

Knowledge in Quantum Systems (Extended Abstract)

R. van der Meyden

Manas Patra

School of Computer Science and Engineering

University of New South Wales

{meyden,manasp}@cse.unsw.edu.au

Abstract

This paper applies to quantum systems a modelling for the logic of knowledge, originally developed for reasoning about distributed systems, but since then applied to game theory, computer security and artificial intelligence. A formal model of quantum message passing systems is developed and the question of how one might define the semantics of a modal operator for knowledge in this model is considered. It is argued that there are at least two plausible semantics, depending on whether the agents are permitted to make use of their quantum state in determining what they know, and on whether one is dealing with single instances of quantum systems, or ensembles. The framework is illustrated using a number of examples from the quantum computing literature, including protocols for quantum key distribution and teleportation.

1 Introduction

The prospect that quantum computing and communications systems potentially have capabilities going beyond those of classical computing and communications is currently generating considerable interest in both computer science and physics. It takes only the most cursory inspection of one of the basic texts in the area [NC00] to find that epistemic locutions such as “agent A knows that” are common in the literature on quantum computing. The use of probability theory in quantum computing is highly sophisticated, but the literature has not developed a standard way to make such informal epistemic language precise. In this paper, we begin an investigation of the extent to which some well established tools of epistemic logic, in particular, a semantic approach for the logic of knowledge [HM90, FHMV95], developed for reasoning about distributed systems, can be applied to give a rigorous semantics for epistemic talk about quantum computing and communications systems. In doing so, we hope to be laying the foundations for formal (i.e., logical) methods for the epistemic analysis of areas such as quantum cryptographic protocols [Ben92], quantum distributed computing [BDHT99] and quantum games [EWL99].

Our main focus in this paper is the question of how one should define the semantics of the logic of knowledge in a quantum system. After a very brief review of quantum mechanics in Section 2, we define in Section 3 a formal model for multi-agent quantum systems in which the agents may communicate both by classical message passing and by exchanging quantum bits. We then turn, in Section 4, to the question of how to define the semantics of a modal logic of

Copyright held by the author

knowledge in such systems. We argue that there are at least two distinct notions of knowledge that make sense in quantum systems. Which one we use depends on the intuitions underlying the scenario we are modelling. One notion is appropriate for protocols implemented using a very fine-grained notion of quantum state (e.g. quantum bits implemented using single photons), and captures the states of information obtained by agents from specific measurements. Another is appropriate for reasoning about knowledge of *ensembles* of quantum states, and also captures the information that agents may *potentially* obtain by performing measurements. In Section 5, we illustrate the difference between these two notions of knowledge with some simple examples from the literature, which provide some intuition for the situations in which these notions are appropriate. Section 6 considers the relations between the two types of knowledge. Section 7 concludes with some discussion of related work and future directions.

2 Quantum Theory

We give a very brief introduction of quantum theory here, and recommend that a reader needing further background consult Chapter 2 of [NC00], whose conventions we follow closely.

According to quantum theory, the state of a physical system S , like an electron or an atom, is represented by a vector of length 1 in a Hilbert space, i.e, a complex vector space with an inner product $\langle \cdot | \cdot \rangle : H^2 \rightarrow \mathbf{C}$. We shall restrict ourselves to finite-dimensional spaces, which are adequate for quantum computing. It is conventional to use the Dirac notation $|x\rangle$ for a vector in H . The inner product induces an isomorphism between H and H^* , the dual space of linear functions $H \rightarrow \mathbf{C}$, where \mathbf{C} is the complex field. The image of a vector $|\alpha\rangle$ under this isomorphism is written $\langle \alpha |$, and is defined to be the function f such that $f(|\beta\rangle)$ is the inner product $\langle \alpha | \beta \rangle$ of $|\alpha\rangle$ and $|\beta\rangle$. Two vectors $|\alpha\rangle$ and $|\beta\rangle$ are said to be orthogonal if $\langle \alpha | \beta \rangle = 0$. An orthonormal basis in H is a basis of unit vectors such that every pair of distinct vectors in it are orthogonal. If the state of the system is $|\alpha\rangle$ then the probability of finding it in a state $|\beta\rangle$ is $|\langle \alpha | \beta \rangle|^2$. That is in any given state if the system is tested for another nonorthogonal state then there is a nonzero probability of success. The adjoint of a linear operator T on H is the (unique) linear operator T^\dagger on H such that $\langle T\alpha | \beta \rangle = \langle \alpha | T^\dagger\beta \rangle$.

Given two systems S_1 and S_2 with respective state spaces H_1 and H_2 , the composite system is described by the tensor product Hilbert space $H_1 \otimes H_2$. The simplest way to define it is through bases. Let $\{\alpha_i\}$ and $\{\beta_j\}$ be bases of H_1 and H_2 , respectively. Then the vectors $\alpha_i \otimes \beta_j$ form a basis of $H_1 \otimes H_2$. The dimension of the tensor product is mn if m and n are the dimensions of H_1 and H_2 . The inner product of $\alpha \otimes \beta$ and $\phi \otimes \psi$ is defined as $\langle \alpha | \phi \rangle \cdot \langle \beta | \psi \rangle$ and extended bilinearly to general vectors in $H_1 \otimes H_2$ making it a Hilbert space. Taking \otimes to be bi-linear allows us to obtain from linear operators T_1 on H_1 and T_2 on H_2 , a linear operator $T_1 \otimes T_2$ on $H_1 \otimes H_2$, defined by $(T_1 \otimes T_2)(\alpha \otimes \beta) = (T_1\alpha) \otimes (T_2\beta)$. Further, a general operator in $H_1 \otimes H_2$ can be written as a sum of such product operators.

Next we discuss the dynamics of quantum states. In the *general measurement* formulation we use, a quantum measurement is a finite sequence of operators $M = [M_1, \dots, M_k]$ on the state space H of the system. These operators are required to satisfy the equation $\sum_m M_m^\dagger M_m = I$, where I is the identity. Each index $m = 1, \dots, k$ of this collection corresponds to one of the possible outcomes of the measurement, occurring with probability $\langle \alpha | M_m^\dagger M_m | \alpha \rangle$ when the state being measured is $|\alpha\rangle$. Unlike a classical system, after a measurement the state of the system changes, being transformed from state $|\alpha\rangle$ to the state $\rho(M_m|\alpha)$, where ρ is the *renormalization* function that transforms a non-zero vector into the unit vector in the same direction.

