Quantum Theory and the Nature of Search

Sachi Arafat and **C. J. van Rijsbergen**
Department of Computer Science
University of Glasgow, UK

Abstract

The conceptual model and mathematical formalism of quantum theory are employed in creating a novel framework for modeling the computational search process addressing problematic issues that restrict information retrieval research. Mapping the mathematical formalism of search to that of quantum theory presents insightful perspectives about the nature of search. However, differences in operational semantics of quantum theory and search restrict the utility of the mapping. An approach is suggested for resolving these semantic differences aiming toward a sound mathematical and conceptual framework for search inspired by quantum theory.

Introduction

Information retrieval (IR) is the the field of research investigating the searching of information in documents, searching for documents themselves, searching for meta-data which describe documents, or searching within databases, whether stand-alone or datasets networked by hyper-links such as the Internet, for text, sound, video, images or other types of data. An IR system is commonly understood as that which deals with the relationship between objects and queries. Queries are formal statements of *information needs* addressed to an IR system by the user. The object is an entity which stores information in a database, known as a document. User queries are matched to documents stored in a database. Often the documents themselves are not kept or stored directly in the IR system, but are instead represented in the system by their pointers. Automated information retrieval systems were originally used to manage information explosion in scientific literature in the last few decades. The value of information is directly related to its ability to be located and used effectively, search engines thereby form a crucial component in the research and understanding of modern times. For something so crucial, IR can be a confusing area of study. Firstly, there is an underlying difficulty, with the very definition of IR, since there exists the adjacent fields of document retrieval, text retrieval, information seeking, information science, information management and others each with their own bodies of literature, theory and technologies which are deeply related to IR and each other to the point where the boundaries are unclear. Secondly, IR is a broad interdisciplinary field, that draws upon secondary fields such as cognitive science, linguistics, computer science, library science and it does so in a loosely organized fashion. It is tempting to refer to this conjunction of diverse areas as *"search science"*, however, due to the presence of ad-hoc techniques used to perform experimentation in IR and the absence of a (general) formal language for definition of IR concepts, components and results, it cannot be called a science. Furthermore, there are no specific definitions of search. With the abundance of methods available for finding information, whether through computer applications, libraries/librarians, a combination thereof or otherwise, a formal definition would need to accommodate a process far more complex than that of traditional web-based querying through systems like Google. The lack of a general formal specification method for search processes, IR research, and the absence of strict a scientific method underpinning it, has posed major barriers to future development and usefulness of research in the field (Arafat, van Rijsbergen, & Jose 2005).

Recent work in (van Rijsbergen 2004) based on ideas borrowed from quantum theory (QT) has suggested methods of formalizing aspects of IR aiming toward a comprehensive theoretical basis in which a search process can be completely defined and reasoned about, and a scientific basis inspired by operational methods in QT. It was subsequently found that there is a potential for QT methods to play a wider role in resolving the above IR issues (of definition and lack of scientific method) than suggested. In addition, it was found that apart from the mathematical formalism of QT which offers analytical tools convenient for representing IR concepts, the scientific method and operational structure (the way QT employs states and state changes) is also very useful. Inspired by these peculiar connections and on attempting to apply these methods and *map* search to QT, it was found that search requires to be re-examined from a perspective quite different from how it is traditionally perceived (see (Arafat, van Rijsbergen, & Jose 2005)) in order to deduce the feasibility, utility and method of the mapping. Thinking about search in this new way also suggests approaches for re-defining the concept of search. The overall goal for our research can be equated to being able to formally refer to IR as *"search science"* by establishing a specific definition of

Copyright © 2007, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

search and deducing scientific methods for the investigation of search, so it can be in all respects, a science.

This paper highlights the nature of the search process and the main problems responsible for IR research being in its current non-ideal state. An outline is given of current work on employing QT to address one of these problems proceeding with an approach to resolve the other causes. In the next section, with reference to the traditional laboratory perspective of IR (Ingwersen & Jarvelin 2005), the nature and scope of the the *evaluation problem* and *user problem* are discussed. Both these problems are dependent on the *definition problem*, which therefore needs to be addressed first. The laboratory view of search is a hindrance to adequate conceptualizations of these problems, thus an alternate view is suggested. The *stack model* provides this view enabling the visualization of the interaction of different research areas, with the advantage that the definition, evaluation, and user problems can be visualized in terms of the model. The section 'A New Perspective' concludes with details of all the key problems particularly elaborating the *conceptual problem* of defining search and with that resolved the section 'The Middle Form' outlines a scientific method for IR using the stack model. Further benefits of adopting QT concepts for IR are also presented. Our approach to solving the research problems in IR with inspiration from QT raises several new and interesting questions, suggesting changes in the method of experimentation, and re-defining boundaries between related research areas.

