

## Simple example of definitions of truth, validity, consistency, and completeness in quantum mechanics

Paul Benioff\*

*Physics Division, Argonne National Laboratory, Argonne, Illinois 60439*

(Received 19 November 1998)

Besides their use for efficient computation, quantum computers and quantum robots form a base for studying quantum systems that create valid physical theories using mathematics and physics. If quantum mechanics is universally applicable, then quantum mechanics must describe its own validation by these quantum systems. An essential part of this process is the development of a coherent theory of mathematics and quantum-mechanics together. It is expected that such a theory will include a coherent combination of mathematical logical concepts with quantum mechanics. That this might be possible is shown here by defining truth, validity, consistency, and completeness for a quantum-mechanical version of a simple (classical) expression enumeration machine described by Smullyan. Some of the expressions are chosen as sentences denoting the presence or absence of other expressions in the enumeration. Two of the sentences are self-referential. It is seen that, for an interpretation based on a Feynman path sum over expression paths, truth, consistency, and completeness for the quantum system have different properties than for the classical system. For instance, the truth of a sentence  $S$  is defined only on those paths containing  $S$ . It is undefined elsewhere. Also  $S$  and its negation can both be true provided they appear on separate paths. This satisfies the definition of consistency. The definitions of validity and completeness connect the dynamics of the system to the truth of the sentences. It is proved that validity implies consistency. It is seen that the requirements of validity and maximal completeness strongly restrict the allowable dynamics for the quantum system. Aspects of the existence of a valid, maximally complete dynamics are discussed. An exponentially efficient quantum computer is described that is also valid and complete for the set of sentences considered here. [S1050-2947(99)02506-8]

PACS number(s): 03.67.Hk, 03.65.Bz, 02.10.By

### I. INTRODUCTION

Most of the activity in quantum computing is supported by the possibility that some problems can be solved more efficiently on quantum computers than on classical machines [1–4]. These possibilities in turn have generated much work towards possible physical realization of quantum computers using such techniques as NMR [5] and trapped ions [6]. Other work on theoretical [7] and experimental [8] error correction codes and other methods [9] to make quantum computers more robust against decoherence resulting from environmental interference [10] and other influences also is part of this activity.

The extreme sensitivity of quantum computers to environmental influences presents a barrier to the practical realization of quantum computation [11]. As a result it is not clear if quantum computers will ever become a practical reality.

The same arguments apply to quantum robots [12]. These are mobile quantum systems that include a quantum computer and other ancillary systems on board that interact with arbitrary environments of quantum systems. The types of environments and their interactions with quantum robots can be quite general. This is unlike the case for quantum computers which either seek to minimize environmental influences or consider very special types of environments such as oracles [13], data bases [2], or additional quantum registers [12].

Another reason for interest in quantum computers and quantum robots is that they represent a basis for beginning the description of quantum systems that make decisions, are aware of their environment, and have important characteristics of intelligence. The existence problem for these intelligent quantum systems is already solved as such systems include the readers (and hopefully the author) of this paper.

It should be noted that the fact that the only examples of intelligent quantum systems we know of are macroscopic ( $\sim 10^{25}$  degrees of freedom) and may be described classically, does not remove the need for a quantum-mechanical description. By study of quantum robots or quantum computers one can find out if such systems must be essentially classical and, if so, in what ways a quantum-mechanical description fails.

From the viewpoint of this paper an essential activity of intelligent systems is the construction of valid physical theories by use of mathematics and physics. The details of this validation process are not important here. What is important is that, if the theory being validated is universally applicable, then the theory is also the same theory that describes the dynamics of the systems carrying out this validation activity.

It follows that a universal theory must include a description in some form of both the mathematical and physical aspects of its own validation. This suggests the need for a coherent theory of mathematics and physics together. Such a theory will refer to its own validity and maximal completeness to the maximum extent possible. It also will be valid and maximally complete. (The importance of maximal com-

---

\*Electronic address: pbenioff@anl.gov

pleteness to these ideas was realized only when the work for this paper was done.)

If quantum mechanics is universal, then such a coherent theory of mathematics and quantum mechanics must necessarily include the description of intelligent quantum systems that can construct and validate the coherent theory. As such the coherent theory should refer to its own validity and maximal completeness to the maximum extent possible, and it should be valid and maximally complete [14].

It is to be expected that such a theory will incorporate or combine aspects of mathematical logic with quantum mechanics. This would require use in a quantum-mechanical context of mathematical logical concepts such as syntactics and semantics and their relation to one another [15–17]. Syntactics deals with expressions as strings of symbols and languages as sets of expressions. This includes the description of constants, variables, terms, formulas, axioms, theorems, and proofs. Semantics is concerned with the meaning of expressions in a language. This includes concepts such as interpretations, models, truth, validity, completeness, and consistency.

There is other work in the literature that recognizes the potential importance of trying to combine mathematical logical concepts with quantum mechanics and of describing intelligent systems in quantum mechanics. The former includes work on formulas in first-order logic [18,19], set theory and quantum mechanics [20–22], and other work [23]. The latter includes work on consciousness and quantum mechanics [29,24,25].

In this paper steps will be taken towards the use of mathematical logical concepts in quantum mechanics by considering a quantum-mechanical system (e.g., head or quantum robot) moving on a lattice of stationary quantum systems where the states of each lattice system are, in general, linear superpositions of symbol states in some basis. As the system moves and interacts with the lattice systems, the system state can be represented as a linear superposition of symbol string states. If one symbol is chosen as a blank, then the state corresponds to a linear superposition of sequences or paths of expressions as sequences of nonblank symbols separated by one or more blanks.

The main new feature added here is that some of the expressions in the superposition paths will be considered as formal sentences or words that are interpreted as having meaning to an outside observer. This is different than the usual state of affairs where the outcomes of measurements considered as numbers (symbol strings) have meaning to the observer carrying out the experiment. However, they are not usually considered to be sentences in some language that may also have meaning to an outside observer.

It is necessary to be quite clear about this point. This paper does not address the more ambitious goal of considering a quantum system that emits sentences that have meaning to the quantum system generating the sentences. Here the selection of which expressions are sentences and how they are to be interpreted is imposed externally. The quantum system knows nothing about which expressions are chosen as sentences or how they will be interpreted.

This is the main reason why the problems raised by Albert [26,27] are not relevant for this paper. This is the case even though, as will be seen, the sentences generated will be in-

terpreted as describing properties of other expressions generated by the quantum system.

Another point is that, as is well known, computers can be and are used to manipulate sentences of languages and axiom systems. An example is a computer that enumerates the theorems of an axiom system [28]. However, all these computer operations deal with the syntactic properties only of the languages. The fact that these sentences may or may not have meaning is outside the realm of what computers, as conceived so far [29], can do.

Here the emphasis is on the semantic properties of the language expressions or their meaning to an external observer. Following a very simple classical example described by Smullyan [16], the sentences will be interpreted as referring to the appearance or nonappearance of other expressions in the superposition. Based on this interpretation, definitions of truth, validity, consistency, and completeness for the set of sentences will be given and some of their properties investigated.

Since the paper is long, a summary of the sections is in order. Following the description of Smullyan's example [16] in the next section, is a description in Sec. III of a quantum-mechanical model of Smullyan's machine. The model consists of a quantum computer or quantum robot moving on a  $k$ -ary quantum register as a one-dimensional (1D) lattice of  $k$ -ary qubytes. (This term is used here instead of qubits for values of  $k > 2$ .) Discrete space and time are assumed. The single time-step generator for the dynamics of the overall quantum system is a unitary operator  $T$  acting on the Hilbert space of system states. A description of the system components is followed by a description of the properties of  $T$ . Various projection operators for expressions and combinations of expressions are also described. A representation of the overall state of the evolving quantum enumeration system is given as a Feynman [30] sum over paths or sequences of expressions.

In Sec. IV a simple subset of the set of sentences in Smullyan's example is chosen. An interpretation is considered which, for each sentence  $S$  in the subset, is based on the measurement at some time step  $n$  for the occurrence or non-occurrence of  $S$  followed by a later measurement at time step  $n + m$  for the presence or absence of the expression to which  $S$  refers.

Based on these measurements, definitions are given for the  $n, m$ -truth of the sentences. The main new feature of the interpretation used here is that the  $n, m$ -truth of  $S$  is defined only on those paths in the path sum for which  $S$  is present at time  $n$ . It is undefined on paths not containing  $S$ . Among other things, this avoids an impossible situation that arises in case  $S$  appears in no paths.

A definition of  $n, m$ -true and  $n, m$ -false is given for the domain of definition for the  $n, m$ -truth of  $S$ . Informally if  $S$  states that some expression  $X_S$  is present, then  $S$  is  $n, m$ -true [ $n, m$ -false] if all [not all] paths containing  $S$  at time  $n$  contain  $X_S$  at the later time  $n + m$ . If  $S$  states that  $X_S$  does not appear then  $S$  is  $n, m$ -true [ $n, m$ -false] if no paths [some paths] containing  $S$  at time  $n$  contain  $X_S$  at time  $n + m$ .

