Quantum Information Dynamics and Open World Science

Dominic Widdows
MAYA Design, Inc.
`widdows@maya.com`

Peter Bruza
Queensland University of Technology
`p.bruza@qut.edu.au`

Abstract

One of the fundamental insights of quantum mechanics is that complete knowledge of the state of a quantum system is not possible. Such incomplete knowledge of a physical system is the norm rather than the exception. This is becoming increasingly apparent as we apply scientific methods to increasingly complex situations. Empirically intensive disciplines in the biological, human, and geosciences all operate in situations where valid conclusions must be drawn, but deductive completeness is impossible.

This paper argues that such situations are emerging examples of *Open World* Science. In this paradigm, scientific models are known to be acting with incomplete information. Open World models acknowledge their incompleteness, and respond positively when new information becomes available. Many methods for creating Open World models have been explored analytically in quantitative disciplines such as statistics, and the increasingly mature area of machine learning.

This paper examines the role of quantum theory and quantum logic in the underpinnings of Open World models, examining the importance of structural features of such as noncommutativity, degrees of similarity, induction, and the impact of observation. Quantum mechanics is not a problem around the edges of classical theory, but is rather a secure bridgehead in the world of science to come.

Introduction

In spite of its great successes, quantum theory has remained something of a scientific enigma. The focus on subatomic phenomena has led many to consider the phrase "quantum effects" to mean "weird things that happen to very small particles". In many fields, this has enabled classical models to be followed without attending to any of the insights and warnings from quantum theory.

This paper takes a different approach. We suggest that many of the ways of thinking encouraged by quantum theory are normal, not exceptional. To see this, we relax our focus from the subatomic domain of traditional quantum mechanics, and instead try to engage the reader in some of the conceptual tools that quantum theory depends upon in order to make its famously accurate predictions. The purpose of this paper is to showcase ideas that are important in quantum theory and are taken for granted in other parts of science. As more information-intensive science develops, mathematical models that are used in quantum mechanics are becoming increasingly vital in other fields. As well as providing a survey, this paper speculates about the relationships between areas of logic, approaches to science, and the information systems required to support different scientific enterprises.

To make the importance of these formal descriptions easier to understand, we introduce a new term, *Open World Science*. By an *Open World*, we mean a scientific system whose description may require a boundless amount and variety of data to produce a complete model. The amount and variety of data is often unknown to the scientist before the modelling process is undertaken, and the model itself often encourages new information gathering to become more complete and to make useful predictions. Like the quantum physicist and philosopher David Bohm (1951, Ch 23), we argue that successful classical models are really limiting cases in which Closed World assumptions are reasonably (sometimes extremely) accurate.

As a case study of Open World methodology, we describe geometric models in knowledge discovery that resonate strongly with ideas from quantum theory (van Rijsbergen 2004; Widdows 2004). We also outline some responsive properties that we believe are beneficial design principles in the practical task of designing information systems capable of supporting Open World Science.

Quantum mechanics, and the challenge to classical physics

Belief in classical determinism was perhaps most famously defined by Pierre-Simon Laplace, in the following statement from the 1820 *Essai Philosophique sur les Probabilités*:

> An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in one single formula the motions of the largest bodies as well as the lightest atoms in the world, provided that its intellect were sufficiently powerful to subject all data to analysis; to it nothing would be uncertain, the future as well as the past would be present to its eyes.

As is now well-known, quantum theory poses great challenges to Laplace's determinism on the subatomic level. The most important challenge to classical mechanics (for the

---

purposes of this paper, at least) came about through problems in observing the position and momentum of a particle simultaneously. To solve the equation of motion and thus predict the behaviour of a simple classical system (for example, the movement of a frictionless pendulum that swings through a small angle), one needs to know the position and the momentum (or velocity) of the particle at a given time t. However, in the framework of quantum theory, it is impossible to measure precisely both the position and the momentum of a particle at the same time, without taking into account which measurement operation was performed first. This result, due to Heisenberg, is known as the Uncertainty Principle, and it applies to other pairs of observables, including the energy and time of a system.

The relevance of such results to scientific and philosophical world-views remains a complex and incomplete topic (Bohm 1980, Ch 4), and this has in many ways allowed quantum theory to be regarded by many practitioners in other fields as something of a scientific curiosity. The usual response to the Uncertainty Principle is not to accept that the Universe is fundamentally unpredictable, but to state (correctly) that in most cases of experience, we are dealing with large enough ensembles of particles that the unpredictabilities of quantum mechanics even themselves out. Philosophically, the suggestion that statistically valid generalizations from a large number of uncertain instances make an acceptable substitute for determinism is an unhappy workaround, but it has enabled the relevance of quantum theory to be relegated to a particularly small region of discourse. The phrase 'quantum effects' has come to mean 'apparent uncertainty regarding the behaviour of individual subatomic particles'. There are some compelling descriptions of the contributions that quantum mechanics might make to modelling consciousness at the microscopic level (Penrose 1989). However, on the macroscopic level, much of physics and philosophy goes on in a classical mode (Malin 2001, p. 89).