Background

The traditional model of a search process is depicted by Figure 1 and still applies to most search systems, see (Ingwersen & Jarvelin 2005). In order to search a data set, the documents in the set must first be indexed. The index is the same data set reduced to contain just the information (collection of words, media, and any metadata) about the set required to represent the collection of documents sufficiently according to a document model. Queries expressed by the user are interpreted according to a query model. A matching sub-process follows which takes the query and for each document assigns a value to the association between the interpretation of that query and the interpretation of the document according to their respective models. This association is termed relevance and can be defined in a multitude of ways (Mizzaro 1997). The results of the matching process are manipulated and shown to the user according to an interface model which is outside the scope of the laboratory model. For a simple list interface like that of Google, the results are ordered according to matching values, so that documents deemed most relevant to the query appear higher than the less relevant ones. Unlike Google there are search engines which allow the user to influence the matching sub-process by feedback. The user feedback expressed through the interface is known as relevance feedback and facilitates learning of user interests. A search system is evaluated according to its effectiveness and computational efficiency. The effectiveness of a search system is usually deduced on analysis of two features, its ability to correctly associate query & document (to judge relevance) and to adequately present results on the interface. Effectiveness of a set of relevance judgments is traditionally deduced using precision $\frac{RETREL}{RET}$/recall $\frac{RETREL}{REL}$ measures (RET is the number of documents retrieved and RELRET is the number of previously judged relevant documents found in RET) to compare against prior human judgments on the same queries and in the same document collection, the majority of evaluation experiments are performed this way. The method of evaluating result presentation, or interface models are equivalent to human-computer interaction evaluations employing questionnaires and usage log analysis. IR research tends to focus on improving the effectiveness of relevance judgments by the creation of novel models for individual components or methods for combining different models. To test a new measure for matching a query to documents, one would select a test collection of previously collected user judgments and run precision/recall experiments for queries in the collection to deduce effectiveness. There are several problems to evaluation done this way, firstly, conclusions to such an experiment are very subjective as they are limited to the scope of the test collection and to the context (factors influencing human perception of information) dependent viewpoints of prior human judgments. Secondly, there are no definitive ways, in general, to deduce why one system performs better than another since prior user judgments assess the system view of relevance; and are informal opinions of the whole system not specific formal reasons attributed to particular components. The latter problem is inescapable when using human judgments. In order to further understand experimental results from test-collection based evaluation, one runs complementary live user-based experiments where users are given tasks to complete on a search system with effectiveness being judged using statistics on questionnaires and usage logs. An inherent weakness is that an experiment cannot be duplicated even if the same users are retained since their context changes. Thus the experimental results are not definitive. Indeed such problems are inherently due to the human factor, and will be referred to as the *evaluation problem.*

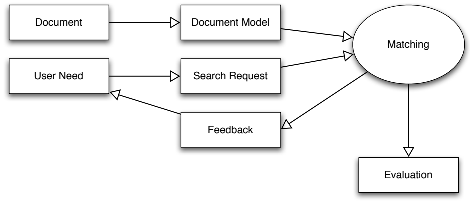

Figure 1: Data flow in the Laboratory model

Unlike relationships between components in a physical system such as in Figure 2 which are made apparent through a prior theoretical framework, IR research does not exhibit general frameworks to deduce such relationships between its components. In the physical system the effect of modifying one parameter on other parameters can be predicted, in an IR evaluation the user is a parameter but there are no ways to determine the effect of its modification on the eval-

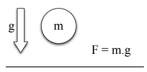

Figure 2: Simple physical model with underlying theory

uation. Since there are no formal methods to relate users, the potential effect on user judgments of modifying a component, is generally too unpredictable relative to the physical case, prior to experimentation. An approach for creating theoretical apparatus for IR with the ability to formally compare components would need to tackle the user issue, since concepts of effectiveness of a search system are inevitably tied to user definitions. If instead there were a way to work with formally specified abstract users which approximate real users then experiments can be duplicated and verified. The problem would then be one related to the effectiveness of the approximation ability of abstract user models. With an accepted user model, the abstract user would then be a controllable experimental factor. However, there are no general formal methods for abstracting user behaviors for creation of abstract user specifications, and it is not common practice in IR research. Instead the research literature uses brief and informal natural language descriptions for users, i.e. "the users are university students with moderate experience in searching". A practical advantage of formal specification is that it would allow relatively economical user simulation type experiments and provide (see (White *et al.* 2005)) definitive results for a specific user specification which can then be verified. With formal specification the effectiveness of a search engine can be identified with specific user types and evaluation results can be reasoned about in terms of the user specifications.