The dependence of these definitions on  $n$  and  $m$  is problematic because the  $n, m$ -truth of a sentence is not preserved under for changes in  $m$  or  $n$ . This problem is removed in the definition of truth for  $S$ . The definition is asymptotic in  $n$  in

that it says informally that  $S$  is true [false] if it is  $n,0$ -true [ $n,0$ -false] in the limit  $n \rightarrow \infty$ . It turns out that the limit definition is independent of  $m$  so  $m$  can be set equal to 0.

A definition of  $n,m$ -validity is given that is quite similar to that used in mathematical logic [15]. The definition connects  $T$  to the  $n,m$ -truth of  $S$  by saying  $T$  is  $n,m$ -valid for  $S$  if the  $n$ -printability of  $S$  implies that  $S$  is  $n,m$ -true. The definition is satisfying in that it is proved that if  $T$  is  $n,m$ -valid for  $S$  and its negation then it is  $n$ -consistent for  $S$  and its negation. That is, no path has both  $S$  and its negation in the region  $[0, n-2]$  of the register lattice.

The problematic dependence on  $n,m$  is removed by a limit definition of validity. That is,  $T$  is valid for  $S$  if the printability of  $S$  implies that  $S$  is true. It follows from this definition that if  $T$  is valid for  $S$  and its negation then  $T$  is consistent for  $S$  and its negation.

A problem with the definition of validity is that one way for  $T$  to be valid for all sentences is to never print any sentences. This possibility is removed by the requirement of completeness.  $T$  is complete for  $S$  if  $S$  is printable (i.e.,  $S$  appears in some path).  $T$  is complete if  $T$  is complete for all sentences; otherwise it is incomplete. Note that unlike the classical case both  $S$  and its negation can be printable; consistency demands that they not appear on the same path but says nothing about their appearance on different paths in the linear superposition.

The set of sentences considered is expanded to include the self-referential sentence that asserts its own nonprintability and its negation to show that if  $T$  is valid for these two sentences then they are not printable. In this case  $T$  is incomplete. This is the equivalent here of Gödel's incompleteness theorem [16,15]. This suggests the introduction of maximal completeness;  $T$  is maximally complete if it is complete for all sentences except those excluded by consistency requirements.

The question of the existence of  $T$  that are valid and maximally complete is discussed in Sec. V. The relations between the truth definitions and correlations between the occurrence of  $S$  and  $X_S$  are noted. An exponentially efficient quantum computer solution to the existence problem is shown for a slight generalization of the  $T$  considered here.

The relation between the existence of a valid, maximally complete  $T$  and the set of expressions taken as sentences is shown by expanding the set of sentences to include more of those in Smullyan's example. It is seen that one must be careful with closed inductive definitions which are widely used in mathematical logic. Some are harmless; others, such as that used to define sentences in Smullyan's example, have nontrivial consequences for quantum-mechanical systems.

A final discussion section includes the point that there are many other interpretations possible for the sentences. An example of another very simple interpretation is briefly discussed. It is noted that, due to the freedom in the choice of a basis for representing the symbols, many more interpretations are possible in quantum mechanics than in classical mechanics.

Another point is that the discussion of the printability both of the sentences and the expressions to which they refer may be of more general applicability than would appear for this special example. This is based on the fact that any quan-

tum system telling us something must do so by means of emitting or printing sentences with meaning.

## II. SMULLYAN'S ENUMERATION MACHINE

Smullyan's example [16] consists of a (classical) machine or computer that prints or enumerates expressions consisting of finite nonempty strings of the five symbols  $\sim PN()$ . If an expression is printable by the machine it will eventually be printed. The norm of any expression  $X$  is defined as the expression  $X(X)$ .

The sentences are defined to be any of the four types of expressions  $P(X), \sim P(X), PN(X), \sim PN(X)$  where  $X$  is any expression. The sentences are interpreted to apply to the enumeration generated by the machine in the sense that  $P(X)$  means  $X$  is printable,  $\sim P(X)$  means  $X$  is not printable,  $PN(X)$  means the norm of  $X$  is printable, and  $\sim PN(X)$  means the norm of  $X$  is not printable. Thus  $P(X)$  is true if and only if  $X$  is printable,  $\sim P(X)$  is true if and only if  $X$  is not printable,  $PN(X)$  is true if and only if the norm of  $X$  is printable, and  $\sim PN(X)$  is true if and only if the norm of  $X$  is not printable. Here and in the following  $X$  denotes either an expression variable or a name for a specific expression. It should be clear from context which is meant.

Under this interpretation the sentences refer to dynamic properties of the machine that generates them in that they describe what the machine does or does not do. More precisely, the interpretation is assumed to be *valid* for the machine in that any sentence that is printed is true or, equivalently, no false sentence is printed. Thus if  $P(X)$  is printed, then  $X$  has been or will be printed eventually, and if  $\sim PN(X)$  is printed then  $X(X)$  will not ever be printed. Similar statements hold for the other two types of sentences.

The implications, printable implies truth (or falseness implies not printable), which hold if the interpretation is valid, are one sided as the converse implications are false. To see this consider the sentence  $\sim PN(\sim PN)$  [16]. This sentence is self-referential in that it refers to its own nonprintability. Thus this sentence is true if and only if it itself is not printable. Since the interpretation is supposed to be valid for the machine, this is a sentence that is true that the machine cannot print. Also the sentence  $PN(\sim PN)$  is not printable as it is false.

The nonprintability of a true sentence, assuming validity, shows that for this machine printability is not equivalent to truth of the sentences. This is similar to Tarski's theorem [17,16] which says that in any formal axiom system the set of true formulas is not definable in the system. Thus the truth or falseness of the sentences is a property not expressible by the machine for the assumed interpretation. In a similar way the system is incomplete in that neither the sentence  $PN(\sim PN)$  nor its negation are printable. This is an example of Gödel's incompleteness theorem [17,15] if printability is interpreted as provability [16].

For use in the following two aspects are worthy of note. One can show that the sentences  $\sim PN(\sim PN)$  and  $PN(PN)$  are the only two self-referential sentences. To see this write  $\sim PN(X) = X(X)$  and require that the number of symbols in the expressions on both sides of the equal sign be the same. This shows that  $X$  must have three symbols. For  $PN(X)$  one shows that  $X$  has two symbols.

The other aspect is that, as will be seen later, the definitions of sentences and their meaning is quite complicated and not really necessary for the purposes of this paper. For this reason all sentences of the form  $PN(X), \sim PN(X)$  will be excluded as will sentences of the form  $P(X), \sim P(X)$  where  $X$  is a sentence. Here sentences will be limited to be of the form  $P(X), \sim P(X)$  where  $X$  is an expression that is not a sentence.

### III. A QUANTUM-MECHANICAL MODEL OF AN ENUMERATION MACHINE

#### A. Component description

A quantum-mechanical model of a symbol enumeration machine as described above consists of a multistate head or quantum robot moving on a lattice or quantum register of 5-ary qubytes. The interaction between the quantum robot and the lattice qubytes is local and includes changing the states of the neighborhood qubytes. For the purposes of this paper it is immaterial whether the whole system is regarded as a multiregister quantum computer or as a quantum robot or as a multistate head moving on a quantum register [12].

The set of five symbols represented by the states of each qubyte are  $\sim, P, (, )$ , and  $0$ . The  $0$  denotes the blank symbol and will be interpreted as a spacer to separate a string of five symbols into a string of expressions separated by spacers. A convenient set of basis states for the quantum register is the set of states  $|s\rangle = \otimes_{j=-\infty}^{\infty} |s_j\rangle$  where  $|s_j\rangle$  denotes site  $j$  qubyte in a state corresponding to any one of the five symbols. The state  $|s\rangle$  describes an infinite symbol string state for which at most a finite number of symbols are nonblank. This limitation, referred to as the  $0$  state tail condition, is used to keep the Hilbert space of the overall system, including the register, separable.

The states of the head or quantum robot can be represented in the form  $|l, j\rangle$  where  $|l\rangle$  denotes any of the  $L$  states of the internal degrees of freedom of the head and  $|j\rangle$  is the lattice position state of the head. For example, if the internal degrees of freedom of the head or quantum robot consist of another  $m$  state head moving on a lattice of  $n$  qubits  $L = mn(2^n)$ . Based on the above a general normalized state of the overall system has the form  $\Psi = \sum_{l,j,s} c_{l,j,s} |l, j, s\rangle$  where the  $c_{l,j,s}$  are arbitrary complex coefficients whose absolute squares sum to unity. The  $s$  sum is over all lattice system basis states that satisfy the  $\bar{0}$  state tail condition.

#### B. System dynamics

The dynamics of the overall system is given by a unitary step operator  $T$  that represents the changes occurring in a

single time step. If  $\Psi(0)$  is the overall system state at time  $0$  then  $\Psi(n) = T^n \Psi(0)$  is the state at time  $n$ .