Measurements typically correspond to projections onto a subspace of the Hilbert space, which *collapses*, or loses, a lot of the state information. A special case of this is measurement in a basis, where each M_m corresponds to a projection onto the one-dimensional subspace spanned by a basis element. Note, however, that another special case of measurement is a singleton $M = \{U\}$. Here U is a *unitary* operator U with $U^\dagger U = I$ the identity. (Thus, this approach captures both types of state transformation, projection and unitary transformation, usually treated separately.)

For quantum computation, we are interested in systems composed from *quantum bits*, which are represented by the two dimensional Hilbert space \mathcal{Q} , with a preferred orthogonal basis, called the *computational basis*, given by the two vectors $|0\rangle, |1\rangle$. An N -qubit Hilbert space is a 2^N -dimensional space \mathcal{Q}^N of the form $\mathcal{Q} \otimes \dots \otimes \mathcal{Q}$, where there are N copies of the space \mathcal{Q} in the tensor product. We write $|i_1, \dots, i_N\rangle$ for the vector $|i_1\rangle \otimes \dots \otimes |i_N\rangle$ in this space, where each i_j is either 0 or 1. The set of vectors $|i_1, \dots, i_N\rangle$ is a basis of \mathcal{Q}^N , called the *computational basis*.

3 Qubit Message Passing Environments

Many of the examples in quantum information theory deal with a computational setting in which the agents communicate by sending classical messages, by transmitting qubits and by operating on (possibly entangled) qubits. (Physically, such a system might be implemented by photons, transmitted between agents along optical fibers.) In this section, we define a class of formal models of such settings. We consider only the simplest possible model of communication: reliable synchronous communication.¹ We call a specific instance of our modelling an *environment*, denoted by \mathcal{E} . An environment specifies the number of agents n in the system, the classical and quantum states that the system can inhabit, the possible initial states of the system, the actions that the agents can perform, and the effect that these actions have on the state of the system. It may also specify an interpretation for some set of basic propositions.

3.1 States

The set of states, also called *global states*, of an environment will be a cartesian product $S = S^q \times S^c$, where S^q is a set of *quantum states* and S^c is a set of *classical states*. If s is a global state, we write s^q and s^c for the elements of S^q and S^c , respectively, such that $s = \langle s^q, s^c \rangle$. The set of quantum states S^q will be \mathcal{Q}^N for some finite number N , modelling the assumption that the agents are able to operate on N qubits. The classical states encode information such as classical bits held by the agents, the location of qubits in the system, measurement outcomes and classical message channels. We now describe each of these elements of the classical state.

To model the classical bits manipulated by the agents, we assume that for each agent i there is a set Var_i of variable names. A *classical bit assignment* is then a function var mapping each agent i to a function $\text{var}(i) : \text{Var}_i \rightarrow \{0, 1\}$ assigning a truth value to each variable of agent i .

In a qubit passing system, each qubit is thought of as being in the possession of some agent, but this agent may change from time to time, as an agent can send some of its qubits to another. We model this by taking a part of a classical state to represent the location of each qubit. Define a *qubit location assignment* in an N -qubit passing environment for n agents to be

¹To keep the model simple, we do not attempt to model adversaries, as would be required for a deeper study of quantum cryptographic protocols than we attempt here.

a function $\text{loc} : [0, N] \rightarrow [0, n]$. Intuitively, for a qubit number x , the value $\text{loc}(x)$ is the agent possessing qubit x .

We model reliable synchronous communication of classical messages between agents as follows. Let Msg be a set of messages. We assume that Msg contains the special value \perp representing a null message (used to represent the fact that no message was sent.) We assume that between each pair of agents i, j , there is a single channel for communications from i to j . (For brevity of exposition, we allow agents to transmit messages to themselves.) At each step of a computation, a single message in Msg may be transmitted along this channel. Thus, we take part of the classical state of a qubit passing system with n agents to be a function $\text{chan} : [1 \dots n]^2 \rightarrow \text{Msg}$. We call such a function a *channel value assignment*. Intuitively, the fact that $\text{chan}(i, j) = \mathbf{m}$ means that the message \mathbf{m} has just been sent from i to j . Initially, $\text{chan}(i, j) = \perp$ for all agents i, j .

Finally, the classical state of the environment is used to record the outcome of quantum operations performed by the agents. Suppose that $\text{loc}^{-1}(i) = \{i_1, \dots, i_k\}$ is the set of indices of qubits located at agent i . Then agent i is able to perform a general measurement on these k qubits. As discussed above, we represent a quantum operation on k qubits by a finite sequence of operators $M = [M_1, \dots, M_l]$, with each M_m operating on \mathcal{Q}^k . Suppose the agents simultaneously perform the quantum measurements M^1, \dots, M^n where each M^i is an measurement on the $k_i = |\text{loc}^{-1}(i)|$ bits located at agent i . Each operation M^i produces some outcome m_i , the index of some linear transformation $M_{m_i}^i \in M^i$, on \mathcal{Q}^{k_i} . We represent a combined outcome of these measurements by a function res from agents to outcomes, such that $\text{res}(i) = (M^i, m_i)$ for each agent i . We call such a function a *measurement result assignment*. Intuitively, each value (M^i, m_i) records the measurement performed and the outcome obtained.

We can now define the classical state of a qubit passing system as a $s^c = \langle \text{var}, \text{loc}, \text{chan}, \text{res} \rangle$ where var is a classical bit assignment, loc is qubit location assignment, chan is a channel value assignment, and res is a measurement result assignment.