Problems with the informatics of the classical model

Classical techniques serve particularly well in some situations, such as the simple harmonic motion of the pendulum, in which the motion of a 'point particle' attached to a 'light inextensible string' moving under the influence of a 'uniform gravitational field' can be successfully predicted based only upon the initial position and momentum of the particle, the length of the string, and the strength of the gravitational field. Such models have certain hallmarks:

- The problem is neatly factored into 'theory' and 'data'. The equation of motion can be deduced from Newton's laws of motion, and solved using the calculus, for an entirely general case, and this general case can be adapted at will to different specific values. In other words, the solution requires no interaction between the data gathering and the mathematical modelling.
- The description length of the specific case is very short. Once one understands the mathematical behaviour of pendula in general, one needs only the length of the string, strength of the gravitational field, and the initial position and momentum.
- These variables can be regarded as independent of one another. The movement of the particle does not affect the strength of the gravitational field, and we even assume without too much inaccuracy that the weight of the particle does not stretch the string.

Many physical systems are not amenable to such simplifications, or such clear cut predictions. Examples are easy to come by. If a puppy is born with a weight of 2 kilos at a particular place and time, how much will the adult dog weigh? The answer is that you don't know for sure, though an expert breeder may be able to give an approximate range of weights which the healthy adult dog may be expected to have. The purpose of biology or veterinary science is not to predict the exact weight of the dog in question at all future times; it is to enable insight into the nature and structure of dogs and mammals more generally, and to recommend courses of action that maximise the health of the creature under consideration.

There are a host of examples where scientific predictions are inexact, but still valid. Economists attempt to predict the general trends and risks to commerce and prosperity, medical studies try to determine that a particular drug is effective and not harmful in a statistically significant number of cases, and so on. Unlike the classical pendulum, it would be a mistake for these disciplines to focus too fixedly on making exact predictions about individual objects. This point is made by Aristotle in the *Nicomachean Ethics* (Bk I, Ch 3):

> It is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits; it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician scientific proofs.

In the face of Aristotle's call to common sense, Laplace's determinism is at least questionable, and at worst, is used as a justification for poor decision making. For example, some politicians and their advisors have dismissed studies that demonstrate that global climate change demands action, on the grounds that the "science of climate change is uncertain" (Eilperin 2005). The claim that there is uncertainty in climatology is of course true, just as there is uncertainty in predicting the adult weight of a newborn puppy. Nonetheless, climatology should certainly be able to make probabilistic predictions and valid recommendations, just as a veterinery scientist should make predictions of the probable range of a creature's healthy adult weight, and recommendations for nutritional practices to achieve this weight. In many fields, such scientific recommendations must be made with probabilistic outcomes in mind, rather than with certain predictions. Totally certain predictions would require complete information about the system being studied, and for complex systems this is simply not available.

Even without scientific examples, we should be somewhat suspicious of any system for reasoning and decision making in general that requires completeness of information. Animals including humans often make decisions in the face of partial information, and our survival demands this of us. For example, animals often flee when startled by a noise, not because they are sure that the noise signifies danger, but because given the risks, flight is a more sensible option than closer investigation. In the hands of Gabbay and Woods (2001), a New Logic is emerging that takes the costs and risks of information discovery into account, as well as the cost of deduction in time and computation, and of course, the risk that a bad decision may lead to a negative consequence for the organism in question. The modelling of logic and reasoning needs to expand beyond the purely deductive steps. Our model must apply costs and benefits to information gathering; it must balance the application of deduction and induction; it must respond to reasonable abductive requests for further information; and it must assess the quality of outcomes for the agent who is performing the scientific task. Gabbay and Woods' New Logic puts forth a cohesive framework for addressing questions such as risk. For example, a logical agent would not in real life wait for a snake to bite her on the grounds that she lacked proof that this particular snake was venomous. If classical physics claims that only determinist predictions are scientifically valid, and therefore, if in doubt, one *should* wait for the snake to bite, then logical creatures naturally need more than classical physics to survive!

Another problem with the classical model is that it does not take the interaction between the observer and the observed into account. A good psychologist or ecologist is carefully taught experimental techniques by which the experimental setup itself has as small an effect on the outcome as is possible. For example, psychological experiments are often constructed with a decoy question, so that the experimental subjects (in this case, humans) do not know what question is really being asked of them. However, it is a widely held rule-of-thumb that only a perfect experiment could have no interference between the experimental setup and the results, and, as in thermodynamics, such a perfect system is not in practice possible.