In order that  $T$  describe enumerations of symbols on the qubyte lattice it is necessary to require that the states of qubytes in finite lattice regions become asymptotically (as  $n \rightarrow \infty$ ) stationary. Any dynamics in which the states of finite regions of the quantum register are always changing does not represent an enumeration. Mathematically this condition can be expressed by the requirement that the expectation value  $\langle \Psi(n) | P_{\underline{s}R} | \Psi(n) \rangle$  has a limit as  $n \rightarrow \infty$ . Here  $P_{\underline{s}R}$  is the projection operator for the symbol string state  $|\underline{s}\rangle$  in a finite region  $R$  of the lattice.

To keep things simple this requirement will be satisfied here by requiring  $T$  to describe motion of the quantum robot in one direction only on the 1D lattice of qubytes. Each iteration of  $T$  will move the quantum robot or head one site to the right. During this motion the internal state of the head and the states of the qubytes at the original and final locations of the head can be changed.

To this end let  $T$  be given by

$$T = U \otimes u. \quad (1)$$

Here  $u$  is the unitary shift operator for moving the quantum robot one lattice site to the right and  $U$  is an arbitrary  $25L$ -dimensional unitary operator on the  $L$  internal head states and states of the two lattice qubytes, one at and one just to the right of the head location.

The action of  $T$  on each overall system state  $|l, j, s\rangle$  is given by

$$\begin{aligned} T|l, j, s\rangle &= |j+1, s_{\neq [j, j+1]}\rangle \\ &\times \sum_{l', s'_j, s'_{j+1}} |l', s'_j, s'_{j+1}\rangle \langle l', s'_j, s'_{j+1} | U |l, s_j, s_{j+1}\rangle, \end{aligned} \quad (2)$$

where  $|s_{\neq [j, j+1]}\rangle$  is the state of lattice qubytes outside of sites  $j, j+1$  and  $u|j\rangle = |j+1\rangle$  has been used.

To be consistent with the  $0$  state tail condition and the choice of  $0$  as the symbol blank or vacuum, the initial state of interest here is  $|0, 0, 0\rangle$ . This state has the head in internal state  $0$  at lattice site  $0$  and a completely blank quantum register. Inclusion of initial wave packet states of different head positions and internal states is not necessary here.

An expansion of  $T^m$  acting on the initial state  $|0, 0, 0\rangle$ , in terms of intermediate states gives

$$\begin{aligned} T^m |000\rangle &= |m, \underline{0}_{\neq [0, m]}\rangle \sum_{s_0, \dots, s_m} \sum_{s'_1, \dots, s'_m} \sum_{l_1, \dots, l_m} |l_m, s_{[0, m-1]}, s'_m\rangle \langle l_m, s_{[0, m-1]}, s'_m | U |l_{m-1}, s_{[0, m-2]}, s'_{m-1}, 0_m\rangle \\ &\times \langle l_{m-1}, s_{[0, m-2]}, s'_{m-1}, 0_m | U |l_{m-2}, s_{[0, m-3]}, s'_{m-2}, 0_{[m-1, m]}\rangle \cdots \langle l_1, s_0, s'_1, 0_{[2, m]} | U |0, 0, 0\rangle. \end{aligned} \quad (3)$$

Here  $s_{[0,b]}$  denotes a string of symbols  $s_0, \dots, s_b$  in the region  $0, 1, \dots, b$  of lattice sites.

This equation shows that the quantum system with dynamics given by  $T$  is a satisfactory enumeration system in that once the head passes a lattice site the states of all qubits in passed regions, denoted by the states  $|s_{[0,b]}\rangle$  are not changed by more iterations of  $T$ . The growth in the passed lattice region by one site per  $T$  iteration is shown by the increase of  $b$  to  $b+1$  in the states  $|s_{[0,b]}\rangle$  appearing in each matrix element. The sum over the unprimed  $s$  represents the completed effect of the passage of the head in generating the states  $|s_{[0,b]}\rangle$ ; the sum over primed  $s$  gives intermediate changes in qubyte states at the location of the head.

### C. Some expression projectors

Expressions  $X$  are defined as consecutive finite strings of any of the four symbols  $P, (, \sim$  with  $0$  excluded within  $X$ . To separate  $X$  from other expressions in a symbol string, it is required that the terminal symbol of  $X$  is followed by at least one  $0$ . Similarly the initial symbol of  $X$  is preceded by at least one  $0$ . Let  $a, b$  be the initial and terminal symbol site location of  $X$  where  $l(X)$  is the length of  $X$ . One has  $a = b - l(X) + 1$ .

Define  $Q_{X,b}$  to be the projection operator for finding  $X$  at lattice sites  $a, a+1, \dots, b$ ,  $0s$  at sites  $a-1, b+1$ , and any symbol, including the blank, at other lattice sites. The quantum robot can be anywhere and in any internal state. This operator is basic to all that follows.

Let  $[m, n]$  with  $n > m$  be a lattice region of sites  $m, m+1, \dots, n$ . Define  $Q_{X,[m,n]}$  to be the projection operator for finding  $X$  somewhere in the region  $[m, n]$ .  $Q_{X,[m,n]}$  is defined by

$$Q_{X,[m,n]} = \sup_{k=m+l(X)-1}^n Q_{X,k}. \quad (4)$$

The least upper bound is used because  $Q_{X,k}$  and  $Q_{X,k'}$  with  $k \neq k'$  are not necessarily orthogonal.  $Q_{X,[m,n]} = 0$  if the region is too short to accommodate  $X$ .

Let  $Q_{\neg X,[m,n]}$  be the projection operator for not finding  $X$  anywhere in  $[m, n]$ . Clearly

$$Q_{X,[m,n]} + Q_{\neg X,[m,n]} = 1. \quad (5)$$

Additional useful properties of these projectors are

$$\begin{aligned} Q_{X,[m,n]} &< Q_{X,[m,n+1]}, \\ Q_{\neg X,[m,n]} &> Q_{\neg X,[m,n+1]}, \end{aligned} \quad (6)$$

and

$$Q_{X,[m,n]} = \sum_{j=l(X)}^n Q_{X,[m,m+j]}^{\text{1st}}. \quad (7)$$

Here  $Q_{X,[m,m+j]}^{\text{1st}}$  is the projection operator for  $X$  occurring in the region  $[m, m+j]$  with the terminal symbol of  $X$  at site  $m+j$ ,  $0s$  at sites  $m+j+1, m+j-l(X)$ , and no  $X$  anywhere else in  $[m, m+j]$ .  $Q_{X,[m,m+j]}^{\text{1st}}$  is the identity on all other lattice sites. The superscript 1st denotes the fact that  $X$  does

not occur in intervals  $[m, m+k]$  with  $k < j$ , i.e., the first occurrence of  $X$ . Note that the projection operators in the  $j$  sum are pairwise orthogonal.

In what follows projection operators are needed for expressions, that, once generated by iterations of  $T$ , are not changed by further iterations. One way to achieve this is to include projection operators for head positions to the right of expressions of interest. To this end let  $Q_{X,[m,n]}^h$  be the projection operator for  $X$  anywhere in the region  $[m, n]$  where  $X$  is separated from other expressions in the region by one or more  $0s$ . If  $X$  ends at site  $n$  there is a  $0$  at site  $n+1$ ; if  $X$  begins at site  $m$ , there is a  $0$  at site  $m-1$ . The head is at site  $n+2$ . That is

$$Q_{X,[m,n]}^h = Q_{X,[m,n]} Q_{n+2}^h, \quad (8)$$

where  $Q_{n+2}^h$  is the projection operator for the head at position  $n+2$  and in any of the  $L$  internal states. A useful generalization is the projection operator  $Q_{X,[m,n],k}^h = Q_{X,[m,n]} Q_{n+k+2}^h$  for  $X$  anywhere in the interval  $[m, n]$  and the head  $k+2$  sites beyond  $n$ . For non-negative  $k$  this projector has the following commutation relation with  $T$  as defined by Eq. (1):

$$T Q_{X,[m,n],k}^h = Q_{X,[m,n],k+1}^h T. \quad (9)$$

The limit projector  $Q_X^h$  defined by

$$Q_X^h = \sum_{n=0}^{\infty} Q_{X,[0,n]}^h \quad (10)$$

corresponds to  $X$  located anywhere to the right of the origin and the head at least two sites beyond the terminal symbol of  $X$ . The projectors in the sum are pairwise orthogonal because of the orthogonality of the head location projectors. Whether or not this limit exists is not important here because the limit operator can always be replaced by  $Q_{X,[0,n]}^h$  for some  $n$  in any of the matrix elements that occur in this work.

Let  $X$  and  $Y$  be two expressions. Then

$$Q_{X \wedge Y,[m,n]}^h = Q_{X,[m,n]}^h Q_{Y,[m,n]}^h \quad (11)$$

is the projection operator for finding  $X$  and  $Y$  in the region  $[m, n]$  and the head at site  $n+2$ . This projection operator is zero if the region is too small to contain  $X$  and  $Y$  without overlap. This is the case even if  $Y$  is a subexpression of  $X$  because the projectors include blank symbols preceding and following each expression.