A user cannot be defined out of context, and is strongly coupled to a set of user-system interactions, and an interface. Unfortunately the interface and the user-system interactions components are usually only specified in informal natural language expressions. Overall, there is no way to formally specify an IR experiment in its entirety. As a result there is no way to formally reason about evaluation results with respect to the interface and interactions. Traditionally, only the document, query and matching models (Figure 1) are formally specified and admit several formal specifications. For example, the association between a document and query termed relevance can be represented in terms of logical implications, conditional probabilities or inner products in a vector space. Unfortunately there exists no unified framework for theoretically comparing between different representations in terms of effectiveness. These problems with defining users, interface, user-system interaction, and the inability to compare different formal representations where they exist, hinders research as it severely limits theoretical conceptualizations of search scenarios and deductions therein. In comparison the simple physical system allows many degrees of freedom for devising hypothetical extensions to already specified scenarios, such as extending Figure 2 with two balls and many walls. Definition problems apply not only to specifying a search scenario but also to the relationships between the IR research field and neighboring fields as stated in the introduction.

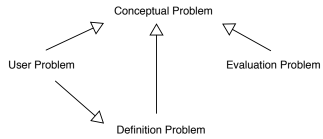

Figure 3: Dependence of Research Problems

The research problems discussed above can be grouped into three broad categories. Firstly there is an inability to verify experimental results in IR and other issues related to evaluation, these shall be collectively known as the *evaluation problem*. Secondly, the evaluation problem is related to inability in formally specifying users, which will be denoted as the *user problem*. Finally, the *definition problem* denotes issues with formally defining all components, their inter-relationships, theoretically reasoning about them and the relationships of IR as a field with other fields. Resolution of the user and evaluation problems depend on the resolution of the definition problem. A fourth problem yet to be discussed is the *conceptual problem* which is addressed in the proceeding section completing the relationship between our research problems as illustrated in Figure 3 where the arrows denote dependence. The dependency between research problems is not formally provable on a general level, the figure serves only to illustrate reasonable relationships as per common experience in IR research. It is at the same time interesting and unfortunate to note that due to the definition problem it is difficult to present without ambiguity these above research problems in a formal way, whether in a mathematical formalism or using formal natural language statements. The stack model of the proceeding section provides a new perspective of a search process that reduces some of the ambiguity. The four categories of problems appear to sufficiently abstract the issues faced in our initial research, which in hindsight attempted to resolve the definition problem. Initially a part of the definition problem was addressed to suggest a unified theoretical basis for comparison between different types of document, query and matching models which had a variety of formal specifications. Such was one of the goals of (van Rijsbergen 2004) which showed that mapping of three types of these models, the vector-space, probabilistic and logical models to the mathematical formalism of Hilbert spaces in the way employed by quantum theory results in a single framework in which one is able to theoretically compare among models, providing greater opportunities for formal analysis than previously available. One important aspect not elaborated in (van Rijsbergen 2004) was that of modeling relevance feedback in the QT formalism. In an attempt to model relevance feedback in the Hilbert Space formalism of quantum mechanics a new set of problems were faced and interesting questions raised which collectively suggest novel inquiries about the nature of IR.

A New Perspective

Our approach to resolving the definition problem for document and matching models (as in (van Rijsbergen 2004)) using QT required an alternative conceptualization of a search process, in terms of states of the search engine and their transformation over the course of a search session. In terms of states and state changes, if the documents and their corresponding relevance judgments (deduced from relevance feedback in the search session) are to be defined as the *state* of the search engine then relevance feedback corresponds to the evolution of the state. The simplest mapping of a document and matching model to the quantum theoretic Hilbert space is the trivial map of the vector space model in IR where documents d_i are represented as vectors with co-ordinates $d_{ij} \in \mathcal{R}$ with values denoting the influence of the respective word in the document. In such a mapping each document vector $|d_i\rangle$ is a possible state of the search engine, and the document with only one term a pure state. The state of a search process can be represented as the mixture state $\rho = \sum_{i\in|Doc|} w_i |d_i\rangle\langle d_i|$ where the w_i are relevance values. It can also be represented as a superposition of document vectors, and in other forms. Therein lies a critical issue which will be alluded to in the proceeding section: What is the best way to represent the state of a search system?