Classical physics has been able to neglect these issues because it works with an information model that is fundamentally *static*. The experimental constants can be measured at the beginning of the experiment, or at any other time, and they will not have changed, even though the system is being observed and other experiments may be underway at the same time. For example, if we set up two pendulum experiments next to one another, we can quite safely assume that the masses of the pendula themselves will not interfere with one another: their mutual gravitational interaction will be negligible compared with the base gravitational field. In practice, such an assumption is more than good enough to send people to the Moon and bring them safely back again. The danger with static classical models arises partly from their spectacular success: they work so well in their appropriate domains, it becomes tempting to take them out of context and apply them unscrupulously. This can lead us to fall foul of Aristotle's warning above, in the hope of achieving objective certainty where it cannot be reached.

Some practical solutions from quantum theory

The heart of our argument in this paper is that quantum theory can help science to bridge the gap between classical physics and the information rich sciences of complex systems, such as medicine, economics, and climatology. Note that this paper is concerned primarily with informatics. The reader interested in the comparison of *physical* primitives between classical and quantum mechanics should consult the excellent and highly readable analysis given by Bohm (1951, Ch. 8).

Though the subatomic results of quantum mechanics such as wave-particle duality and the Uncertainty Principle are well-known in the scientific mainstream, the *logical* reasons for these phenomena are still specialist topics. The logic of quantum mechanics was introduced in full in the 1930's (Birkhoff & von Neumann 1936), though one can find clear descriptions of the conjunction (meet) and disjunction (join) operators in Grassmann's original *Audehnungslehre* that founded the theory of vector spaces (Grassmann 1862). By representing quantum particles and their wave-functions as 'state vectors', and exploring the ordered relationships between linear combinations of these state vectors, Birkhoff and von Neumann were led to consider the lattice of subspaces of a vector space V as a logic. The conjunction or meet in this logic (logical *AND*) is defined by the intersection of two subspaces, the disjunction or join (logical *OR*) is represented by their linear sum, and the negation or complement of a subspace (logical *NOT*) is represented by its orthogonal complement. The idea that logical concepts should be represented as linear subspaces, instead of just arbitrary subsets, leads to key differences between the logic of quantum mechanics on the one hand, and the Boolean logic of classical physics on the other. A definitive text on the nature of these logics and their physical models is that of Varadarajan (Varadarajan 1985). Some of the important features of information representation in quantum logic and quantum theory generally are described below.

Degrees of similarity

An important mathematical feature of quantum geometry is that the angles between subspaces give rise to probabilities of observing different experimental outcomes. Now, if perpendicular angles in a geometric space are defined, it is possible to define other angles in terms of these. It follows that knowledge of orthogonal complements (i.e., the quantum logical structure) is enough to define angle or similarity between subspaces, from which the well-known probabilistic outcomes of quantum mechanics can be derived (van Rijsbergen 2004, p. 24, citing an unpubished paper by Von Neumann, 1957).

It follows that quantum logic holds some states to be more similar to one another than others, even if they are not equal. This is different from classical Boolean logic, in which states are either equal or unequal (Varadarajan 1985, Ch 1). Degrees of similarity or belief are of course important in many disciplines, such as medicine. It is possible, for example, to give a definition for the phenotypic condition "being underweight" as a Boolean predicate (for example, using a threshold cutoff in the possible ranges of Body Mass Index), but this should not disguise the fact that there are varying degress of underweightness or overweightness.

Smoothing and Induction

Once degrees of similarity (or conversely distance) are defined, it is a simple next step to define clusters of more similar (or conversely less distant) objects. This can lead to the definition of geometric 'closure' operators, whereby a few exemplars of a particular concept can be used to generate a more broad-ranging concept, and some concepts are more stable than others. The relationship between geometric closure, inductive hypotheses in machine learning, and the distributive law (present in classical logic but notably absent from quantum logic) was explored in earlier work (Widdows & Higgins 2004). Boolean logic and set theory have no natural analogue of this behaviour, so all sets or concepts are equally stable. This can lead to logical problems, such as Quine's (1964) famous suggestion that the word *gavagai* could refer equally well to the set of all rabbits, or to the set of all rabbits plus the Eiffel Tower. In quantum logic, the latter set would be unstable without also including concepts in the linear span of *rabbits* and *The Eiffel Tower*. Quantum logic in vector spaces is not the last word here: an alternative and sometimes more practical closure condition may be convexity (Gärdenfors 2000). Nonetheless, such ideas can still be motivated by contrasting the spatially expansive quantum wavefunction with the isolated classical particle.