More generally let  $Q_{\bigwedge_{k=1}^m Y_k}^h$  be defined by

$$Q_{\bigwedge_{k=1}^m Y_k}^h = \prod_{k=1}^m Q_{Y_k}^h = \sum_{j=0}^{\infty} Q_{\bigwedge_{k=1}^m Y_k, [0,j-2]} Q_j^h. \quad (12)$$

This is the projection operator for finding expressions  $Y_1, \dots, Y_k$  anywhere to the right of the lattice origin and to the left of the head (and a  $0$  just left of the head).

It is important to note that the symbols  $\neg, \wedge$  appearing in the subscripts are in the metalanguage used to describe the properties of the system. They do not appear in the expressions or sentences described here. The operator  $Q_{\neg X}^h$  projects

out all expression path states that do not contain  $X$  anywhere. This is the case whether  $X$  is or is not a sentence.

#### D. Sums over paths of expressions

At this point it is worthwhile to look more closely at iterations of  $T$  and the generation of an exponentially growing tree of paths or sequences of expressions separated by strings of  $0s$ . To this end define the projection operators

$$Q_0 = \sum_{j=-\infty}^{\infty} Q_{0,j-1} Q_j^h,$$

$$Q_{\neq 0} = \sum_{j=-\infty}^{\infty} Q_{\neq 0,j-1} Q_j^h. \quad (13)$$

Here  $Q_j^h$  is the projector for the head at site  $j$ ,  $Q_{0,j-1}$  is the projector for a 0 at site  $j-1$ , and  $Q_{\neq 0,j-1}$  is the projector for any one of the four symbols ( $\cdot$ ),  $P$ ,  $\sim$  at site  $j-1$ . It is clear that  $Q_0 + Q_{\neq 0} = 1$ .

These operators can be used to separate  $T$  defined by Eq. (1) into the sum of two operators  $T_0, T_{\neq 0}$ :

$$T = (Q_0 + Q_{\neq 0})T = T_0 + T_{\neq 0}. \quad (14)$$

The projectors are chosen so that for any overall system state  $\Psi$ ,  $T_0\Psi$ , and  $T_{\neq 0}\Psi$  show, respectively, a 0 or an expression symbol state for the qubyte at the site last visited by the head. These states have the property that iteration of  $T$  on these states does not change these qubyte states as changes are limited to those qubytes at and immediately to the right of the head location. Iteration of  $T_0$  generates a spacer state

as a finite string of  $0s$ , and iteration of  $T_{\neq 0}$  generates a linear superposition of expression states. Note that  $T_0$  and  $T_{\neq 0}$  do not commute.

Using the fact that  $T^n = (T_0 + T_{\neq 0})^n$  and collecting together powers of  $T_0$  and  $T_{\neq 0}$  gives

$$T^n = \sum_{t=1}^n \sum_{h_1, \dots, h_t}^{\delta(\Sigma, n)} [(T_{\neq 0}^{h_t} T_0^{h_{t-1}} T_{\neq 0}^{h_{t-2}} \dots T_0^{h_2} T_{\neq 0}^{h_1} + T_0^{h_t} T_{\neq 0}^{h_{t-1}} \dots T_{\neq 0}^{h_2} T_0^{h_1}) \delta_{t, \text{odd}} + (T_0^{h_t} T_{\neq 0}^{h_{t-1}} \dots T_{\neq 0}^{h_2} T_0^{h_1} + T_{\neq 0}^{h_t} T_0^{h_{t-1}} \dots T_0^{h_1}) \delta_{t, \text{even}}]. \quad (15)$$

The  $t$  sum is over the number of expressions and intervening spacers, and the  $h$  sums are over the length of the expressions and spacers in an alternation with  $t$  spacers and expressions. The upper limit of the  $h$  sums denotes the requirement that  $h_1 + h_2 + \dots + h_t = n$ . The above shows that four types of alternations are possible; they may begin or end with either  $T_0$  or  $T_{\neq 0}$ . How they end depends on how they begin and whether  $t$  is even or odd. The two types for  $t$  odd are shown in the first line above multiplied by a  $\delta$  function for  $t$  odd. The second line gives the two types for  $t$  even.

The above can be used to expand  $T^n|0,0,0\rangle$  into a Feynman sum [30] over sequences or paths of expression states separated by  $0s$  similar to the sum over phase paths used elsewhere [12]. In order to keep things simple the expansion will be given for the first alternation type only with  $t$  odd and full account will be taken of the fact that once the head passes a lattice region no further interactions occur with the qubytes in the region. One has

$$T^n|0,0,0\rangle = \sum_{t=1}^n \sum_{h_1, \dots, h_t}^{\delta(\Sigma, n)} \sum_{l_1, \dots, l_t} \sum_{s'_1, \dots, s'_t} \sum_{X_1, X_2, \dots, X_{[(t+1)/2]}} |l_t, n, \underline{0} * X_1 * \underline{0} * X_2 * \underline{0} * \dots * \underline{0} * X_{[(t+1)/2]} * s'_t * \underline{0}\rangle$$

$$\times \langle l_t, X_{[(t+1)/2]} * s'_t | T_{\neq 0}^{h_t} | l_{t-1}, s'_{t-1} * \underline{0} \rangle \langle l_{t-1}, \underline{0} * s'_{t-1} | T_0^{h_{t-1}} | l_{t-2}, s'_{t-2} * \underline{0} \rangle$$

$$\times \dots \langle l_2, \underline{0} * s'_2 | T_0^{h_2} | l_1, s'_1 * \underline{0} \rangle \langle l_1, X_1 * s'_1 | T_{\neq 0}^{h_1} | 0,0,0 \rangle. \quad (16)$$

The  $l$  sums are over head or robot internal states, and the  $s'$  sums are over intermediate states (including 0) of qubytes at the head location. As shown earlier they may be changed by the next iteration of  $T$ . The  $X$  sums are over all possible completed expressions of length specified by the  $h$  sum terms. The head position state has been suppressed in the matrix elements. The asterisk denotes concatenation of symbols and expressions.

Each matrix element shows the state changes resulting from one alternation. For example,  $T_0^{h_j}$  is active on the lattice region extending from  $\sum_{k=1}^{j-1} h_k$  (the initial head position) to  $\sum_{k=1}^j h_k$  (the final head position). It converts the state of qubytes in the lattice region from  $|s'_{j-1} * \underline{0}\rangle$  to  $|\underline{0} * s'_j\rangle$  where  $\underline{0}$  denotes a string of  $h_j$   $0s$  and a sum over  $s'_j$  is implied. (Here the subscript  $j$  is an alternation index, not a lattice site.)

The action of  $T_{\neq 0}^{h_j}$  differs only in that the final state is  $|X_j * s'_j\rangle$  where  $X_j$  is an expression of length  $h_j$ . Sums over  $s'_j$  and  $X_j$  are implied. Note that this represents the generation of a linear superposition of length  $h_j$  expression states in the specified lattice region. Except for future entanglements with states of qubytes not yet reached by the head, the expression states are not changed by more iterations of  $T$ . This is shown by the fact that each  $X_j$  also appears in the final output state  $|l_t, n, \underline{0} * X_1 * \underline{0} * X_2 * \underline{0} * \dots * \underline{0} * X_{[(t+1)/2]} * s'_t * \underline{0}\rangle$  which shows the head in the internal state  $l_t$  at position  $n$ . The state shows the sequence of expressions  $X_1, X_2, \dots, X_{[(t+1)/2]}$  separated by finite strings of  $0s$ . The terminal expression  $X_{[(t+1)/2]}$  is complete for those components in which one more iteration of  $T$  converts  $s'_t$  to 0. It is incomplete in the

other components of the sums of Eq. (16). The qubits at lattice sites not in the region  $[0, n]$  are in state  $|0\rangle$ .

Equation (16) shows a sum over all expression sequence states consistent with the requirement that the lengths of all expressions and spacers sum to  $n$ . It includes components for one expression of length  $n-2$  up to components for  $[n/2]$  expressions each containing one symbol. Other components with expressions of varying length are also included.

The amplitude associated with each expression sequence state  $|l_t, n, 0 * X_1 * 0 * X_2 * 0 *, \dots, * 0 * X_{[(t+1)/2]} * s'_t * 0\rangle$  is given by a partial sum over products of matrix elements shown in Eq. (16). It depends sensitively on the properties of the  $25L$  dimensional unitary operator  $U$ , Eq. (1), which shows the changes in the lattice system states as the quantum robot moves along the lattice.