The QT Hilbert space presents some new mathematical features which are useful for modeling details about the state of search. For example the influence of a word in a document is often denoted by the product of the word frequency within the document with its rarity among the collection of documents. In the Hilbert space these features of a word need not be amalgamated into a real-number product and can instead separately 'stored' in complex number co-ordinates, which would increase the analytical power of the representation of state. The reason for this is that keeping these term features separate would not decrease the analytical power of representation, as the traditional representation of terms can be established by multiplying the weights. Thus the analytical power is at least the same. As the semantics behind the weights, frequency and rarity, are different concepts, it is useful to keep them separate for addressing research inquiries about document models with respect to one parameter and then comparing to inquiries pertaining to the other parameter. Also consider the way a real and complex number are semantically related when used to represent concept pairs like amplitude and frequency of a wave. There is a striking similarity between these pairs and the pair word frequency and word rarity, for a term. The frequency of a term in a document is like amplitude (within that document), and its 'rarity' or 'inverse document frequency' is its frequency, like that of a wave. Whether there lies any benefit in exploring these relationships between the representation of a term in a collection and an electric signal among a set of signals, is open to research. Finally, the argument of a complex number $r.e^{i\theta}$ has some simple mathematical structure pertaining to symmetry, which can be exploited to relate the rarity values of terms (further elaboration is outside the scope of the paper). Thus, using complex numbers increases the analytical power of term representation suggesting an area of investigation in itself.

With the trivial map of the vector space model to the QT formalism there are many mathematical features that present interesting analytical possibilities for analysis of hypothetical search processes. The difficulty stems from representing any concept of dynamics. On the mathematical level a set of transformation matrices applied to the state represented by a density matrix ρ, would update it to a succeeding state $U\rho U^{-1}$. However what is the method for deducing transformation matrices? One can map the traditional method of relevance feedback in the vector space model which is a low-level approach, it involves keeping a ideal query vector and updating it on relevance feedback according to a simple learning function. There are many interesting opportunities to do such mappings and corresponding mathematical manipulations in the Hilbert space which could upgrade the traditional low-level feedback within an elegant geometrical framework (see (van Rijsbergen 2004)). These opportunities present initial advantages of mapping to a QT formalism, and are due only to the mathematics. How can one create transformation matrices at a higher level? Physical systems have underlying theory, or at least universal facts embedded in its design such as ideas about energy, which attach semantics to a state and allow characterization of its change according to the theory. The theory is used to deduce change operators which are used to for predicting physical events. For a search process state containing relevance judgments, state change would mean a change in the values of these judgments. General interests of the user can be deduced through feedback, and while one can devise machine learning (or otherwise) models for predicting future interests on the low-level there are no general underlying higher-level principles. Any such underlying principles useful for predicting change in user interests would need to model user behavior thereby requiring to address the user and definition problems. There are several questions raised in the pursuit of conceptualizing a search process in terms of states and state changes. In QT the evolution operators can have physical meaning and are rich in algebraic and geometric properties, for example a group structure for a set of unitary transformations. What could be corresponding mathematical properties for relevance feedback operators, and what would, for example, a group structure for a set of feedbacks mean in terms of search concepts? These type of inquiries are inherent in the state based conceptualization of QT and its mathematical formalism. There are no IR frameworks which inquire in this way. IR does not offer a general method to answer these inquiries, especially those about a higher-level approach to relevance feedback without first resolving the user and definition problems.