Quantization

Once it is understood that concepts can be represented by a more or less dispersed region of a space, one needs to explain how a given object is discovered to be in one of a discrete set of states. In quantum theory, the set of states and their associate probabilities is expressed accurately in terms of angles, though the actual mechanism of the collapse is still something of a mystery. In language, a similar quantization occurs along many physically measurable dimension, such as the description of a person as *young, middle-aged, old*, a temperature as *cold, cool, warm, hot*, or a musical tempo as *largo, andante, moderato, allegro, vivace*. Similar (sometimes more formally defined) definitions of particular states are employed in medicine ("is an epidemic"), economics ("is in recession"), or meteorology ("is a hurricane"). Classical physics offers no explanation of how concepts may naturally arise through defining different regions along a scale of measurement. Such an explanation is natural in quantum logic (formally, pure states are given by the Eigenvectors or stationary vectors of projection operators, though this may be too literal a definition for many applications).

Projections and Non-Commutativity

The measurement operators of quantum logic are famously *projections* (van Rijsbergen 2004, p. 56). The probability of observing a particle with (normalized) state vector a in a pure state represented by the subspace B is given by $||\pi_B(a)||$, the magnitude of the projection of a onto B.

Such projection operators do not commute with one another: in generic situations, it is not the case that for subspaces B and C, $\pi_B \pi_C = \pi_C \pi_B$. To visualize this, suppose that the vector space is the plane, B is the subspace given by the line $x = 0$, and C is given by the line $y = x$. The image of the projection π_B is the line $x = 0$, whereas the image of the projection π_C is the line $y = x$, from which it follows that the order of projections is important for all nonzero points.

Non-commutativity is an important principle in many cognitive and practical disciplines. More strongly, one might argue that commutativity is a foolish assumption in many everyday situations. In spite of the mathematical difficulties arising from commutativity, we argue that any conceptual theory wherein the order of play is unimportant cannot adequately model human or natural phenomena. Measurement in classical mechanics is always commutative, but quantum mechanics naturally allows for non-commutativity.

Composition of projections is not the only possibility for combining concepts or operators in quantum mechanics. For example, the tensor product may be used, which may lead to concepts becoming entangled (Aerts & Gabora 2005).

The Impact of Observation

In many disciplines, including quantum mechanics, there is no way of gathering data without affecting the system that is being observed. When posing a questionaire or carrying out a psychological study, it is important to realise that the nature of the questionaire or study may directly affect the outcome.

Quantum mechanics naturally embodies this principle, partly by virtue of its non-commutative formal structure: the fact that the survey has been carried out interferes with the data resulting from the survey. Classical mechanics of course assumes that it is possible to observe a system without the observation process biasing the results in any way.

Summary

Many of the properties of quantum theory and quantum logic are examples of more general and often quite typical scientific considerations. On the other hand, some of the classical counterparts of these properties are oversimplistic and impractical. While there is much insight to be found in classical models, we hope that the examples given in this section are enough to make the reader question the dominant teaching whereby classical physics is normal and quantum physics is strange.

Open World Science and Quantum Information Dynamics

As Laplace suggested, the predictive certainty of classical models is achieved on the assumption that all information necessary to describe an entire system can be simultaneously gathered. Remembering that all good scientific models make carefully chosen simplifying assumptions, we should not regard this determinism as a failing of the classical model. However, we must understand the nature of the classical assumptions, in order not to apply them by mistake and create inappropriate scientific models and stipulations in other domains.

Successful classical assumptions are based on the presumption that we are describing a closed system, and that, once observed, we can fully predict the system, without ever returning our attention to observe its evolution. Such a system may be described as a *Closed World*. The part played by assumptions about the isolation of physical systems, and the importance of these assumptions to classical determinism, is of philosophical importance (Hoefer Summer 2005). A successful example of Closed World Science may be the creation of models in celestial mechanics in which only the masses of the Sun, Moon and Earth are taken into account, models that are perfectly appropriate for many aspects of satellite navigation.

We believe that the behaviour of many systems of interest to science (including the behaviour of cognitive agents) is better described by *Open World* models. By an Open World model, we mean a scientific model that assumes that the observer must observe the system from time-to-time, in order to keep her knowledge about the system up-to-date, and her model as accurate as possible. Open World scientists need to presume that systems are complex, and that the boundary around systems is not always clear or fixed. New information may become relevant, or it may always have been relevant and its relevance not perceived. Important examples of Open World Science may include many models in epidemiology and population dynamics, where it is often necessary to presume that the entire population cannot be simultanously observed, and that fresh information may arrive and need to be organized and incorporated at any time.

