belkin would have it, "Anomalous" states) we can follow the mathematical steps.

One way of thinking about information needs is to characterize them by the range of possible and satisfactory "answers." In the world of Information Retrieval we may formulate that as a set of *documents* which together provide a complete and irreducible (that is, none can be deleted) resolution of the user's specific information need.[6] Such "satisfying sets" may also have associated costs (in the worst case, some documents are in a language the user does not know, while others teach the grammar and lexicon of that language).

These ideas can be represented in QM using a language sometimes called "second quantization" but with a few crucial changes. Instead of a "vacuum state" |0>, we posit a state called the "satisfied state" |S>, which describes the user who either has no needs, or had the specific need under discussion, and it has been satisfied. We also posit that there is, corresponding to each document that might be presented[7], d, "creation operator" $a_d^\dagger$. (This notation is sometimes called "second quantization." See for example, [2]. Note, however, that we do not use the usual commutation relations, and are not concerned with phase or number operators.

---

[3] We use the "bracket" notation, which has become the standard in current discussions of QM.

[4] This phrase was a favorite of the late Prof. Eugene Wigner, who always remained aware of the fact that QM remains puzzling to those who dare to think carefully about it.

[5] This is the observer's view of the "collapse of the wave function."

[6] The user imagined here is distinct from the typical user who reads this paper, and who, as a researcher, seeks to remain perpetually *unsatisfied*, by positing new questions and information needs as the research progresses. ☺

[7] If the system presents something other than documents, such as facts, or assertions, the same reasoning will apply.

If a user's need can be satisfied by the specific document d we represent the user's "state of need" by $a_d^{\dagger}|S>$, where the operator can be thought of as "creating the quantum of need for document d.". If the user's need could be met by any of the individual documents $\in D$, we can represent the user's need by a sum over the projection operators corresponding to the states of needing the several different documents: .

$$\rho_{0-mixed} = \{\sum_{d \in D} v_d \, a_d^{+}|\mathrm{S}><\mathrm{S}|a_d \} \quad (1)$$

Here v_d is intended to represent the value, to the user, of satisfying the information need by means of document *d*. This may depend a lot on the document. For example, a paper by a verbose author may provide a more painful way of obtaining the needed information, for one user, while another user may be uncomfortable with a concise presentation, and prefer the leisure of having a long document to read.

Now, how are we to represent the presentation, by the system, of the specific document *d*? We propose that this is represented by the annihilation operator, as it is expected to remove the need for that document. So the "measurement operator" corresponding to the document "presenting the document d" is precisely.a_d We need to model how this operator acts on the states of the user. We propose these three equations:

$$a_d \, a_d^{\dagger} = 1 \quad (2a)$$

$$a_d \, |S \geq |S> \quad (2b)$$

$$a_e \, a_d^{\dagger} = a_d^{\dagger} a_e \quad (2c)$$

The first means that presenting *d* annihilates the need for *d*. The second and third together stipulate that if *d* was not needed, presenting it does not change the user's state. (Specifically, the role of the third equation is to commute the annihilation operator for a document *e* all the way over to the state of need, which then remains unchanged.)

With these commutation relations, the result of presenting the document d is to put the user into a state in which his density matrix contains one term representing complete satisfaction, |S> and a sum of other terms representing the fact that other documents would have also met the need. This is not totally satisfying, but it is not apparent how the strict formulation can enable the "other terms" to know that the user is now satisfied. But when this is the formulation, the system does not "know that it can stop, as its description of the user still contains, with some probability, terms representing alternative ways of satisfying the user." While the conversation will be terminated by the user, one of the long-standing challenges in Information Retrieval is whether we (that is, the system) can interpret a terminated conversation as a "satisfied user." To put it simply, the completely dissatisfied user will, with a few pathological exceptions, also terminate the conversation.