Stack Model

There are paradigmatic differences between QT and IR in their operational methods and semantics as discussed above. It is clear that the mapping to QT has mathematical benefits, with features like complex numbers and algebraic structures on the Hilbert space but it is unclear as to how one can use the QT formalism, what QT concepts cannot be used and most importantly what some QT concepts corresponding to the formalism (such as algebraic structure) mean in IR terms. A bottom-up approach would be to assess the benefits of each feature of the QT framework individually, however it is difficult to deduce IR meanings for IR models created in a QT formalism. Instead a top-down approach is attempted with the premise that the apparent paradigmatic differences between IR and QT can be reduced if one no longer thinks of a search process in the traditional sense, according to the laboratory model, but instead in the way QT would perceive a process, as a physical process. The search process can be abstracted as a physical process in which there are a set of interactions between two physical systems, the user and system. The stack model in Figure 4 is a visualization inspired by this perspective and is visually equivalent to a traditional network architecture diagram illustrating the design of the protocol of communication between two agents. It corresponds to a design method for visualizing arbitrary search scenario designs. Stack labels and slices are according to the purpose of a scenario design. For example in an investigation focusing on document/matching models which is at the memory/reasoning level, one may not need to discern between the gestures and physical layers, combining them instead. The stacks method of design is unlike the laboratory model, since it freely includes (by design) other aspects of the user and system, including as in Figure 4, the interface (gestures layer), hardware (physical layer), search strategy (session layer) and also details the user. According to Figure 4 a user interacting with a system involves a reasoning sub-process instigated by the memory layer, which then influences a search strategy (session layer) and activates corresponding gestures expressed by the physical layer representing their physical expression tools. The system in Figure 4 setup with the same design, observes the user's physical action, interpreting a gesture in the context of other factors (such as prior user feedback) and updating its notion if user interests (memory) according to a user-interest update policy (reasoning). Memory and reasoning layers in the stack correspond respectively to the document and document model in the laboratory model. In the stack design a search process is a set of sub-processes either between agents which is of type *observational* or between components within an agent which are of two types *expression* and *interpretation*; these are common principles for any stack based model of a search process. The sub-process of type expression denotes a general flow of activity toward the memory level, it corresponds to one set of changes that are internal to an agent upon another agent interacting with it. The other set of changes internal to an agent are those that lead it to react to the prior interaction, these changes are of the expression type and generally flow from the memory layer toward the agent's expressive faculty, the interface, denoted by the physical layer. By a 'flow' of changes it is meant a set of effects which end at a destination (i.e. the memory layer for interpretation flows) and are caused by entities in a preceding layer with the order corresponding to the type of flow (see Figure 4).

Stack design is a shift in the traditional conceptualization of IR toward the state-based conceptualization for physical systems in QT. It was inspired by the QT way of abstracting physical sub-systems and their relationships. In terms of the physical semantics of QT, each layer is a physical subsystem and can only directly influence adjacent sub-systems which preserve the order inherent in search systems. The stack model does not presume any specification language for any layers. A clearer relationship between the user, definition and evaluation problems can be observed when these problems are in terms of the stack visualization. First, as the research in (van Rijsbergen 2004) shows, the mathematical formalism of QT can be used as a language for modeling document and matching models, corresponding to the memory layer and reasoning layers; although a QT specification of the reasoning in a search engine is theoretically limited due to the relevance feedback issue (as discussed previously). Second, retrieval research provides no general frameworks for modeling interfaces and interactions, thus the gestures and session (containing search strategies) layers have no general modeling language. Similarly on the user side there are no general modeling languages for any aspects. Instead the literature of IR and related disciplines provide several specific models for each layer often specified in natural language form, making it difficult to compare models theoretically. Hence the definition problem for IR on which the user problem depends is that of having several models specified in different languages for each layer of each agent and no analytical way to compare between them. It is clear that there requires to be a methodology for modeling corresponding to the visual stack diagram. An adequate formalism for specifying each layer, inter-layer communication and between agent communication need to be deduced; this corresponds exactly to the definition problem and is addressed in the next section following a discussion of the *conceptual problem* on which it is dependent.

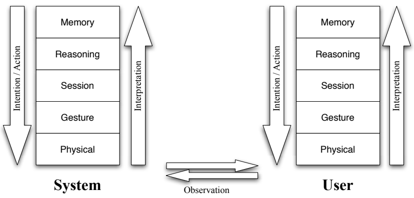

Figure 4: A Stack Model

Modeling of Diverse Scenarios

Unlike in traditional IR research, the stack design is used to model the evaluating agent, and hence the evaluation process itself as illustrated in Figure 5. The practice of modeling the measurement device in QT inspired this way of using stacks. An information seeking process is one which considers the broader experience of the user in the search process which can involve more agents, and possibly human agents. An example would be the case of a user interacting with librarians, automated search systems and other agents in a library. The information seeking process is easily visualized by adding more stacks to the design of the typical information retrieval model of two stacks. A distributed search scenario such as

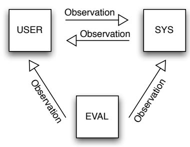

Figure 5: Model of an Evaluation

in peer-to-peer searching or searching through multiple indexes representing an intranet, can also be visually depicted in the same way by a set of stacks.