Figure 1 shows some contrasting properties and methods that we believe are often correlated with the differences between Open World and Closed World models. We do not intend this figure to be definitive or exhaustive in any way, but we do believe that some of the correlations suggested are ones that many working scientists encounter in many domains. On the one hand, Closed World models assume static information models, and once initiated, the methods for following the model to its conclusions are largely deductive and rational. On the other, Open World models are initiated and updated throughout their lifespan, and so the methods for following the model include induction from repeated information gathering, abductive suggestions of new information that may be sought, and are often empirical. At the heart of Open World Science, there must be an information model that is dynamic. The domain examples given in Figure 1, citing mechanics as mainly Closed World, medical informatics as mainly Open World, and linguistics as more of a mixture, are somewhat anecdotal: though many motivating examples in these contrasting sciences can be found, so can counterexamples, and we should anyhow be wary of any attempt to classify sciences along any single dimension.

As argued in the previous section, there are many reasons for thinking of quantum theory as one of the sciences that demands a Open World approach. The observation of the system is of paramount importance in the theory, and observation affects the system itself. The predicted behaviour of a quantum system is not completely known, and the system interreacts in complex ways so that all relevant information about the system cannot be simultaneously known. The information about a quantum system is a dynamic, not a static bundle of information, hence our use of the term 'Quantum Information Dynamics.'

Quantum Information Dynamics and Knowledge Discovery

In recent years, connections between certain geometric models of information and quantum theory have received growing attention (Gabora & Aerts 2002; Widdows & Peters 2003; Widdows 2004; van Rijsbergen 2004; Aerts & Gabora 2005; Bruza & Cole 2005a). Of particular note is Van Rijsbergen (2004). Information retrieval systems are some of the earliest technologies to have made conscious use of the idea of geometric (Hilbert) spaces, whose relevance to topics such as knowledge discovery and text mining is becoming increasingly but slowly understood. By means of an example, we describe some of the benefits arising from quantum approaches to knowledge discovery, which we believe exhibit some of the hallmark qualities of an Open World approach to science.

In the mid-nineteen eighties, Don Swanson, an information scientist, made a chance discovery by connecting two disparate on-line medical literatures, one dealing with Raynaud's disease, the other with dietary fish oil (Swanson 1986a; 1987). Patients with Raynaud's disease suffer from intermittent blood flow in the extremities — fingers, toes, and ears. At the time, there was neither an effective general treatment, nor cure. Swanson formulated the explanatory hypothesis that fish oil may be a beneficial treatment, which was later verified by clinical trials.

Swanson's serendipitous discovery highlights a more widely occurring phenomenon. In order to deal with the information explosion, disciplines and expertise are becoming increasingly specialized and insular with little awareness of kindred, or potentially allied, specializations. As a consequence, disparate bodies of knowledge form, and with them "undiscovered public knowledge" (Swanson 1986b). Moreover, in our view, the Swanson discovery illustrates the need to move to Open World Science. The system is complex, and the boundary around the system is not clear or fixed. New information may become relevant, or it may always have been relevant and its relevance not perceived. In addition, the role of the human cognition is central and cannot be ignored. In this light, Swanson's discovery is an example of an abductive scientific discovery. Gabbay and Woods (2005) have convincingly argued that abduction has its roots in cognitive economy. Put crudely, it is cheaper to hypothesize, than to pursue a deductive agenda in relation to a problem at hand. Broadly stated, abduction is a mode of information gathering relevant to the Open World perspective, whereas deduction is a mode of reasoning suited to a Closed World perspective. In relation to this contrast, it is important to note that deduction is a mode of inference which has largely been studied independent of the (human) agent performing the reasoning (Gabbay & Woods 2005).

It is interesting to briefly consider Gabbay and Woods'

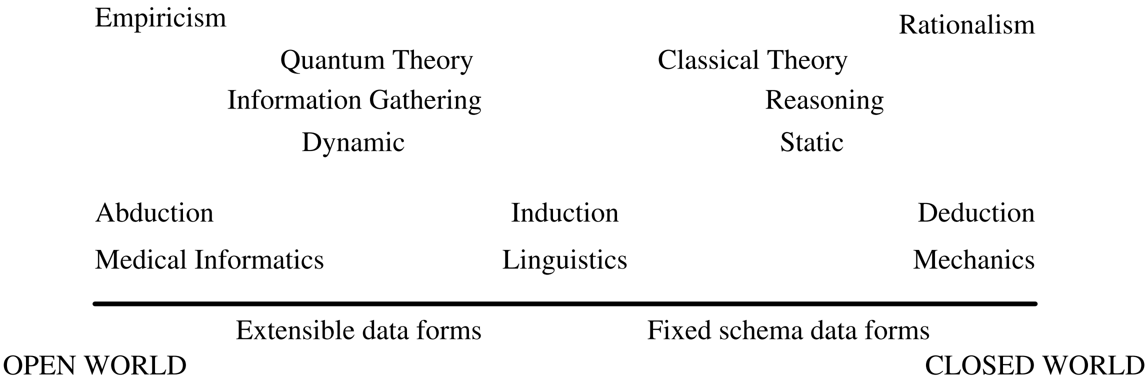

Figure 1: Cartoon showing some of the contrasting emphases of Open World and Closed World science

conjecture about the cognitive roots of abduction within the framework of Gärdenfors' (2000) three level model of cognition, and then relate it to Quantum Theory.