It is clear from the above that each expression path may contain many different expressions  $X, Y, Z, \dots$  as well as repetitions of expressions. Let  $Q_{Y, [0, n-2]}^h$  be the projection operator for the expression  $Y$  appearing somewhere in the region  $[0, n-2]$  and the head at site  $n$  in any internal state, Eq. (8). Then the state  $Q_Y^h T^n |0, 0, 0\rangle = Q_{Y, [0, n-2]}^h T^n |0, 0, 0\rangle$ , expanded as a sum over expression paths [Eq. (16)], contains all expression path states with  $Y$  appearing somewhere in the region  $[0, n-2]$ . For later times  $m+n$  the state  $T^m Q_{Y, [0, n-2]}^h T^n |0, 0, 0\rangle$ , expanded as a path sum, shows a sum over expression path states with  $n+m$  symbols, including blanks, that have in common the appearance of  $Y$  somewhere in the region  $[0, n-2]$ .

#### IV. TRUTH, VALIDITY, CONSISTENCY, AND COMPLETENESS

##### A. Sentences and their interpretation; truth

So far expressions have been discussed without any mention of which ones are sentences and how the sentences should be interpreted. Here sentences are those expressions with the form  $P(X), \sim P(X)$  where  $X$  is any expression that is not a sentence. In the following  $S$  will often be used to denote a sentence and  $X_S$  the expression to which the sentence refers.

A sentence  $S$  is defined to be  $n$ -printable if  $\langle 0, 0, 0 | (T^\dagger)^n Q_S^h T^n | 0, 0, 0 \rangle > 0$ . It is not  $n$ -printable if the matrix element equals 0.  $S$  is printable if  $\lim_{n \rightarrow \infty} \langle 0, 0, 0 | (T^\dagger)^n Q_S^h T^n | 0, 0, 0 \rangle > 0$ . Otherwise it is not printable.  $Q_S^h$  is given by Eq. (10) with  $X=S$ . The limit clearly exists because the matrix element is nondecreasing as  $n$  increases and is real, positive and bounded above by 1.

The definition of printability means that if  $S$  is printable it will appear somewhere in the sum of Eq. (16). The limit  $n \rightarrow \infty$  is needed because, in general, there is no upper bound on when a sentence must first appear in the different paths. Conversely a sentence is not printable if it never appears in any path.

There are many possible ways to interpret the sentences  $S$ . Here the interpretation of  $P(X)$ , stated informally, is that all paths in the path sum of Eq. (16) that contain  $P(X)$  also contain  $X$ . The informal interpretation of  $\sim P(X)$  is that no path containing  $\sim P(X)$  also contains  $X$ .  $P(X)$  or  $\sim P(X)$  are true if their informal meaning statements are true. The goal of the following is to make precise these informal expres-

sions of the meaning and truth of the sentences.

A feature introduced by these interpretations is that the meaning of a sentence  $S$  is limited to those paths containing  $S$ .  $S$  has no meaning in paths not containing  $S$ . It says nothing about the presence or absence of  $X_S$  in these paths. This is a consequence of the presence of many paths in the path sum of Eq. (16). Classically, where there is just one path, this partial definability is reflected in the fact that, for the interpretation used here, any sentence not appearing in the enumeration path also has no meaning for the expressions in the path.

The meaning of  $S$  is closely connected to the carrying out of measurements on the quantum enumeration system. Suppose one carries out a measurement of the observable  $S Q_S^h + \neg S Q_{\neg S}^h$  at time  $n$  on the state  $\Psi(n) = T^n |0, 0, 0\rangle$  and  $m$  time steps later carries out a measurement of the observable  $X_S Q_{X_S}^h + \neg X_S Q_{\neg X_S}^h$  on the component state  $Q_S^h |\Psi(n)\rangle$ . This can be described by adding two ternary qubits to the quantum enumeration systems to record the outcomes of these operations. The qubyte states  $|i\rangle_S, |i\rangle_{X_S}$  denote the initial or no measurement state and the other states denote the possible two outcomes of each measurement.

The overall process is described by unitary operators  $U^S$  and  $U^{X_S}$  that establish correlations between the states of the enumeration system and the qubits. The result of carrying out the two measurements is given by

$$\begin{aligned} U^{X_S} T^m U^S \Psi(n) |i\rangle_S |i\rangle_{X_S} &= Q_{X_S}^h T^m Q_S^h \Psi(n) |1\rangle_S |1\rangle_{X_S} \\ &+ Q_{\neg X_S}^h T^m Q_S^h \Psi(n) |1\rangle_S |0\rangle_{X_S} \\ &+ T^m Q_{\neg S}^h \Psi(n) |0\rangle_S |i\rangle_{X_S}. \end{aligned} \quad (17)$$

In the above the amplification or decoherence by interaction with the environment [10] necessary to complete the measurement process is ignored as it is not needed here. In Peres' language [31] the above is a premeasurement.

The limitation of the  $X_S$  measurement to component states containing  $S$  is made because extension of the measurement of  $X_S$  to component states not containing  $S$  is not needed here. However, if desired, the extension can be included. The only effect is to add an additional term to Eq. (17). No conclusions are affected.

In order to keep things simple it has been assumed in the above that the projection operators  $Q_X^h$  and  $Q_{\neg X}^h$  where  $X$  is an expression or sentence can be measured in one time step. This is unrealistic as the measurement involves searching for  $X$  in an arbitrarily large region of the lattice. A more realistic approach would be to replace this one step operation with a measurement described as a multistep search task carried out by a quantum robot moving along the lattice of qubits [12]. For each time  $n$  the search would terminate as a finite lattice region  $[0, n-2]$  only needs to be searched. However the number of steps in the search is polynomial in  $n$ .

Eq. (17) clearly supports the idea that for a measurement at time  $n$  the meaning or interpretation of  $S$  is limited to those component path states in which  $S$  occurs somewhere in the region  $[0, n-2]$ .  $S$  has no meaning for those component states in which  $S$  occurs nowhere in  $[0, n-2]$ . This would remain the case even if the measurement of  $X_S$  were ex-

tended to paths not containing  $S$ . The equation also shows that for the measurement at time  $n$  the domain of states for which  $S$  has meaning is the Hilbert space spanned by the set of normalized states  $T^m Q_S^h \Psi(n) / \|Q_S^h \Psi(n)\|$  for  $m = 0, 1, \dots$ . This Hilbert space is empty if  $S$  has more than  $n - 2$  symbols.

This Hilbert space is also the state space or domain of truth definition for the sentence  $S$  at time  $n$ . It corresponds to the states shown in the first two right-hand terms of Eq. (17).

In other words, the truth at time  $n$  for a sentence  $S$  is definable on the (unnormalized) states  $T^m Q_S^h \Psi(n)$  for  $m = 0, 1, \dots$ . It is not definable on the states  $T^m Q_{\sim S}^h \Psi(n)$ .

Many definitions of  $n, m$ -truth for the sentences are possible. Here  $n, m$ -truth will be defined as follows:  $S = P(X)$  is  $n, m$ -true [ $n, m$ -false] if the amplitude of the second right-hand term of Eq. (17)  $= 0$  [ $> 0$ ].  $S = \sim P(X)$  is  $n, m$ -true [ $n, m$ -false] if the amplitude of the first right-hand term  $= 0$  [ $> 0$ ]. That is

$$P(X) \begin{cases} \text{is } n, m\text{-true} \\ \text{is } n, m\text{-false} \end{cases} \text{ if } \langle \Psi(n) | Q_{P(X)}^h (T^\dagger)^m Q_{\sim X}^h T^m Q_{P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ > 0, \end{matrix} \quad (18)$$

$$\sim P(X) \begin{cases} \text{is } n, m\text{-true} \\ \text{is } n, m\text{-false} \end{cases} \text{ if } \langle \Psi(n) | Q_{\sim P(X)}^h (T^\dagger)^m Q_X^h T^m Q_{\sim P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ > 0. \end{matrix} \quad (19)$$

These definitions follow the practice in mathematical logic [15] of defining the truth of formal sentences relative to the informal truth of the statements that are the interpretation of the sentences. Here the interpretation of  $P(X)$  is the mathematical statement  $\langle \Psi(n) | Q_{P(X)}^h (T^\dagger)^m Q_{\sim X}^h T^m Q_{P(X)}^h | \Psi(n) \rangle = 0$  that describes a property of the quantum system shown by the measurement of Eq. (17).  $P(X)$  is  $n, m$ -true [ $n, m$ -false] if the mathematical statement is true [false] for the system. Similarly  $\sim P(X)$  is  $n, m$ -true [ $n, m$ -false] if the statement  $\langle \Psi(n) | Q_{\sim P(X)}^h (T^\dagger)^m Q_X^h T^m Q_{\sim P(X)}^h | \Psi(n) \rangle = 0$  is true [false] for the system.

The above definitions also have different properties from the usual definitions. Besides the need for a state domain of truth definition, it is *not* the case that  $P(X)$  is  $n$ -true if  $\sim P(X)$  is false and conversely. This is a consequence of the quantum-mechanical nature of the enumeration system with the presence of multiple paths in the path sum. In fact there is no reason why these sentences cannot both be  $n, m$ -true on their respective domains of definition.