In conclusion, our attempt at resolve the definition problem for document and matching models (as initiated in (van Rijsbergen 2004)) led to a different way of thinking about a search process, in terms of a stack design, which admits a visual description (Figure 4). It also created questions about high-level models of change based on relevance feedback, and about semantic mappings between QT and IR which are necessary for interpreting mathematical manipulations on the QT formalism. Inquiries about paradigmatic differences between QT and IR are due to the conceptual problem which is addressed in the next section. Addressing the definition problem (hence user and evaluation problems) is the same task as creating a formal framework corresponding to the stack visualization in which it is possible to deduce how and when one can use the QT formalism (or any other) for modeling components of a search process, i.e. relevance feedback.

The Conceptual Problem

The stack model depicts a search process as a physical process, a set of interactions between two or more physical systems through subprocesses of the observation, interpretation and expression types. On a higher level the depiction is that of a communication process between agents. What then differentiates a search process from any other communication process? Is it that the system interface, as would be represented by the gestures and physical layers is a search engine interface? What constitutes a search interface, is not the user always searching for something whether implicitly or otherwise? For example a user using any interface can be said to be searching for results of his interaction - instead of formal documents. This line of inquiry can be reformulated in the low-level language of interaction between physical systems: what is peculiar about the physical process in which search is done and any other physical process? What is search? This is the conceptual problem for IR on which the definition problem depends since to define search, a search interface and other components, it is required to lay down principle interpretations about a search process.

Resolving the Conceptual Problem

The conceptual problem in Figure 5 must be initially addressed as all other problems depend on it. Initially we make some observations which illustrate the conceptual problem and then proceed to resolve it. First, it is important to realize that relevance feedback can be implicit, any interaction with a search system is a relevance feedback. Second, a search interface is not necessarily explicit, and can be defined as an interface by which a user's cognition is influenced, which thereby includes any interface that can be sensed assuming any sensory activity modifies cognition in some way. As exemplified by information seeking, a user does not require a computerized search engine to search, any interaction with another human or agent is also search. Hence any communication process is a search. Generalizing further in terms of the stack model, in any agent an internal sub-processes (of type expression or interpretation) which is an interaction between any two layers is a search process. Finally, any process in which change occurs is a search process. It is assumed that every physically existing process has change occurring in it, however insignificant, until it ceases to exist. It is also assumed that an abstract process, one not directly physically existing, exists indirectly, as it requires the existence of physical processes to express it. For example, a mathematical description of a process is an abstract description which can only exist when expressed, either in the mind of a mathematician by physical cognitive processes or in physical reality. Thus every process, physical or abstract is a search process. Defining information as that which causes representations to change (see (Mackay 1969)), it is deduced from prior statements that every process can potentially accept information. In terms of this definition and the stack model, the directed sub-processes in Figure 4 denote *information flow*.

A common supposition of IR research is that the purpose of users interacting with a search engine is to fulfill their *information need* (IN). In light of current discussion, a user can never fulfill the IN associated with him, as he is defined by a set of physical processes which exist as long he does, meaning that the user is always ready to accept information and therefore always has an IN. The concept of IN applies unambiguously to all agents, with information defined as in (Mackay 1969) since all agents change over time. Agents have a potential to change therefore have an IN meaning that the IR system has IN according to the generalized definition of IN and not in the traditional sense. All processes change over time, and are therefore search processes. Defining information and search in this way directly addresses the conceptual problem. The definition problem now has a much larger scope, instead of referring to traditional computer based search it now refers to any process. This allows one to try any pre-existing formalisms for defining processes whether it be QT methods for defining physical processes or otherwise. Previously, aspects of IR were being mapped to QT for certain mathematical, and some operational benefits, now, since a physical process and a search process are defined as the same thing, there are no conceptual problems to the mapping. The definition problem otherwise remains the same.