In Gärdenfors' model, information representation varies greatly across the different levels. Within the lowest level, information is pre- or sub-conceptual and is carried by a connectionist representation. Within the uppermost level information is represented symbolically, for example, by higher order linguistic structures. It is the intermediate, conceptual level (or "conceptual space"), which is of particular relevance to this account. Here properties and concepts have a geometric representation in a dimensional space. We subscribe to the view that associations and analogies generated within conceptual space play an important role in hypothesis generation which lies at the heart of Swanson's discovery. Gärdenfors (2000, p. 48) alludes to this point when he states, "most of scientific theorizing takes place within the conceptual level". His conjecture is aligned with Gabbay and Woods' insights regarding the cognitive economic basis of abduction in the following way: Within the conceptual space, inference takes on a decidedly associational character because associations are often based on context-sensitive similarity (e.g., semantic or analogical similarity), and notions of similarity are naturally expressed within a dimensional space (Open World). Inference at the symbolic level, however, is transacted in a linear, deductive fashion (Closed World). It may well be that because associations are formed below the symbolic level of cognition, significant cognitive economy results.

The field of cognitive science has recently produced an ensemble of semantic models which have an encouraging, and at times impressive track record of replicating human information processing, such as human word association norms (Lund & Burgess 1996; Burgess, Livesay, & Lund 1998; Lowe 2000; 2001; Landauer & Dumais 1997; Landauer, Foltz, & Laham 1998; Patel, Bullinaria, & Levy 1997; Levy & Bullinaria 1999; Sahlgren 2002). The term "semantic" derives from the intuition that words seen in the context of a given word contribute to its meaning. Colloquially expressed, the meaning of a word is derived from the "company it keeps", a famous quote originally from the linguist J.R. Firth (1890-1960). Although the details of the individual models differ, they all process a corpus of text as input and represent words, or concepts, in a (reduced) high dimensional space. These models are interesting in light of the Swanson discovery as they open the door to gaining operational command of cognitive semantics and associated human pragmatic inference mechanisms, like abduction (Bruza & Cole 2005b; Bruza, Song, & McArthur 2004). Semantic space models can be considered a computational approximation, albeit primitive, of Gärdenfors conceptual spaces described above (Bruza *et al.* 2006).

Recently a highly speculative but potentially far reaching discovery was made by some physicists: The formalization of quantum mechanics shows a strong connection with the mathematical basis of semantic space models (Aerts & Czachor 2004). The discovery raises the following speculation: Given that semantic space models have such an encouraging record of replicating human cognitive performance, can quantum mechanics be used to model human sub-symbolic reasoning, like abduction? Put more broadly in the context of this article, does quantum theory underpin information gathering from the Open World perspective, and if so, how? We don't have a definitive answer to this question, but our earlier work tries to set out the first steps (Bruza & Cole 2005a; Bruza, Widdows, & Woods 2006).

One way to illustrate where further investigations may lead is to reconsider the Raynaud/fish oil connection in the light of quantum theory. The illustration derives from the following speculation: Can the concepts "Raynaud" and "fish oil" be viewed as particles which exhibit something akin to quantum entanglement? Previous research cited above has shown a semantic space model to be closely related to a density matrix which represents the state of the quantum system. For example, let S_r denote the semantic space computed from the corpus of documents around Raynaud's disease, and let σ_r denoted the associated density matrix. Similarly, let S_f and σ_f denote the semantic space and associated density matrix of the corpus of information around dietary fish oil. The two are then considered to be a combined quantum system. If the combined state σ can be written in the form $\sum_i p_i(\sigma_r^i \otimes \sigma_f^i)$, then the combined quantum system is deemed "separable", otherwise it is "entangled". Framing the connection between the concepts "Raynaud" and "fish oil" via entanglement is not simply our own fanciful speculation (Aerts, Broekaert, & Gabora 2005; Aerts, Czachor, & D'Hooghe 2005). Nelson & McEvoy (2007, this issue) are looking to quantum entanglement to explain "spooky action-at-a-distance" effects between word associates in human memory. This speculation has two important aspects. The first is cognitive, namely, that entanglement in semantic space parallels entanglement in conceptual space (that is in human cognition). The second is the potential bearing on (semi-)automated knowledge discovery systems. It has been shown that the statistical connection between the concepts "Raynaud" and "fish oil" is statistically weak (Bruza & Cole 2005b). As a consequence, it is challenging to build automated knowledge discovery systems using models based on classical probability theory. Assuming that the quantum entanglement of concepts does manifest in semantic space models, and furthermore, the entangled concepts represent potentially meaningful connections, then this may lead to radically different information retrieval and knowledge discovery technology than currently exists. It is important to note in passing that quantum entanglement may manifest between concepts whereby the distance between them in semantic space may be large. This parallels the quantum entanglement of particles across large physical distances. In the end, the hope is that the entanglement of concepts may underpin information retrieval technology which can span disparate islands of knowledge and thus enhance human awareness. The motivation behind and the conception of such technology arises from Open World Science. There are many challenges to surmount, for example, determining whether a given state is entangled is NP-Hard, but this need not preclude the discovery of effective, heuristically-driven solutions.