This completes our description of the structure of the global states S in a qubit message passing environment \mathcal{E} . The environment also specifies a nonempty set I of global states, representing the possible initial configurations of the system. When the environment also specifies an interpretation function $\pi : S \times \text{Prop} \rightarrow \{0, 1\}$ for some set of propositions Prop , we say that the environment is *interpreted*. As quantum states differing by a constant factor $z \in \mathbb{C}$ with $|z| = 1$ are considered to be identical, we require that $\pi(\langle z \cdot s^q, s^c \rangle, p) = \pi(\langle s^q, s^c \rangle, p)$ for all global states $\langle s^q, s^c \rangle$ and propositions p .

3.2 Actions

We now describe the actions that agents are able to perform in a qubit message passing environment, and their effect on the global states.

For each classical bit variable $v \in \text{Var}_i$, and boolean value $b \in \{0, 1\}$ the agent has an action “ $v := b$ ”, the effect of which is to change $\text{var}(i)(v)$ to have value b . Additionally, the agent is able to perform an action $\text{flip}(v)$ on the bit v , the effect of which is to change $\text{var}(i)(v)$ to have value either 0 or 1, with equal probability. This models a fair coin flip.

In order to transmit qubits, agent i is able to perform the action $\text{transmit}(b, j)$, where b is a qubit index in $\text{loc}^{-1}(i)$. The effect of this action is to change the value of $\text{loc}(b)$ to equal j . (Note that this does not necessarily mean that j knows the value of qubit b .)

In order to transmit classical messages, agent i has an action $\text{send}_{i,j}(\mathbf{m})$, where j is an agent and $\mathbf{m} \in \text{Msg}$. The effect of agent i performing this action is to set $\text{chan}(i, j) = \mathbf{m}$ in the next

state. For all $j' \neq j$, we have $\text{chan}(i, j') = \perp$ in the next state. Moreover, if agent i performs any other action than one of the form $\text{send}_{i,j}(\mathbf{m})$, we have $\text{chan}(i, j) = \perp$ in the next state for all agents j .

Finally, agents have the ability to perform measurements on the qubits in their possession. To capture the combined effect when each agent performs such a measurement, it is convenient to introduce some notation for shuffling the order of qubits. Given a qubit location assignment loc , we define a unitary transformation U_{loc} as follows. For each agent i , let $\text{loc}^*(i)$ be the sequence i_1, \dots, i_{k_i} where $\text{loc}^{-1}(i)$ contains the values $i_1 < \dots < i_{k_i}$. For each $i = 1 \dots N$, define $p(i)$ to be the i -th element of the sequence $1 \dots N$ to $\text{loc}^*(1) \dots \text{loc}^*(n)$. The function p is a permutation on $1 \dots N$. We define U_{loc} by describing how it acts on each element $|b_1 \dots b_N\rangle$ of the computational basis of \mathcal{Q}^N , viz., $U_{\text{loc}}|b_1 \dots b_N\rangle = |b_{p(1)} \dots b_{p(N)}\rangle$. Intuitively, effect of U_{loc} is to shuffle the order of the components of the components of the tensor product of qubit Hilbert spaces, to ensure that all the qubits located at agent 1 occur first, then the qubits located at agent 2, etc.

The combined outcome m_1, \dots, m_n of measurements M^1, \dots, M^n corresponds to a state transformation T on the quantum state $|\psi\rangle \in \mathcal{Q}^N$ of the system, given by $T = U_{\text{loc}}^{-1}(M_{m_1}^1 \otimes \dots \otimes M_{m_n}^n)U_{\text{loc}}$. In state $|\psi\rangle$, this transformation occurs with probability $\langle \psi | T^\dagger T | \psi \rangle$, and transforms the quantum state of the system from $|\psi\rangle$ to $\rho(T|\psi)$, where ρ is the renormalization function.

A *joint action* in an environment \mathcal{E} is a tuple $\langle a_1, \dots, a_n \rangle$ where each a_i is an action of agent i . We say that there is a *transition* of \mathcal{E} from global state s to global state t if one of the possible combined effects of the agents performing the actions a_1, \dots, a_n , as described above, is to transform the state s into the state t , with non-zero probability.

3.3 Protocols

Define a *run* to be a function $r : \mathbb{N} \rightarrow S$ mapping the natural numbers to the set of global states S of a qubit passing environment \mathcal{E} . A *point* is a pair (r, m) consisting of a run r and a time m in the natural numbers. Intuitively, a run describes a potential evolution of the system, with $r(m)$ representing the global state of the system at time m . We will consider systems comprised of specific sets of runs, generated by the agents engaging in a particular pattern of behaviour, which we call a *protocol*.

Intuitively, at each each point of time, an agent has acquired some information, which it uses to make a decision concerning its next action. We may model the information that an agent acquires when the environment is in a particular global state s by means of a function O_i mapping global states to some set of *observations* \mathcal{O}_i . We defer discussion of some possible definitions of O_i to the next section. We may now represent the information that the agent has acquired after a number of steps of computation by a sequence of observations, i.e. an element of \mathcal{O}_i^+ . Thus, we represent a protocol for agent i by a function $P : \mathcal{O}_i^+ \rightarrow \text{Act}_i$. A *joint protocol* is a tuple $\mathbf{P} = \langle P_1, \dots, P_n \rangle$, where each P_i is a protocol for agent i .²

Define the *perfect recall local state* of agent i at a point (r, m) , denoted $r_i(m)$, to be the sequence of observations made by the agent to time m in r , i.e., $r_i(m) = O_i(r(0)), \dots, O_i(r(m))$.

²Note that our definitions make protocols *deterministic*, but the effect of an action (e.g. a coin flip or a measurement) is not deterministic. We could also consider non-deterministic and probabilistic protocols, but will not attempt this here. It would be natural to require the protocol to be a computable function, but we do not require this in the present paper, in order to focus on the purely information theoretic aspects of our framework. Amongst the issues to be faced in formalizing computability is the fact that we have an uncountable set of possible quantum measurements that the agent can perform.