The Concept of Relevance

Relevance is a label and a reason attached to a particular set of changes. In a traditional search process a user is expected to interact with some items, the semantic of these interactions can be embedded in the interface, i.e. the Google ranked list where higher rank indicates increased likeliness of *relation* to the given query. The reason a user interacts with a particular item, such as a document link in the search results which are presented as a ranked list is complicated and uncertain as it pertains to user cognition. It is usually assumed that the user interacted with an element as they thought it was 'relevant', or at least that they thought their interaction was 'relevant'. Thus, the word 'relevant' can be taken as a primitive, replacing otherwise complex descriptions of reasons for user interaction. In our approach where the user is modeled as an agent, one can state the reason for an interaction, in terms of the stack entities, changes within them, and influences between them. Upon a set of user interactions, the user interest as a general concept, is to be determined from the 'relevant items' which have been involved in interaction. The word 'relevance' is assigned as a label for the general concept of user interest. Thus the words 'relevant' and 'relevance' are labels *referring* to complex reasons and concepts (respectively) pertaining to user cognition. The labels can be formally defined in terms of particular user models. However, these labels are not particularly important, as we have a formally defined artificial user agent, so those complex reasons/concepts referred to by these labels are well defined in our framework. User interest is used in the definition of the user agent and is inscribed in the definitions of individual entities and their inter-relationships. One can still refer to artificial agents interacting with relevant data or make statements like "an artificial agent's concept of relevance has changed" except now these statements can be refined in terms of the stack model. Information need with respect to the user agent, retains its traditional meaning in our framework, changes of IN are not described in terms of concepts of relevance but in terms of the stack model, which in turn can be *translated* (see next section) into statements using the terms 'relevance' or 'relevant'.

Role of Quantum Theory

In the new conceptual model, every process is a search process, what is then the role of QT? Firstly, the operational methods of QT inspired the stack design to allow thinking about search in terms of states and state changes. Second, the QT formalism can be used to specify the memory and reasoning layers for the system, and the user in a simple user model as in (Aerts & Gabora 2002). The main problem with the mapping of change in document/matching models was that while QT is associated to physical theories which it can use to create evolution matrices and give meaning to them, IR has no such high-level theory - and in order to have such a theory, one would need to first resolve the user and definition problems. A specification language is then required in which a user can be defined with respect to other components. It must be a single language for describing all layers of the stack in Figure 4 and their inter-relationships. We deduced a language of *flows/changes* (information flows) general enough to completely specify components during search. However for such a general language to be of use it must admit translations into other languages. A complete specification of the language of flows is outside the scope of this paper, a brief account proceeds the current section. A third role of QT is its semantics or physical theory, as our general specification language uses the physics semantic of energy for assigning general meanings to changes. Recalling from the introduction, the aim is to create a search science. Search is now defined and the science is the topic of the next section.

The Middle Form

The language of flows is a typed language corresponding to the stack, it is a 'middle form' borrowing ideas from QT to model IR. A search process is specified by a number of agents, who are in turn specified by their component stack. To each process belongs a set of flows (changes) and a set of states, which are defined in formal natural language in terms of other components as necessary. The challenge is then to translate each flow and state description to a language which admits the desired theoretical framework. With a further abuse of terminology, the following terms are equated: a flow, 'flow of energy', potential to change, information flow (in that there is an IN of a component within an agent or a separate agent). IN admits properties which make it conceptually similar to physical energy. First there is the notion of IN conversion, each entity in a stack has a different *type* of IN (and 'flow') associated with it, the flow to an entity from an adjacent entity is a conversion of some IN from the source entity to that of the destination entity or qualitatively it is a type conversion. It simply denotes cause and effect of changes. The type structure is then used to describe semantics at each layer of the stack. Secondly there is a concept of IN conservation. In the case of a human user, this supports the idea in the prior section that one is always searching and their IN is never fulfilled only changing, hence any quantitative conceptualization of IN must reflect the constancy of IN. For an artificial user modelled on a human user with 'user interests', its IN is translated as its 'potential to change'. In the artificial case, 'user interest' as inscribed in its stack model can be fixed to test hypothesis about that interest, however user interest is only part of the IN. During the lifetime of a simulated search process involving artificial user agents, the user agent always has the potential to change state until its pre-defined stopping condition prevents it. Therefore its 'internal activity' which forms its IN ensures the IN is always changing and that the agent is always 'searching' according to the definition of search in above sections. The constancy in the artificial agent refers to the fact that its potential to change always exists until its terminating condition is reached, therefore its overall IN is always conserved. The intention of defining IN as 'energy' is to have a simple and universal underlying theoretical principle for all search processes on which further theory can be built eventually allowing an analytical scenario like that of the physical case (Figure 2). Further elaboration on IN and reasons for relating it to energy are outside the scope of this paper.