An equivalent definition of  $n, m$ -truth follows from the fact that  $Q_X^h$  and  $Q_{\sim X}^h$  are orthogonal and sum to unity:

$$P(X) \begin{cases} \text{is } n, m\text{-true} \\ \text{is } n, m\text{-false} \end{cases} \text{ if } \langle \Psi(n) | (T^\dagger)^m Q_X^h T^m Q_{P(X)}^h | \Psi(n) \rangle - \langle \Psi(n) | Q_{P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ < 0, \end{matrix} \quad (20)$$

$$\sim P(X) \begin{cases} \text{is } n, m\text{-true} \\ \text{is } n, m\text{-false} \end{cases} \text{ if } \langle \Psi(n) | (T^\dagger)^m Q_{\sim X}^h T^m Q_{\sim P(X)}^h | \Psi(n) \rangle - \langle \Psi(n) | Q_{\sim P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ < 0. \end{matrix} \quad (21)$$

The fact that the projectors  $Q_S^h$  and  $(T^\dagger)^m Q_{X_S}^h T^m$  commute has been used to obtain these expressions.

The connection of these definitions of  $m, n$ -truth to the measurement is shown by the fact that the definition correlates the  $n, m$ -truth of  $S$  to the amplitudes of states appearing in the first two terms of Eq. (17). Also the domains of definability and undefinability of  $n, m$ -truth are shown in the equation.

These definitions also illustrate another reason for use of a domain of definability for the truth of a sentence at time  $n$ . For the special case where the probability of finding  $S$  at time  $n$  is 0 then the normalized measurement state corresponding to finding  $S$  at  $n$ ,  $Q_S^h \Psi(n) / \|Q_S^h \Psi(n)\|$  is undefined. In addition, if the definitions of  $n, m$ -truth also applied for this case, it would be impossible for  $S$  to be  $n, m$ -false. This untenable situation is avoided by the use of domains of definability.

A problem with these definitions is their dependence on  $m$  and  $n$ . To see this let  $T$  be such that there is no upper bound on the distance of first appearance of  $X_S$  from the origin. That is, for each  $m$  there exist paths in Eq. (16) such that for some  $m' > m$   $X_S$  is not in  $[0, n + m - 2]$  but is in  $[0, n + m' - 2]$ . This means that  $P(X)$  can be  $n, m$ -false and  $n, m'$ -true. Also  $\sim P(X)$  can be  $n, m$ -true and  $n, m'$ -false. This follows from Eq. (6).

It is also the case that  $n, m$ -truth depends on  $n$ . For example,  $T$  may be such that  $S$  is  $n, m$ -true but is  $n', m$ -false for some  $n' > n$ . This is possible because as  $n$  increases, extensions and branching of paths occurs. Since  $S$  may first occur in some of these extensions or branches, it is possible for  $S$  to be  $n', m$ -false on some of these extensions. Additional dependence on  $n$  arises from the above argument for  $m$  dependence because the size of the region for the  $X_S$  measurement, Eq. (17), depends on both  $n$  and  $m$ .



The definition of truth should be such as to exclude the dependence on  $n$  and  $m$ . It should also maximize the domain of definability of truth for a sentence. These goals are attained by taking the limit  $n \rightarrow \infty$  in Eq. (17) and in the definitions of  $n, m$ -truth. The desired maximization is achieved because  $\langle \Psi(n) | Q_S^h | \Psi(n) \rangle$  is a nondecreasing function of  $n$  bounded above by 1.

In addition, for the limit definition, the value of  $m$  is arbitrary as the limit values of the matrix elements are independent of  $m$ . This is shown in the Appendix. Because of this the value of  $m$  will be chosen to be 0. Based on these aspects true and false are defined as follows for the two types of sentences:

$$P(X) \begin{cases} \text{is true} \\ \text{is false} \end{cases} \text{ if } \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{-X}^h Q_{P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ > 0, \end{matrix} \quad (22)$$

$$\sim P(X) \begin{cases} \text{is true} \\ \text{is false} \end{cases} \text{ if } \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h Q_{\sim P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ > 0. \end{matrix} \quad (23)$$

Equivalent definitions corresponding to those of Eqs. (20) and (21) are

$$P(X) \begin{cases} \text{is true} \\ \text{is false} \end{cases} \text{ if } \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h Q_{P(X)}^h | \Psi(n) \rangle - \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ < 0, \end{matrix} \quad (24)$$

$$\sim P(X) \begin{cases} \text{is true} \\ \text{is false} \end{cases} \text{ if } \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{-X}^h Q_{\sim P(X)}^h | \Psi(n) \rangle - \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim P(X)}^h | \Psi(n) \rangle \begin{matrix} = 0 \\ < 0. \end{matrix} \quad (25)$$

These definitions remove the  $n$  dependence in that if a sentence  $S$  is  $n, 0$ -true for each  $n$  then it is true. The converse, namely that if  $S$  is true, it is  $n, 0$ -true for each  $n$ , is true for  $S = \sim P(X)$ . This follows from Eq. (23) as the matrix element is nondecreasing as  $n$  increases. The converse is not necessarily true for  $S = P(X)$ . However, one can show from Eq. (7) and the material in the Appendix that if  $P(X)$  is true then it is  $n, \infty$ -true for each  $n$ .

### B. Validity and consistency

The definitions of  $n, m$ -truth and truth for the sentences given above clearly depend on the system dynamics  $T$ . However they place no restrictions on  $T$ . Each sentence with a nonempty domain of  $n, m$ -truth definition (i.e.,  $n$ -printable) can be  $n, m$ -true or  $n, m$ -false, (or true or false) on the domain. The main idea here is to use validity to connect  $T$  to the truth of the sentences.

Informally the idea is that  $T$  is  $n, m$ -valid for a sentence  $S$  if  $S$  is in fact  $n, m$ -true on its domain of definition. That is, all paths containing  $P(X) [\sim P(X)]$  in the region  $[0, n-2]$  also contain [do not contain]  $X$ . This is captured in the following definition:

*Definition 1:*  $T$  is  $n, m$ -valid for a sentence  $S$  if  $S$  is not  $n$ -printable or  $S$  is  $n$ -printable and  $n, m$ -true on its domain of definition.

An equivalent statement of the definition is that  $T$  is  $n, m$ -valid for  $S$  if  $n$ -printability of  $S$  implies that  $S$  is  $n, m$ -true on its domain of definition. As noted the domain is nonempty if and only if  $S$  is  $n$ -printable.

This definition is quite similar to that used in mathematical logic where a sentence is valid for some interpretation if it is true for the interpretation [15,17]. (It is also similar to the notion of accuracy or correctness used by Smullyan [16].) The main difference is that here the domain of definability of  $n, m$ -truth plays an important role. Note that for  $T$  as defined here no sentence with more than  $n-2$  symbols is  $n$ -printable. However,  $T$  is  $n, m$ -valid for all these sentences.

The definition of  $n, m$ -validity for  $S$  is closely connected with the measurement of  $S$  and  $X_S$ , Eq. (17). If  $S$  is  $n$ -printable and  $S = P(X) [S = \sim P(X)]$ , then the  $n, m$ -validity of  $T$  for  $S$  means that the second [first] right-hand term of Eq. (17) is 0.

The equation also shows that  $T$  is  $n, m$ -valid for  $S$  if the probability that  $S$  does not appear in the region  $[0, n-2]$  plus the probability that  $S$  appears in the region and is  $n, m$ -true on its domain of definition equals 1. In terms of matrix elements for the two types of sentences this is expressed by

$$\text{for } P(X) \quad \langle \Psi(n) | (T^\dagger)^m Q_X^h T^m Q_{P(X)}^h | \Psi(n) \rangle + \langle \Psi(n) | Q_{\sim P(X)}^h | \Psi(n) \rangle = 1,$$

$$\text{for } \sim P(X) \quad \langle \Psi(n) | (T^\dagger)^m Q_{-X}^h T^m Q_{\sim P(X)}^h | \Psi(n) \rangle + \langle \Psi(n) | Q_{\sim P(X)}^h | \Psi(n) \rangle = 1. \quad (26)$$

The definition of  $n, m$ -validity for  $S$  can be extended to all sentences.  $T$  is  $n, m$ -valid if it is  $n, m$ -valid for all sentences  $S$ . In terms of single sentence measurements described by Eq. (17) if  $T$  is  $n, m$ -valid then Eqs. (26) hold for each sentence. Measurements can be limited to just those sentences that will fit in the region  $[0, n-2]$  as  $T$  is automatically  $n, m$ -valid for all other sentences because they are not  $n$ -printable.