We say that a run r is a *run of the joint protocol* $\mathbf{P} = \langle P_1, \dots, P_n \rangle$ in the environment \mathcal{E} if for each time m , there exists a transition of \mathcal{E} on the joint action $\langle P_1(r_1(m)), \dots, P_n(r_n(m)) \rangle$ from global state $r(m)$ to global state $r(m+1)$. We write $\mathcal{R}(\mathcal{E}, \mathbf{P})$ for the set of all runs of the joint protocol \mathbf{P} in \mathcal{E} .

4 A Logic for Knowledge and Time in Quantum Systems

We now define a modal logic that can be used to formally state claims about knowledge and time in a quantum system. We define the set of formulas of the logic to be the smallest set such that each basic proposition p in Prop is a formula, and if ϕ_1 and ϕ_2 are formulas, then so are $\neg\phi_1$, $\phi_1 \wedge \phi_2$, as are $K_i^c\phi_1$ and $K_i^q\phi_1$, for each agent i , and $\Box\phi_1$ and $\text{init}(\phi_1)$. We use the customary abbreviations of propositional logic. The formula $K_i^c\phi$ is read “agent i classically knows ϕ ”, and the formula $K_i^q\phi$ is read “agent i quantumly knows ϕ ”. We provide intuitions for these constructs shortly. The formula $\Box\phi$ is read “at all times in the future, ϕ ”, and $\text{init}(\phi)$ is read “initially ϕ ”. (Other temporal operators could be added, but these two suffice for the examples we consider.)

To give a semantics to this language, we evaluate formulas with respect to triples $\mathcal{E}, \mathbf{P}, (r, m)$ where \mathcal{E} is an interpreted qubit passing environment, \mathbf{P} is a joint protocol for \mathcal{E} and (r, m) is a point on a run $r \in \mathcal{R}(\mathcal{E}, \mathbf{P})$. In order to interpret the two knowledge operators, we will define, for each agent i , two equivalence relations on points: \sim_i^c and \sim_i^q , which, intuitively, capture different notions of two points being *indistinguishable* to the agent. Let π be the interpretation function of \mathcal{E} . The semantics is then given in the familiar pattern for the logic of knowledge and time, by the recursion

1. $\mathcal{E}, \mathbf{P}, (r, m) \models p$ if $\pi(r(m), p) = 1$, when $p \in \text{Prop}$;
2. $\mathcal{E}, \mathbf{P}, (r, m) \models K_i^c\phi$ if $\mathcal{E}, \mathbf{P}, (r', m') \models \phi$ for all points (r', m') of $\mathcal{R}(\mathcal{E}, \mathbf{P})$ such that $(r, m) \sim_i^c (r', m')$;
3. $\mathcal{E}, \mathbf{P}, (r, m) \models K_i^q\phi$ if $\mathcal{E}, \mathbf{P}, (r', m') \models \phi$ for all runs r' of $\mathcal{R}(\mathcal{E}, \mathbf{P})$ such that $(r, m) \sim_i^q (r', m')$;
4. $\mathcal{E}, \mathbf{P}, (r, m) \models \Box\phi$ if $\mathcal{E}, \mathbf{P}, (r, m') \models \phi$ for all $m' \geq m$;
5. $\mathcal{E}, \mathbf{P}, (r, m) \models \text{init}(\phi)$ if $\mathcal{E}, \mathbf{P}, (r, 0) \models \phi$.

plus the obvious clauses for the propositional operators.

Recall that we defined the notion of an agent’s *perfect recall local state* $r_i(m)$ as the sequence of observations the agent has made in reaching the point (r, m) . Using this notion, we may define an equivalence relation \sim_i on points by $(r, m) \sim_i (r', m')$ if $r_i(m) = r'_i(m')$.³ The definitions of the equivalence relations \sim_i^c and \sim_i^q arise by using different definitions of the observation function O_i in this definition.

The first notion of observation we consider corresponds to what an agent is able to learn from that portion of the *classical state* of the system that it is able to control. Given a global state $s = (s^q, s^c)$, where the classical state $s^c = (\text{var}, \text{loc}, \text{chan}, \text{res})$, we can define the observation

³Perfect recall corresponds to an optimal notion of knowledge that suffices for the points we wish to make. We note, however, that other definitions (e.g. assuming that agents have no recall of their past observations) would be equally sound, depending on the application at hand.

of agent i by $O_i^c(s) = \langle \text{var}(i), \text{loc}^{-1}(i), \text{chan}(i), \text{res}(i) \rangle$. Intuitively, the presence of $\text{var}(i)$ means that the agent is aware of the values of its classical bits. The presence of $\text{loc}^{-1}(i)$ in the observation means that the agent is aware of which qubits it possesses. The presence of $\text{chan}(i)$ means that the agent is aware of what messages it has just received from other agents. Finally, the presence of $\text{res}(i)$ models the fact that the agent is aware of the outcome of the measurement that it has just performed on the qubits in its possession. We write \sim_i^c for the notion of indistinguishability derived from the definitions above when we use the function O_i^c for the observation function.

This notion of indistinguishability assumes that while the agent makes use of the past and present values of the classical bits in its possession, its knowledge about which qubits it possess or has possessed, and the past outcomes of measurements it has performed on these qubits, it does not make use of the “values” of the quantum bits in its possession. It makes intuitive sense to impose such a restriction, since any attempt to obtain such additional information would perturb the quantum state, leaving the system in a *different state*. (Moreover, when some of the agent’s qubits are entangled with the qubits of another, the perturbation is not local to the agent!)

On the other hand, it is possible to provide intuitive justifications for allowing an agent to make use of its local qubits to determine what it knows. One way to justify this is to note that the classical notion of knowledge, familiar in the literature on epistemic logic, is an information theoretic idealization, inasmuch as $K_i\phi$ does not state that agent i is in a position to perform a computation on its local state which establishes that ϕ holds. Rather, it says that the agent *could*, in principle, decide ϕ based on its local state, were it to be given *unlimited* computational resources, extending even beyond recursive enumerability. (Consider, e.g., the (valid) formula stating that the agent knows the truth or falsity of Goldbach’s conjecture.) From this perspective, it is reasonable to consider similarly idealized notions of knowledge, that take into account what an agent *could* determine from its local state, including qubits, given sufficient powers.