Translation

With a general specification language in place what requires to follow are translations of a search scenario specification

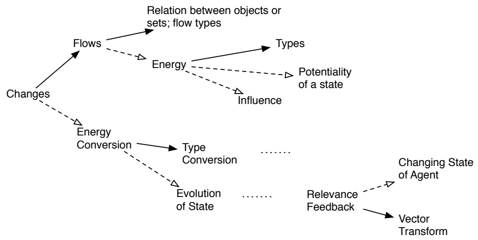

Figure 6: Relationship between languages

in flow language to a destination language which offers the desired analytical capability. A translation can be trivial, for example, the representation of the memory layer as anomalous state of knowledge in (Belkin, Oddy, & Brooks 1982) has many similarities to the 'potentiality of concepts' representation in (Aerts & Gabora 2002). Overall in order to unify the multitude of IR models for each layer both in their syntax and semantics (termed as representation and method space in (Arafat, van Rijsbergen, & Jose 2005) substantial further research is necessary. A map of this research in Figure 6 shows some of the goals. Note that a translation of that part of a description language referred to as 'syntax' will exhibit both a resultant syntax and a semantics. For example a relevance feedback algorithm translates to a matrix on the Hilbert space, a syntax, but this also carries QT semantics. Thus the definition problem, along with its dependents still exist but there is now a framework to systematically resolve them. The problem of creating a 'search science' is now transformed into the problem of defining appropriate translations of the scenario to a formal language where there is a 'science' to be used, if any.

New Problems

The language of flows is a formal specification method, and therefore a particular search scenario can itself be analyzed using concepts from computability/complexity theory. An important issue about recursion needs to be addressed here. As if every process is a search process, then indeed the researcher modeling a search process is also doing search. Are there any theoretical gains or problems due to this self-referential issue? Further questions of a definitive type are also within the scope of the formal specification, some elaborate ones can be formulated: "Is the effectiveness of a particular IR system for some specified set of abstract users always limited in some way unless the user *learns* to effectively use some interface feature?" These types of questions are new inquiries for IR.

Conclusions

In this paper we have detailed the key research problems in IR explaining how thinking of IR as a physical process using the operational setup of QT offered a new perspective on these problems and resulted in the stack model. The conceptual problem in IR was resolved by re-defining the concept of search and IN. With the conceptual problem resolved a framework for addressing the definition problem was introduced. The problem of creating a science for search is equivalent to translating from the flow language representation to a formalism with the desired analytic tools. A substantial research effort is now required to formalize a field as interdisciplinary as information retrieval, but it is hoped that the framework presented will form a foundation to that effort. There are new interesting issues concerned with self-reference, which may need to be addressed first. The immediate research concern is to show the practical benefit of the flow-language based specification by illustrating how its analytic ability can be used to influence design decisions in a particular IR system specification with specific user definitions.

A crucial practical implication of our work is formal user simulations which could reduce experimentation costs. A related implication is that if researchers can use pre-defined user models for their experiments then the need for live user studies is reduced. Other sub-parts of the research community, the more user orientated, would then concentrate on creating realistic user models. In any case, due to formal specification there are likely to be clearer boundaries between related research areas and some standardization of the research practices therein.

Acknowledgements

We gratefully acknowledge Peter Bruza, Matt Leifer and Alex Wilce for comments on our work and EPSRC for partial funding.

References

Aerts, D., and Gabora, L. 2002. Contextualizing concepts using a mathematical generalization of the quantum formalism. *Journal of Experimental and Theoretical Artificial Intelligence* 14:327–358.

Arafat, S.; van Rijsbergen, C. J.; and Jose, J. 2005. Formalising evaluation in information retrieval. In *Workshop on Evaluating User Studies in Information Access Fifth International Conference on Conceptions of Library & Information Science - Context: nature, impact and role.*

Belkin, N. J.; Oddy, R. N.; and Brooks, H. M. 1982. A.s.k for information retrieval part two: Results of a design study. *Journal of Documentation* 38(3).

Ingwersen, P., and Jarvelin, K. 2005. *The Turn: Integration of Information Seeking and Retrieval in Context.* Springer.

Mackay, D. M. 1969. *Information, Mechanism and Meaning.* The M.I.T Press.

Mizzaro, S. 1997. Relevance: The whole history. *Journal of the American Society of Information Science* 48(9):810–832.

van Rijsbergen, C. J. 2004. *The Geometry Of Information Retrieval.* Cambridge University Press.

White, R. W.; Ruthven, I.; Jose, J. M.; and van Rijsbergen, C. J. 2005. Evaluating implicit feedback models using searcher simulations. *ACM Transactions on Information Systems (ACM TOIS)* 23(3):325 –361.