It may be possible to combine all the single sentence measurements into one measurement observable and all  $X_S$  measurements into another observable. If this is the case then  $n, m$ -validity can be described in terms of one type of measurement. These observables are much more complex than those in Eq. (17). The complexity is shown even for the measurement for just two sentences  $S$  and  $S'$ . In this case the left-hand side of Eq. (17) is replaced by  $U^{X_S} U^{X_{S'}} T^m U^{S'} U^S \Psi(n) |i\rangle_{S'} |i\rangle_S |i\rangle_{X_S} |i\rangle_{X_{S'}}$ . The right-hand side is a sum of nine terms, four for components containing both  $S$  and  $S'$  as each sentence can be  $n, m$ -true or  $n, m$ -false, two for each of the two components containing just one of the two sentences in which  $n, m$ -truth is defined for the sentence appearing, and one for components containing neither sentence. For this term,  $n, m$ -truth is undefined for both sentences. Note that the two unitary operators for  $X_S$  and  $X_{S'}$  commute as do the two for  $S$  and  $S'$ .

If  $T$  is  $n, m$ -valid for both these sentences, then the  $n, m$ -truth conditions give the result that at most four of the nine terms are nonzero. These are the terms that exclude paths showing the  $n, m$ -falseness of either sentence. As a specific example let  $S = P(X)$  and  $S' = \sim P(Y)$ . If  $T$  is  $n, m$ -valid for these two sentences, then only four terms may be nonzero. The term of most interest here is that containing all paths in which both  $P(X)$  and  $\sim P(Y)$  appear and are  $n, m$ -true. This is the state  $Q_{X \wedge \sim Y}^h T^m Q_{P(X) \wedge \sim P(Y)}^h \Psi(n) |1\rangle_{P(X)} |1\rangle_{\sim P(Y)} |1\rangle_X |0\rangle_Y$ . Conservation of probability and the unitarity of  $T$  gives the result:

$$\begin{aligned} & \langle \Psi(n) | (T^\dagger)^m Q_{X \wedge \sim Y}^h T^m Q_{P(X) \wedge \sim P(Y)}^h | \Psi(n) \rangle \\ &= \langle \Psi(n) | Q_{P(X) \wedge \sim P(Y)}^h | \Psi(n) \rangle. \end{aligned}$$

This equation is interesting because if  $X = Y$  then the left-hand matrix element is 0 because the projection operator  $Q_{X \wedge \sim X}^h = 0$ . Thus the following theorem has been proven.

*Theorem 1:* Let  $T$  be  $n, m$ -valid for  $P(X)$  and for  $\sim P(X)$ . Then  $\langle \Psi(n) | Q_{P(X) \wedge \sim P(X)}^h | \Psi(n) \rangle = 0$ .

This theorem is satisfying since it shows that if  $T$  is  $n, m$ -valid for the sentences  $P(X)$  and  $\sim P(X)$ , then these sentences are  $n$ -consistent for  $T$ . Here a sentence and its negation are defined to be  $n$ -consistent for  $T$  if no path in the path sum of Eq. (16) contains both the sentence and its negation in the lattice region  $[0, n-2]$ . That is  $\langle \Psi(n) | Q_{P(X) \wedge \sim P(X)}^h | \Psi(n) \rangle = 0$ .

It is clear that  $n, m$ -validity of  $T$  for  $P(X)$  and  $\sim P(X)$  is a sufficient condition for  $T$  to be  $n$ -consistent for these sentences because without the condition of  $n, m$ -validity for  $T$  there is no reason why a sentence and its negation could not

both appear in some path. It is not a necessary condition since it is possible for  $T$  to be  $n$ -consistent but not  $n, m$ -valid for these two sentences. Note that  $n, m$ -validity of  $T$  for  $P(X)$  and  $\sim P(X)$  does not prevent them from appearing in some paths (i.e., both can be  $n$ -printable). They cannot, however, both appear on the same path.

The dependence of  $n, m$ -truth on  $n$  and  $m$  also results in a similar dependence for  $n, m$ -validity. For example if  $S = P(X)$  then  $T$  may be  $n, m'$ -valid for  $S$  but not  $n, m$ -valid for  $S$  where  $m' > m$ . If  $S = \sim P(X)$  then  $T$  may be  $n, m$ -valid for  $S$  but not  $n, m'$ -valid for  $S$ . A similar dependence holds for changes in  $n$ .

As was done for  $n, m$ -truth, this undesirable  $n, m$ -dependence can be removed by setting  $m = 0$  and taking the limit  $n \rightarrow \infty$ . To this end one defines  $T$  to be valid for  $S$  by the following definition.

*Definition 2:*  $T$  is valid at  $S$  if either  $S$  is not printable or  $S$  is printable and true on its domain of truth definition.

An equivalent statement of the definition is that  $T$  is valid for  $S$  if printability implies truth on the domain of truth definition. Here printability is defined as in Sec. IV A by  $\lim_{n \rightarrow \infty} \langle \Psi(n) | Q_S^h | \Psi(n) \rangle > 0$  and truth by Eqs. (22) and (23) [or Eqs. (24) and (25)].

The definition of validity for a sentence can be extended to all sentences by defining  $T$  to be valid if  $T$  is valid for all sentences. That is  $T$  is valid if for all sentences  $S$  printability of  $S$  implies that  $S$  is true on its domain of definition. This definition can be given in an equivalent form based on Eq. (26):

*Definition 3:*  $T$  is valid if for all expressions  $X$  that are not sentences,

$$\begin{aligned} \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h Q_{P(X)}^h | \Psi(n) \rangle + \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim P(X)}^h | \Psi(n) \rangle &= 1, \\ \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim X}^h Q_{\sim P(X)}^h | \Psi(n) \rangle & \\ + \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim \sim P(X)}^h | \Psi(n) \rangle &= 1. \end{aligned} \quad (27)$$

Note that the limit  $n \rightarrow \infty$  commutes with the sum in Eq. (27).

The definition of  $n$ -consistency can also be extended to the limit  $n \rightarrow \infty$  as follows:

*Definition 4:*  $T$  is consistent for  $P(X)$  and  $\sim P(X)$  if  $\lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{P(X) \wedge \sim P(X)}^h | \Psi(n) \rangle = 0$ .

Since the matrix element in this definition is nondecreasing as  $n$  increases, this definition is equivalent to defining  $T$  to be consistent for  $P(X)$  and  $\sim P(X)$  if for each  $n$ ,  $T$  is  $n$ -consistent for these two sentences.

It should be noted that the use of consistency here seems different from that used in the consistent histories theory of quantum mechanics [32–34]. Here consistency refers to properties of certain pairs of sentences generated by a quantum system. In the consistent histories approach consistency is a property of projectors on a tensor product of Hilbert spaces associated with multitime events or histories of a quantum system [32].

One can use the above definition of consistency and the definition of validity to obtain the following theorem:

*Theorem 2:* If  $T$  is valid for  $P(X)$  and  $\sim P(X)$  then  $T$  is consistent for  $P(X)$  and  $\sim P(X)$ .

To prove this theorem note that by the properties of the projection operators the following two relations hold:

$$\begin{aligned} & \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h Q_{P(X)}^h Q_{\sim P(X)}^h | \Psi(n) \rangle \\ & \leq \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h Q_{\sim P(X)}^h | \Psi(n) \rangle, \end{aligned}$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim X}^h Q_{P(X)}^h Q_{\sim P(X)}^h | \Psi(n) \rangle \\ & \leq \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim X}^h Q_{P(X)}^h | \Psi(n) \rangle. \end{aligned}$$

Since  $T$  is valid for these two sentences, by Eqs. (22) and (23), the right-hand sides of these two inequalities are both equal to 0. Addition of these two inequalities and noting that the limits commute with the sum, and using  $Q_X^h + Q_{\sim X}^h = 1$  gives  $\lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{P(X)}^h Q_{\sim P(X)}^h | \Psi(n) \rangle = 0$  which proves the theorem.

$T$  is said to be consistent if for all expressions  $X$  that are not sentences,  $T$  is consistent for  $P(X)$  and  $\sim P(X)$ . It is clear from the above that if  $T$  is valid then it is consistent.

### C. Completeness

As noted if  $T$  is valid then each sentence is either not printable or it is printable and true. Thus any  $T$  which does not print any sentence at all is valid. Such  $T$  are easy to construct as it is easily decidable which expressions are sentences.

Completeness is used to remove this possibility:

*Definition 5:*  $T$  is complete for a sentence  $S$  if  $S$  is printable.  $T$  is complete if it is complete for all sentences. Otherwise it is incomplete.

As is known from the Gödel incompleteness theorem, [15–17] there exist sets of sentences with axioms that are incomplete. The same situation can occur here. To see this add to the set of sentences [of the form  $P(X)$  and  $\sim P(X)$ ] the two sentences, in Smullyan's example,  $PN(X)$  and  $\sim PN(X)$  where  $X = \sim PN$ . For this value of  $X$   $\sim PN(X)$  is self-referential.

For the interpretation of the sentences used here (and addition of  $N$  to the set of expression symbols) it is easy to prove the following theorem:

*Theorem 3:* Let  $T$  be valid for the sentences  $\sim PN(\sim PN)$  and  $PN(\sim PN)$ . Then neither of these sentences is printable.