A further observation that lends credence to the idea that an agent should be able to make use of its qubits in determining what it knows is that individual quantum systems, such as single photons, are extremely difficult to isolate in the laboratory. Instead, one typically deals in practice with *ensembles*, consisting of many instances of the physical system (e.g. packets consisting of multiple photons). It is possible to prepare such ensembles in such a way that each element is in the same quantum state (e.g. each photon in the packet is polarised in the same direction.) Under unitary transformations on the system, this identity of state for members of the ensemble is preserved. This makes it possible to treat the ensemble like a single copy of the system. For measurements in general, however, this identity is not preserved. Different members of the ensemble will yield different measurement outcomes, yielding a probability distribution (a spectrum) on outcomes of the measurement, rather than a single outcome. Moreover, when dealing with an ensemble, there is a sense in which one can measure the quantum state while leaving it intact. This is to perform the measurement on a subset of the ensemble, leaving the remainder in the original state. Indeed, given a sufficiently large ensemble, this approach allows one to make multiple measurements.

These considerations lead us to consider a different observation function O_i^q that takes into account the possible measurements on the qubits possessed by an agent. Suppose we are given a global state with quantum state component $|\psi\rangle$ and a measurement M performable by agent i , i.e., a measurement on the qubits $\text{loc}^{-1}(i)$ local to i in the global state. Define $d_{i,|\psi\rangle}(M)$ to be the distribution over the possible outcomes of M . (That is, if M_m is the operator corresponding

to a possible outcome then the probability associated to that outcome is $\langle \psi | M_m^\dagger M_m | \psi \rangle$.) By performing the measurement often enough, the agent is able to estimate these probabilities. This models what an agent is able to observe by making a single measurement. To capture what the agent is able to obtain from *all* measurements it is able to perform on the state, we treat $d_{i,|\psi\rangle}$ as a function mapping the possible measurements on qubits $\text{loc}^{-1}(i)$ to their outcome distributions.

We now define $O_i^q(s)$ as the pair $\langle O_i^c(s), d_{i,s^q} \rangle$. That is, we make both the classical observation and the results of all possible quantum measurements that the agent could perform observable to the agent. We write $r_i^q(m)$ and \sim_i^q for, respectively, the notion of perfect recall local state, and indistinguishability, obtained when $O_i = O_i^q$. That is, the two points are \sim_i^q -indistinguishable if *both* the sequence of classically observed information obtained by agent i in the past, *and* the sequence of all possible information it could gather by quantum measurement are the same in the two points.⁴

There is an aspect of our modelling for which we need to account when we interpret quantum states as representing ensembles: we have said that the result of a measurement (as encoded in a measurement result assignment) is a *single outcome* rather than a distribution, and this does not seem to fit the ensemble interpretation. One way to reconcile the two is to view a measurement of an ensemble as randomly selecting a single copy of the system from the ensemble, and performing the measurement on this. Indeed, this is what is done in implementations of quantum cryptography, where photon packets are attenuated to a single photon before measurement [IY94]. However, we will see from the examples in the next section that this “selection” interpretation of our modelling is not appropriate for all protocols, and return to this point below.

5 Examples

To illustrate our definitions, we now formalise a number of examples from the quantum computing literature using our framework. We use the following technical notion. If ϕ is a formula of our logic, and \mathbf{P} is a joint protocol and \mathcal{E} is an interpreted environment, then we say that \mathbf{P} *realizes* ϕ in \mathcal{E} if for every run $r \in \mathcal{R}(\mathcal{E}, \mathbf{P})$, we have $\mathcal{E}, \mathbf{P}, (r, 0) \models \phi$. That is, the protocol ensures that ϕ holds at the initial point of every run in the system generated by the protocol.

5.1 Distinguishability of states

Our first example concerns what a single agent is able to learn about a quantum state, which may initially have one of two possible values. Consider an environment for the agent in which $S^q = \mathcal{Q}^N$, and in which the agent has no local classical bits, and possesses all qubits, i.e., $\text{loc}(j) = 1$ for all $j = 1 \dots N$. We remove the message channels from the definition of the environment since messages sent by the agent to itself are uninformative. For the initial states l of the environment, take the two (distinct) global states s_1 and s_2 in which the quantum portion is $|\psi_1\rangle \in \mathcal{Q}^N$ or $|\psi_2\rangle \in \mathcal{Q}^N$, respectively. Note that the classical portions of the state are uniquely determined by the definitions above, and identical in s_1 and s_2 . Let π be the interpretation of the propositions p_j , for $j = 1, 2$, such that $\pi(s_k, p_j) = 1$ iff $k = j$. That is, p_j

⁴We note that identity of d_{i,s^q} and d_{i,t^q} quantifies over the uncountable set of all possible measurements, but quantum theory has some finely honed tools, (which will be the more familiar notion to the reader versed in quantum theory than the above), viz., density operators and the *partial trace*, whereby this identity can be very conveniently expressed.

expresses that the global state is s_j . We write $\mathcal{E}_{|\psi_1\rangle,|\psi_2\rangle}$ for the interpreted environment obtained from the choice of $|\psi_1\rangle, |\psi_2\rangle$ in this definition. In this setting, the agent cannot transmit qubits, and protocols describe a strategy for performing measurements on the quantum state. The strategy may be adaptive, i.e., depend on the outcome of measurements already performed.

Proposition 1 *Let \mathbf{P} be any protocol for $\mathcal{E}_{|\psi_1\rangle,|\psi_2\rangle}$. Then for $j = 1, 2$,*

1. \mathbf{P} realizes $p_j \Rightarrow \neg K^c(p_j) \wedge K^q(p_j)$ in $\mathcal{E}_{|\psi_1\rangle,|\psi_2\rangle}$, and
2. if $|\psi_1\rangle, |\psi_2\rangle$ are not orthogonal, then \mathbf{P} realizes $p_j \Rightarrow \Box \neg K^c(\text{init}(p_j))$ in $\mathcal{E}_{|\psi_1\rangle,|\psi_2\rangle}$.