In general, the creation of information systems that provide effective support for Open World Science still contains great challenges, though developments in recent years (such as RSS, Really Simple Syndication, and other publish subscribe systems) are encouraging. Just to combine existing, related information from a variety of different data sources is often acknowledged as a semantically difficult problem *par excellence*, without adding the difficulty of seeking new information and supporting updates on changing information. Much of MAYA Design's research has been devoted to solving these problems, in the hope of enabling the creation of the *Information Commons*, a global scientific database of accepted facts and observations (Lucas, Senn, & Widdows 2005). The Information Commons is implemented as a network of database repositories, which share and replicate universally identified information objects called *u-forms*. U-forms are extensible and have no fixed schemata: in general, whereas classical systems may be described by a fixed set of variables according to a particular schema (as used in a relational database), Open World Science will make heavy use of extensible data forms such as u-forms and XML (Figure 1). The universal identity or UUID of each u-form enables any user to request or refer to any u-form: the boundaries of an information system depends on the availability, not on the definition or classification of information.

Such an information system incurs very real challenges in the location of u-forms, and in ensuring that each replica of a u-form is as up-to-date as possible with changes committed in other venues. However, the epistemological considerations that must be faced when building such a system are deeply rooted in the nature of the real world, which is tremendously varied and constantly changing. The design of information systems must not be bound by Closed World assumptions, but must embrace the evolving Open World challenges presented by quantum information dynamics.

Conclusion

Many scientific endeavors rely on dynamic information models, and on information systems that are capable of responding well in evolving situations with extensible needs. The discovery and implementation of models that adequately describe evolving and incompletely known situations is the purpose of Open World Science. Several structural properties of quantum theory, such as noncommutativity, underspecification, and spatial distribution and induction, are central to the needs of Open World Science. For these reasons, we believe that quantum theory should not be seen as a curiosity confined to the subatomic domain. Mathematical models developed in quantum theory have proved valuable in other branches of science. Far from being a peripheral curiosity, quantum theoretic ideas may pave the road to the creation of successful dynamic information systems.

References

Aerts, D., and Czachor, M. 2004. Quantum Aspects of Semantic Analysis and Symbolic Artificial Intelligence. *Journal of Physics A-Mathematical and General* 37:L123–L132. http://uk.arxiv.org/abs/quant-ph/0309022.

Aerts, D., and Gabora, L. 2005. A Theory of Concepts and Their Combinations II: A Hilbert Space Representation. *Kybernetes* 34:176–205. http://uk.arxiv.org/abs/quant-ph/0402205.

Aerts, D.; Broekaert, J.; and Gabora, L. 2005. A Case for Applying an Abstracted Quantum Formalism to Cognition. In Bickhard, M., and Campbell, R., eds., *Mind in Interaction*. John Benjamins: Amsterdam.

Aerts, D.; Czachor, M.; and D'Hooghe, B. 2005. Towards a quantum evolutionary scheme: violating Bell's inequalities in language . In Gontier, N.; Van Bendegem, J. P.; and Aerts, D., eds., *Evolutionary Epistemology, Language and Culture*. Springer.

Birkhoff, G., and von Neumann, J. 1936. The logic of quantum mechanics. *Annals of Mathematics* 37:823–843.

Bohm, D. 1951. *Quantum Theory*. Prentice-Hall. Republished by Dover, 1989.

Bohm, D. 1980. *Wholeness and the Implicate Order*. Routledge Classics, republished 2002. Routledge.

Bruza, P., and Cole, R. 2005a. Quantum Logic of Semantic Space: An Exploratory Investigation of Context Effects in Practical Reasoning . In Artemov, S.; Barringer, H.; d'Avila Garcez, A. S.; Lamb, L. C.; and Woods, J., eds., *We Will Show Them: Essays in Honour of Dov Gabbay*, volume 1. London: College Publications. 339–361.

Bruza, P., and Cole, R. 2005b. A Bare Bones Approach to Literature-Based Discovery: An analysis of the Raynaud's/fish-oil and migraine-magnesium discoveries in semantic space. In Hoffman, A.; Motoda, H.; and Scheffer, T., eds., *Discovery Science, 8th International Conference, DS 2005, Singapore, October 8-11, 2005, Proceedings*, volume 3735 of *Lecture Notes in Artificial Intelligence*. Springer. 84–98.