While there may be a more clever solution, for now the best I can think of is to include the system in the description of the density matrix, and allow for a transition, corresponding to this presentation, which lets the system "know" that the user is now satisfied.

With some reflection, it appears that this formulation will be adequate for this specific case, and also for the case in which satisfaction of the need can be achieved by any of several irreducible sets of documents, $d \in D$. In that case the (user) density matrix takes the more complex form:

$$\rho_0 = \sum_{d \in D} v_d \, \{\prod_{d \in D} a_d^{\dagger} \, |S >< S| \prod_{d \in D} a_d \quad (3)$$

We further suppose that there is some kind of feedback from the user, to the system, about the value that has been delivered by the specific document *d*. The notation then becomes rather cumbersome in two ways: first, the change in each set *D* to which *d* belongs will bring an associated change in the value of delivering the remainder of the set. This will, in general, be an increase in that value. As a first guess we might propose that the value is always inverse to the size of the set, perhaps according to law such as:

$$v_d^{hyp}(d \in D) = \frac{1}{|D|} \quad (4)$$

Or

$$v_d^{hyp}(d \in D) = e^{-|D|} \quad (5)$$

On the other hand, popular formulations of cumulated discounted gain might suggest slower drop off, related to the logarithm of |D|. Briefly, the density matrix corresponding to the information system will always be a pure state, as we presume that the system (which is the observer here) "knows its own state." [8] These states will have a simple binary flag indicating whether the user's need is satisfied, but they must also retain as much as possible of the information about the sequence of documents presented, and the resulting change in the user's state, encoded in information about the values *v*.

4. DISCUSSION AND FUTURE WORK

This work is motivated by two observations. The first is that the state of the user, which is surely the most vital thing that an information system can know[9], is not well determined. This leads to the idea that it might be represented in the language of quantum mechanics. That language allows for even a "definite state" (such as "spin pointing up") to be indeterminate (that is, representable only as a probability distribution) when measured in any "non-

---

[8] This could become a tricky philosophical point, related to the question of whether Schrodinger's Cat knows whether he is still alive. We propose not to engage.

[9] This understanding leads to an astonishing drive for systems and their developer to read our email, and mine the contents of our computers and phones for information that will make their services even more useful. The author finds himself not yet ready to be the biological bearer of a really clever computer chip, which will run his life.

commuting" ways. However, we also know that exposure to the presented materials (and, in fact, even to such things as document titles, or author lists, or even the short snippets presented by a search engine) *changes the state of the user*. This suggests the notion that perhaps the discontinuous change of the density matrix, or "collapse of the wave function" may provide a useful language in which to describe the users of systems.

We are proposing here to explore whether this formulation can remain conceptually consist *and also provide some guidance to system designers* as to what they might ask for, in the way of feedback, to satisfy the joint goal of bringing the user to the pure state represented by |S><S|. Ideally, the conceptual framework will also support the notion of retaining a "user model" but to do that we will need to link the set of conversations to their outcomes.

Ideally, the state of the system will transition in ways that respond to the history of the conversation. It is here that the possible order dependence of "set value" will naturally enter. But it is unclear how the system will "reason with this information," which, in the language of this model means that the system should be described, after each exchange, as having a "state" which effectively induces a ranking of the documents that are available for presentation. This is where all of the classical IR work on document-document relations, and possible semantic[5] or generative [4] representations will come into play.

At this point the author's optimism is tempered by the fact that he finds himself thinking in terms of classical probabilities. That is, after the measurement, the user is in one of several states, and the only thing the system has added to its knowledge is that the user assigned that value *v* in his most recent transition, in response to the "measurement" by document *d.* Can we find a way to reason about this knowledge, in any way other than trying to estimate the probability that a specific alternative document *d'* will induce a specific increase in value to the user. In a sense we face a problem that familiar from the world of fuzzy control systems: no matter how clever the internal manipulation of information, in those problems the control system must, in the end, speed or slow down.