On the other hand, if $|\psi_1\rangle, |\psi_2\rangle$ are orthogonal, then there exists a protocol \mathbf{P} that realizes the formula $\bigwedge_{j=1,2} p_j \Rightarrow \Diamond K^c(\text{init}(p_j))$ in $\mathcal{E}_{|\psi_1\rangle,|\psi_2\rangle}$.

That is, initially, the agent does not know which state it is in based on its classical information alone, but using our idealized notion of quantum knowledge K^q , the agent is able to distinguish whether it is in state s_1 or state s_2 . If the states are non-orthogonal, then no measurement protocol will give it classical knowledge of the initial state. On the other hand, when the states are orthogonal, it is able learn the initial state (based on its classical observations) by some measurement protocol. (In fact, a single measurement, in a basis containing $|\psi_1\rangle, |\psi_2\rangle$, suffices.)

5.2 A quantum key distribution protocol

Our next example concerns the key ingredient of a quantum cryptography protocol [Ben92], the objective of which is to establish a shared secret key (a finite sequence of classical bits) between two parties (Alice and Bob), who can only communicate via a channel controlled by an eavesdropper (Eve). The protocol itself is complex, using a variety of mechanisms for error correction and for detecting interference by Eve. We consider just a fragment (which we call B92) that shows how a single shared classical bit is established. We simplify Eve's capabilities by eliminating Eve's ability to interfere with message transmissions. Because of space limitations in this abstract, our presentation of even this simplified fragment is necessarily sketchy: the full paper will contain a more complete description.

The protocol assumes that there is a single qubit, initially possessed by Alice. Additionally, Alice has a classical bit a and Bob has two classical bits a' and b . The protocol involves two orthonormal bases for \mathcal{Q} : the computational basis $|0\rangle, |1\rangle$ and the basis consisting of $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$.

The protocol proceeds as follows. First, Alice flips her classical bit a . If the outcome is 0, she prepares the qubit in the state $|0\rangle$ (this can be done by performing a measurement and applying a unitary transformation if the desired outcome $|0\rangle$ is not produced as a result). If the outcome is 1, she prepares the qubit in the state $|+\rangle$. Alice then transmits the qubit to Bob. Upon receipt of the qubit, Bob flips his classical bit a' . If the outcome is 0 he measures the qubit received in the computational basis, otherwise, he measures it in the basis $|+\rangle, |-\rangle$. If the measurement result was to transform the qubit to first element of the basis measured, Bob sets his classical bit b to 0, otherwise he sets it to 1. Bob then sends a classical message to Alice stating the value of b .

Not all runs of the protocol are considered successful, only those in which $b = 1$ after the above steps have been performed. In these runs, it can be shown that $a = 1 - a'$ in the final

state reached, which we may consider as Alice and Bob having reached agreement on the value of a .

In order to model Eve's ability to eavesdrop on this protocol, we "double up" all messages and qubit transmissions. That is, we have Alice prepare two qubits in the same state, sending one to Bob and one to Eve. Similarly, we have Bob send his classical message both to Alice and to Eve. Let \mathcal{E} be an environment suitably equipped for this doubled-up version. We assume all classical bits are initially equal to 0. We say that \mathbf{P} is an *eavesdropping version of B92* in \mathcal{E} if it prescribes that Alice and Bob behave as discussed above, and Eve waits to receive the copy of the qubit transmitted to Bob and Bob's classical message, and then runs some protocol in which she performs measurements on her copy of the qubit. (Eve is not allowed to engage in any communication with Alice and Bob in this stage of her protocol.) Write $k_i^x(a)$ for $K_i^x(a=0) \vee K_i^x(a=1)$, i.e., agent i knows the value of the classical bit a , where x is either c or q . Then we can show the following.

Proposition 2 *If \mathbf{P} is an eavesdropping version of B92 in \mathcal{E} then \mathbf{P} realizes the following formulas in \mathcal{E} :*

1. $\Box(b = 1 \Rightarrow (k_A^c(a) \wedge k_B^c(a)))$
2. $\Box(b = 1 \Rightarrow \neg k_E^c(a))$
3. $\Box(b = 1 \Rightarrow k_E^q(a))$

The first formula states that in successful runs, Alice and Bob come to classically know the value of bit a . The second formula says that Eve never comes to know the value of this bit, based on her classical observations alone. However, the final formula states that, were we to allow Eve repeatable quantum measurements on the qubit (which corresponds to K_E^q), she could come to know the value of the bit a (indeed she would know it as soon as she receives her copy of the qubit from Alice.) This relates to known fact that the security of the protocol relies on its being implemented using single photons. However, because of present limitations on our ability to control individual photons, in practical implementations, what is transmitted is not a single photon, but an ensemble (packet) of photons. This allows Eve to undetectably shave off a part of the ensemble and perform some measurements. The protocol is not secure in such a setting! (See [BBB⁺92] for a discussion of how such problems are overcome in another protocol).

5.3 Teleportation

Our final example is a protocol of Bennet et al [BBC⁺93], that illustrates a quite astounding difference in the dynamics of knowledge in classical communications protocols and quantum communications protocols.

To state this difference, we need some technical notions. Define two runs r, r' of a joint protocol $\mathbf{P} = \langle P_1, \dots, P_n \rangle$ in an environment \mathcal{E} to be *communication indistinguishable* if for all times m and agents i , either $P_i(r_i(m)) = P_i(r'_i(m))$, or both $P_i(r_i(m))$ and $P_i(r'_i(m))$ are not communications actions (i.e., neither is a send or transmit). Intuitively, this means that the agents perform exactly the same communications actions in r and r' .

Say that a joint protocol \mathbf{P} is *communication independent of the initial state* in an environment \mathcal{E} if for all runs $r \in \mathcal{R}(\mathcal{E}, \mathbf{P})$, and all initial global states s of \mathcal{E} , there exists a run $r' \in \mathcal{R}(\mathcal{E}, \mathbf{P})$ such that $r'(0) = s$ and r and r' are communications indistinguishable.