Bruza, P.; Cole, R.; Abdul Bari, Z.; and Song, D. 2006. Towards operational abduction from a cognitive perspective. *Logic Journal of the IGPL* 14(2):161–177.

Bruza, P.; Song, D.; and McArthur, R. 2004. Abduction in semantic space: Towards a logic of discovery. *Logic Journal of the IGPL* 12(2):97–110.

Bruza, P.; Widdows, D.; and Woods, J. 2006. A Quantum Logic of Down Below. In Engesser, K.; Gabbay, D.; and Lehmann, D., eds., *Handbook of Quantum Logic, Quantum Structures and Quantum Computation*. Elsevier. (In Press).

Burgess, C.; Livesay, K.; and Lund, K. 1998. Explorations in context space: words, sentences, discourse. *Discourse Processes* 25(2&3):211–257.

Eilperin, J. 2005. Severe hurricanes increasing, study finds. *Washington Post*.

Gabbay, D. M., and Woods, J. 2001. The New Logic. *Journal of the Interest Group in Pure and Applied Logics* 9 (2):157–190.

Gabbay, D., and Woods, J. 2005. *The Reach of Abduction: Insight and Trial*, volume 2 of *A Practical Logic of Cognitive Systems*. Elsevier.

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

Gärdenfors, P. 2000. *Conceptual Spaces: The Geometry of Thought*. MIT Press.

Grassmann, H. 1862. *Extension Theory*. History of Mathematics Sources. American Mathematical Society, London Mathematical Society. Translated by Lloyd C. Kannenberg (2000).

Hoefer, C. Summer 2005. Causal determinism. In Zalta, E. N., ed., *The Stanford Encyclopedia of Philosophy*.

Landauer, T., and Dumais, S. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review* 104:211–240.

Landauer, T.; Foltz, P.; and Laham, D. 1998. An introduction to latent semantic analysis. *Discourse Processes* 25(2&3):259–284.

Levy, J., and Bullinaria, J. 1999. Learning lexical properties from word usage patterns: Which context words should be used? In French, R., and Sounge, J., eds., *Connectionist Models of Learning, development and Evolution: Proceedings of the Sixth Neural Computation and psychology Workshop*, 273–282. Springer.

Lowe, W. 2000. What is the dimensionality of human semantic space? In *Proceedings of the 6th Neural Computation and Psychology workshop*, 303–311. Springer Verlag.

Lowe, W. 2001. Towards a theory of semantic space. In Moore, J. D., and Stenning, K., eds., *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*, 576–581. Lawrence Erlbaum Associates.

Lucas, P.; Senn, J.; and Widdows, D. 2005. Distributed knowledge representation using universal identity and replication. Technical Report MAYA-05007, MAYA Design.

Lund, K., and Burgess, C. 1996. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments & Computers* 28(2):203–208.

Malin, S. 2001. *Nature Loves to Hide*. Oxford University Press.

Patel, M.; Bullinaria, J.; and Levy, J. 1997. Extracting semantic representations from large text corpora. In French, R., and Sounge, J., eds., *Connectionist Models of Learning, Development and Evolution: Proceedings of the Fourth Neural Computation and Psychology Workshop*, 199–212. Springer.

Penrose, R. 1989. *The Emperor's New Mind*. Oxford University Press.

Quine, W. V. O. 1964. *Word and Object (Studies in Communication)*. MIT Press.

Sahlgren, M. 2002. Towards a Flexible Model of Word Meaning. Paper presented at the AAAI Spring Symposium 2002, March 25-27, Stanford University, Palo Alto, California, USA.

Swanson, D. 1986a. Fish Oil, Raynaud's Syndrome, and Undiscovered Public Knowledge. *Perspectives in Biology and Medicine* 30(1):7–18.

Swanson, D. 1986b. Undiscovered public knowledge. *Library Quarterly* 56:103–118.

Swanson, D. 1987. Two medical literatures that are logically but not bibliographically connected. *Journal of the American Society for Information Science* 38:228–233.

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

Varadarajan, V. S. 1985. *Geometry of Quantum Theory*. Springer-Verlag.

Widdows, D., and Higgins, M. 2004. Geometric ordering of concepts, logical disjunction, learning by induction, and spatial indexing. In *Compositional Connectionism in Cognitive Science*. Washington, DC: AAAI Fall Symposium Series.

Widdows, D., and Peters, S. 2003. Word vectors and quantum logic: experiments with negation and disjunction. In *Proceedings of Mathematics of Language 8*, 141–154.

Widdows, D. 2004. *Geometry and Meaning*. CSLI Publications.