Any such axiomatic theory is ultimately validated either by yielding improved performance of systems in extended conversations with users, or by exhibiting a kind of "generative potential" through making it easier for developers and engineers to conceive of potential innovations, and put them into hardware and software. The presentation given here had potential to be coherent and extensible, but converting it into a language that will speak to the engineers who create today's IR systems will be a substantial challenge. The pitfall to be avoided is mathematical intensity which is, to paraphrase Marvin Goldberger "like the breasts on a brass monkey: ornamental but useless."

5. ACKNOWLEGMENTS

The author thanks his teachers, Eugene Wigner, Sam Treiman and Gian Carlo Wick; all no longer with us, and support from the National Science Foundation under Grant #1247696**,** and the **Industrial and Systems Engineering Department of the University of Wisconsin Madison**, for hospitality and a courtesy appointment.

REFERENCES

1. Mike H. Anderson, Jason R. Ensher, Michael R. Matthews, Carl E. Wieman, and Eric A. Cornell. 1995. Observation of Bose-Einstein condensation in a dilute atomic vapor. *science* 269, 5221: 198–201.
2. F. A. Berazin. 2012. *The method of second quantization*. Elsevier.
3. Garrett Birkhoff and John Von Neumann. 1936. The logic of quantum mechanics. *Annals of mathematics*: 823–843.
4. D. Blei, A. Ng, and M. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993–1022.
5. Scott C. Deerwester, Susan T Dumais, Thomas K. Landauer, George W. Furnas, and Richard A. Harshman. 1990. Indexing by latent semantic analysis. *JASIS* 41, 6: 391–407.
6. Lev Davidovich Landau and Evgenii Mikhailovich Lifshits. 1965. *Quantum Mechanics Non-relativistic Theory: Transl. from the Russian by JB Sykes and JS bell. 2d Ed., rev. and Enl*. Pergamon press.
7. Massimo Melucci. 2012. Deriving a quantum information retrieval basis. *The Computer Journal* 56, 11: 1279–1291.
8. Massimo Melucci and Keith Rijsbergen. 2011. Quantum mechanics and information retrieval. *Advanced topics in information retrieval*: 125–155.
9. Albert Messiah. 1966. *Quantum mechanics. 2 (1966)*. North-Holland Publishing Company.
10. Benjamin Piwowarski, Ingo Frommholz, Mounia Lalmas, and Keith Van Rijsbergen. 2010. What can quantum theory bring to information retrieval. *Proceedings of the 19th ACM international conference on Information and knowledge management*, ACM, 59–68.
11. Benjamin Piwowarski and Mounia Lalmas. 2009. A quantum-based model for interactive information retrieval. *Advances in Information Retrieval Theory*: 224–231.
12. Benjamin Piwowarski and Mounia Lalmas. 2009. Structured information retrieval and quantum theory. *Quantum Interaction*: 289–298.
13. Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. *Introduction to recommender systems handbook*. Springer.
14. Dawei Song, Mounia Lalmas, C. J. van Rijsbergen, et al. 2010. How quantum theory is developing the field of information retrieval. .
15. Cornelis Joost Van Rijsbergen. 2004. *The geometry of information retrieval*. Cambridge University Press.
16. John Von Neumann. 1955. *Mathematical foundations of quantum mechanics*. Princeton university press.
17. Juan Yin, Yuan Cao, Yu-Huai Li, et al. 2017. Satellite-based entanglement distribution over 1200 kilometers. *Science* 356, 6343: 1140–1144.
18. Guido Zuccon, Leif Azzopardi, and Keith van Rijsbergen. 2009. The Quantum Probability Ranking Principle for Information Retrieval. *ICTIR* 9: 232–240.
19. Guido Zuccon, Benjamin Piwowarski, and Leif Azzopardi. 2011. On the use of Complex Numbers in Quantum Models for Information Retrieval. *ICTIR*, Springer, 346–350.