We first note a property of classical communications protocols. Say \mathcal{E} is a *classical message passing environment* if it is a qubit message passing environment with zero qubits, i.e., in which agents can communicate only by sending classical messages. Then we have the following result. (Recall that the truth value of atomic propositions is a function of the global state, and that $k_i^x(\phi)$ means that i knows whether ϕ is true or not.)

Proposition 3 *If \mathcal{E} is a classical message passing environment then there is no protocol \mathbf{P} for \mathcal{E} that is communication independent of the initial state and realizes the formula $k_A^c p \wedge \neg k_B^c p \wedge \diamond k_B^c \text{init}(p)$, where p is an atomic proposition.*

A corresponding proposition is false for qubit message passing environments. There is a protocol \mathbf{P} for a qubit message passing environment \mathcal{E} that is communication independent of the initial state and realizes the formula $k_A^q p \wedge \neg k_B^q p \wedge \diamond k_B^q \text{init}(p)$ where p is an atomic proposition. So stated, this is not quite as surprising as it may look. For suppose that A initially possesses qubit 1, which is in state $|\psi_0\rangle$ or in state $|\psi_1\rangle$. We may realize the formula simply by performing the action $\text{transmit}(1, B)$ as the first step of A 's protocol, and taking p to mean that qubit 1 is in state $|\psi_0\rangle$. But what is truly amazing is that there exists a protocol independent of the initial state in which the *only* communications actions are classical send operations! That is, by sending only classical messages that depend in no way on the initial qubit value, A may "inform" B of the initial state of this qubit. Indeed, by sending exactly two classical bits, A may transmit to B a continuum of possible initial values of the qubit! The protocol that achieves this is known as the *teleportation* protocol [BBC⁺93]. The trick it uses to achieve the result is to exploit *entanglement*, which is intuitively a type of correlation of information between qubits. The state $(|00\rangle + |11\rangle)/\sqrt{2}$ of a two-qubit system is an example of an entangled state. If, in this state, one or both of the two qubits are measured in the computational basis, we get a collapse to $|00\rangle$ or $|11\rangle$ with equal probability, even if the qubits are far apart. The famous Einstein-Podolski-Rosen argument about the adequacy of quantum theory was based on this correlation and the apparent faster-than-light communication that it involves, but this behaviour has been experimentally verified, and entanglement is now understood as one of the fundamental sources of the power of quantum computing.

The teleportation protocol uses entanglement in a system with three qubits, with A in possession of qubits 1 and 2 throughout the protocol, and B in possession of qubit 3 throughout the protocol. Qubit 1 initially has any value $|\psi\rangle = x|0\rangle + y|1\rangle$, and qubits 2 and 3 are initially in the entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$. That is, the initial states are of the form $|\psi\rangle \otimes (|00\rangle + |11\rangle)/\sqrt{2}$. At the end of the protocol, B 's qubit has the value $|\psi\rangle$, even though the only operations that have been performed are local measurement and transmission of classical messages. The overall structure of the protocol is as follows: first, A performs a joint measurement on her qubits 1 and 2, and gets one of 4 possible results. She sends her result to B classically (encoded in two classical bits). Finally, B performs a measurement on his bit, depending on the message received from A . The effect of this is to transform the state of B 's qubit to the original value of A 's qubit.

The measurement A performs on bits 1 and 2 is a measurement in the *Bell basis*, which consists of the vectors: $(|10\rangle + |01\rangle)/\sqrt{2}$, $(|10\rangle - |01\rangle)/\sqrt{2}$, $(|11\rangle + |00\rangle)/\sqrt{2}$ and $(|11\rangle - |00\rangle)/\sqrt{2}$. When the outcome corresponding to the first measurement is obtained, the state is transformed to $(|10\rangle + |01\rangle)/\sqrt{2} \otimes (x|0\rangle + y|1\rangle)$. The second outcome transforms the state to $(|10\rangle - |01\rangle)/\sqrt{2} \otimes (x|0\rangle - y|1\rangle)$. We see that in these resulting states, B 's qubit is unentangled with A 's qubits, but that its value is an unitary function of the original value of qubit 1. Similar forms are

obtained for the other two measurement results, and the unitary function transforming the original value of qubit 1 to the value of B 's qubit corresponds one to one to the measurement outcome obtained. Thus, once B receives Alice's two bit message describing which measurement outcome was obtained, B can reconstruct the original value of qubit 1 by applying the inverse of the corresponding operation to recover $|\psi\rangle$. Taking the proposition p to mean "the state of A 's qubit 1 is $|\psi_0\rangle$ " for some fixed value $|\psi_0\rangle$, we see that the formula $k_A^q p \wedge \neg k_B^q p \wedge \diamond k_B^q \text{init}(p)$ is realized.

The fact that B 's qubit changes immediately after A 's measurement makes it appear that there has been faster than light communication from A to B . This nonlocal behaviour is another unusual characteristic of quantum theory. Certainly, there is not any communication from A to B in the classical sense. Before receiving the two bit message, B does not classically know which of the four possible states the system is in, since B does not observe the measurement outcome. However, in our richer sense of knowledge, B knows the value of its qubit. This indicates that we need to understand this notion of knowledge as a kind of potential knowledge, or upper bound on attainable knowledge, in order for it to be consistent with Einstein's theory of relativity.

While we have discussed both classical and quantum knowledge in the setting of the teleportation protocol, we note that it raises some difficulties for the ensemble interpretation of states and the "selection" interpretation of measurement discussed above. One way to make sense of this interpretation for teleportation is that A selects a single copy from the ensemble $|\psi\rangle$ for measurement in the first step, and that A and B share a single entangled pair (rather than an ensemble). But then B has a single copy of $|\psi\rangle$ in the final state of the protocol, which weakens our ensemble justification for the use of K_B^q in this state. Other interpretations similarly lead to difficulties. We discuss this issue at greater length in the full version of the paper.

6 Quantum Knowledge and Potential Knowledge

Can we relate the two types of knowledge? It is plain from the definitions that for every formula ϕ , the formula $K_i^c \phi \Rightarrow K_i^q \phi$ is valid, hence also the contrapositive $\neg K_i^q \phi \Rightarrow \neg K_i^c \phi$. That is, if an agent does not quantumly know ϕ , then it does not know ϕ based on its classical information alone.

In fact, in some cases we can make a stronger statement than this: if ϕ is a formula that talks only about the current state in the system, and an agent does not quantumly know ϕ , then no amount of measurement can bring about a situation where the agent knows that ϕ was the state of affairs that held initially. That is, quantum knowledge provides an upper bound on the information that the agent is able to obtain by measurement alone. In order to obtain more information than this, the agent needs to communicate with other agents.

These claims can be made precise as follows. Define a protocol P to be a *measurement protocol* for agent i if for all sequences σ of observations for agent i , we have that $P(\sigma)$ is a measurement operation on the set of qubits $\text{loc}^{-1}(i)$ indicated by the final observation in σ to be local to the agent running the protocol. Define a protocol P to be *trivial*, if for all if for all sequences σ of observations, we have $P(\sigma)$ equal to the null measurement $\{I_S\}$, where I_S is the identity operator over the Hilbert space associated to the set of qubits $S = \text{loc}^{-1}(i)$ local to agent i , (as indicated by the final observation in σ .) Say that a *joint* protocol $\langle P_1, \dots, P_n \rangle$ is a *joint measurement protocol for agent i* if P_i is a measurement protocol and P_j is trivial for all agents $j \neq i$. Note that in a joint measurement protocol for agent i , the set of qubits local to

any agent is invariant.

Then we can show the following:

Proposition 4 *If ϕ is a formula that depends only on the current state and \mathbf{P} is a joint measurement protocol for agent i in environment \mathcal{E} , then \mathbf{P} realizes the following formulas in \mathcal{E} :*

1. $\neg K_i^q \phi \Rightarrow \Box \neg K_i^q \text{init}(\phi)$, and
2. $\neg K_i^q \phi \Rightarrow \Box \neg K_i^c \text{init}(\phi)$.

That is, if the agent would not know ϕ even with complete information about the statistics of all possible single measurements, then there is no measurement protocol it could run that could ever result in its knowing that the initial state satisfied ϕ , in either the classical or the quantum sense of knowledge. This shows in a strong sense that quantum knowledge captures everything that the agent can learn by measurement alone.

Note that the result for the second formula is immediate from that for the first, by the fact that $\neg K_i^q \phi \Rightarrow \neg K_i^c \phi$ is valid. The proof of the result for the first formula involves showing that general measurements are sufficiently expressive to capture deterministic sequences of general measurements, in the sense that if $[M_1, \dots, M_k]$ and $[N_1, \dots, N_l]$ are two general measurements then the composition $[M_1 N_1, \dots, M_i N_j, \dots, M_k N_l]$, with outcomes equal to pairs (i, j) , is a general measurement. This makes the result seem to depend on the use of the general measurement formalism. However, the same result can be shown for a more restricted model in which we allow unitary operations and projective measurements in a single measurement step only.

7 Conclusion

The question of how one should interpret the formalism of quantum mechanics has been the subject of considerable debate. We have taken a very pragmatic approach, treating the quantum state as a physical state predictive of measurement outcomes. We note that a recent strand of work in the interpretation of quantum mechanics [CFS02] has begun to explore the question of whether quantum states themselves can be understood as the states of knowledge of some agent. We have also been cavalier in mixing classical and quantum states. The question of the boundary between classical and quantum is a deep one that has engendered acrimonious debates. If, ultimately, the physical world is a purely quantum system, one might wish to eliminate our use of classical state. There have been attempts in the literature on quantum foundations to explain the apparent stability of the classical world from the perspective of quantum mechanics, and there is an example known as “Wigner’s friend” that deals with epistemic concerns in such attempts [Wig62]. We will further discuss this literature in the full paper.

We have been primarily concerned with laying down and justifying some definitions in this paper, but having done so, a vast number of interesting questions open up. We list a few here. Can one develop a theory of automated verification of knowledge assertions in quantum protocols? What do notions like common knowledge and distributed knowledge from the literature on knowledge in distributed systems mean in a quantum setting? How should the combination of the logic of knowledge and probability be defined in quantum systems? Can quantum knowledge be justified as a limiting Bayesian notion of classical knowledge? Elsewhere, we have already begun to develop a logic for reasoning about quantum probabilities [MP03], but much remains to be done to answer such questions.

References

- [BBB⁺92] C. H. Bennet, F. Bessette, G. Brassard, L. Salvail, and J. Smolin. Experimental quantum cryptography. *Journal of Cryptology*, 5(1):3–28, 1992.
- [BBC⁺93] C. H. Bennet, G. Brassard, C. Crepeau, R. Josza, A. Peres, and W. K. Wootters. Teleporting unknown quantum states via dual classical and EPR chnnels. *Physical Review Letters*, 70(13):1895–1899, 1993.
- [BDHT99] H. Buhrman, W. van Dam, P. Hoyer, and A. Tapp. Multiparty quantum communication complexity. *Physical Review A*, 60(4):2737–2741, 1999.
- [Ben92] C. H. Bennet. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68(21):3121–3124, 1992.
- [CFS02] C. M. Caves, C. A. Fuchs, and R. Schack. Quantum probabilities as Bayesian probabilities. *Physical Review A*, 65:022305, 2002.
- [EWL99] J. Eisert, M. Wilkens, and M. Lewenstein. Quantum games and quantum strategies. *Physical Review Letters*, 83:3077–3088, 1999.
- [FHMV95] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge, Mass., 1995.
- [HM90] J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.
- [IY94] A. Imamoglu and Y. Yamamoto. Turnstile device for heralded single photons: Coulomb blockade of electron and hole tunneling in quantum confined p-i-n heterojunctions. *Physical Review Letters*, 72(2):210–213, 1994.
- [MP03] R. van der Meyden and M. Patra. A logic for probability in quantum systems. submitted for publication, 2003.
- [NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. CUP, Cambridge, UK., 2000.
- [Wig62] E.P. Wigner. Remarks on the mind-body question. In I.J Good, editor, *The Scientist Speculates*, pages 284–301. Heinemann, London, 1962.