To prove this theorem assume  $\sim PN(\sim PN)$  is printable. Since  $T$  is valid for this sentence it must then be true. By the definition of truth, Eq. (23),

$$\begin{aligned} & \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim \sim PN(\sim PN)}^h Q_{\sim PN(\sim PN)}^h | \Psi(n) \rangle \\ & = \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{\sim PN(\sim PN)}^h | \Psi(n) \rangle. \end{aligned}$$

Since the two projection operators appearing in the left-hand matrix element are orthogonal, the matrix element is zero for each  $n$ . Thus both matrix elements in the equality must equal 0 which contradicts the assumption that  $\sim PN(\sim PN)$  is printable. So this sentence is not printable.

For the sentence  $PN(\sim PN)$  the properties of the projection operators give the result

$$\begin{aligned} & \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{PN(\sim PN)}^h Q_{\sim PN(\sim PN)}^h | \Psi(n) \rangle \\ & \leq \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{PN(\sim PN)}^h | \Psi(n) \rangle. \end{aligned}$$

Since  $T$  is valid for this sentence, then either the right-hand matrix element = 0 or it is > 0 and equals the left-hand matrix element (by the definition of truth). But, by theorem 2 on consistency, this is impossible. Thus the right hand matrix element is zero in either case, so  $PN(\sim PN)$  is also not printable, which proves the theorem.

It follows that for this expanded set of sentences  $T$  is not complete. Note that this holds for all  $U$  in the definition of  $T$ , Eq. (1). This suggests that one define a concept of maximal completeness. A valid  $T$  is *maximally complete* if it is complete for all sentences subject only to the requirements of consistency. For the interpretation considered here for the sentences, all uses of consistency including theorem proofs, ultimately depend on the fact that the projection operators  $Q_X^h$  and  $Q_{\sim X}^h$  are orthogonal for any expression  $X$ .

The other self-referential sentence,  $PN(PN)$  (Sec. II), can be included at no cost. The reason is that if it is printable it is trivially true. Thus the requirement that  $T$  be valid for  $PN(PN)$  imposes no restrictions on  $T$ . Completeness for  $PN(PN)$  just requires that the sentence be printable.

### V. THE EXISTENCE OF VALID AND MAXIMALLY COMPLETE $T$

The above considerations raise the issue of the existence of step operators (or generators of the dynamics)  $T$  that are valid and maximally complete for the set of sentences and interpretation considered here. This is not trivial because it is clear from the above that these requirements are quite restrictive. (For the set of sentences used here maximal completeness is equivalent to completeness.)

It is an open question if there exist  $T$  satisfying Eq. (1) that are valid and complete for the sentences considered here. One aspect is that the requirement that the sentences be true requires the presence of strong correlations between the occurrences of a sentence  $S$  and the expression  $X_S$ . If  $S = \sim P(X)$  then  $\sim P(X)$  is true if and only if there is a complete negative correlation between the occurrence of  $\sim P(X)$  and  $X$  no matter how far apart they are.  $P(X)$  is true if and only if there is a complete positive correlation between the occurrence of  $P(X)$  and  $X$ . However, unlike the case for  $\sim P(X)$ , this correlation can be of finite length because once  $X$  occurs in a path containing  $P(X)$  the truth definition is satisfied for all paths that are extensions of the path segment containing  $X$ .

These correlation requirements are shown by Eqs. (24) and (25). In essence, these equations show the deviations

from the condition of no correlation imposed by the truth definitions. No correlation is expressed by

$$\lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{X_S}^h Q_S^h | \Psi(n) \rangle = \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{X_S}^h | \Psi(n) \rangle \times \langle \Psi(n) | Q_S^h | \Psi(n) \rangle.$$

If one relaxes the requirement that  $T$  satisfy Eq. (1) to allow limited backward motion, then there exist both classi-

cal and quantum computer solutions to the existence question. The quantum computer (or robot) proceeds as follows: For each  $n=1,2,\dots,2^M$ , generate the superposition  $1/\sqrt{2}(|0\rangle_a + |\sim\rangle_a) |P(X)0\rangle_{[a+1,b]}$  of states of all  $P(X), \sim P(X)$  where  $X$  is any length  $n$  expression. Then for each  $X$  correlate whether  $X$  is [is not] a sentence (a decidable property) with the states  $|1\rangle_q [ |0\rangle_q ]$  of a qubit  $q$ . Change an adjacent state of  $n$  0s to  $X$  only for the sentence  $P(X)$  and only for those  $X$  that are not sentences. Increase  $n$  by 1 and move to the next region of  $2(n+1)+5$  blank sites.

For each  $n$  the overall state transformations are given by

$$\begin{aligned} |0\rangle_q |0\rangle_{[a,b]} |0\rangle_{[b,c]} &\rightarrow \frac{1}{\sqrt{4^n}} \sum_{X=1}^{4^n} |0\rangle_q \frac{(|0\rangle_a + |\sim\rangle_a)}{\sqrt{2}} |P(X)0\rangle_{[a+1,b]} |0\rangle_{[b,c]} \\ &\rightarrow \frac{1}{\sqrt{4^n}} \left( \sum_X^{-S} |0\rangle_q + \sum_X^S |1\rangle_q \right) \frac{(|0\rangle_a + |\sim\rangle_a)}{\sqrt{2}} |P(X)0\rangle_{[a+1,b]} |0\rangle_{[b,c]} \\ &\rightarrow \frac{1}{\sqrt{4^n}} \sum_X^{-S} |0\rangle_q \frac{(|0P(X)0\rangle_{[a,b]} |X\rangle_{[b,c]} + |\sim P(X)0\rangle_{[a,b]} |0\rangle_{[b,c]})}{\sqrt{2}} \\ &\quad + \sum_X^S |1\rangle_q \frac{(|0\rangle_a + |\sim\rangle_a)}{\sqrt{2}} |P(X)0\rangle_{[a+1,b]} |0\rangle_{[b,c]}. \end{aligned} \tag{28}$$

Here  $b=n+4-a$  and  $c=n-1-b$ . The sums  $\sum_X^{-S}, \sum_X^S$  are over all length  $n$  expressions that are not ( $-S$ ) or are ( $S$ ) sentences. The last line shows that nothing is done to the  $\sum_X^S$  components. The number of length  $n$  expressions that are not sentences is  $4^n - \Delta$  where  $\Delta = 4^{n-3} + 4^{n-4}$ . The quantum robot starts at position  $a$  in internal state  $|0\rangle$  and ends in state  $|0\rangle$  at position  $c+1$  to repeat the cycle for  $n+1$ . The value of  $M$  reflects the presence of a quantum computer with a register of at least  $M$  qubits on board the quantum robot. In this way the quantum robot has with it a record of the active value of  $n$ .

This quantum computer is exponentially efficient in that the number of steps required to generate the final state for a  $T$  that is valid and complete for all sentences  $S$  where the length of  $X_S$  is  $\leq 2^M$  is polynomial in  $(2^M+1)2^M/2$ . The efficiency is shown by the fact that the number of sentences included in the final state is given by  $N = \sum_{n=1}^{2^M} 2(4^n - \Delta)$ .

This efficiency is lost if one wants to determine by measurement of  $X$  values if  $T$  is in fact valid and complete as  $\approx N$  repetitions of the preparation and measurement of the state shown above would have to be carried out. A much more promising approach is to carry out a Fourier transform over the qubytes in region  $[a+3, b-1]$  that give the argument  $X$  of  $P(X), \sim P(X)$  in the  $\sum_X^{-S}$  final-state part of Eq. (28). This gives the state

$$\begin{aligned} \frac{1}{4^n \sqrt{2}} \sum_{Y=0}^{4^n-1} \sum_X^{-S} |0\rangle_q e^{2\pi i Y X / 4^n} [ |0P(Y)0\rangle_{[a,b]} |X\rangle_{[b,c]} \\ + |\sim P(Y)0\rangle_{[a,b]} |0\rangle_{[b,c]} ]. \end{aligned}$$

Here the probability distribution of  $0P(Y)$  as a function of  $Y$  is completely different from that of  $\sim P(Y)$ . For  $0P(Y)$  the distribution is uniformly distributed over all  $Y$  whereas for  $\sim P(Y)$  the distribution is to a good approximation [of order  $(\Delta/4^n)^2$ ] a  $\delta$  function at  $Y=0$ . (Here, depending on context, expressions are either strings of  $n$  symbols or 4-ary numbers). Since these probability distributions are so different it is likely that, as is the case for other quantum algorithms [1,3], they can be determined to good accuracy in polynomially (in  $n$ ) many repetitions of preparation and measurement of  $0P(Y)$  and  $\sim P(Y)$ .

This quantum computer solution for the existence problem refers to an example for which the same quantum system generates both the sentences and the expressions to which the sentences refer. Of more general interest is the case where the quantum system generating the sentences is distinct from the quantum system to which the sentences refer. This is the usual case in physics where the system carrying out measurements is distinct from the system being measured. Study of these systems is deferred to future work.

It is worth noting that the existence problem is strongly related to the set of expressions admitted as sentences. Suppose, following Smullyan [16], one expands the set of sentences by dropping the requirement that  $X$  is not a sentence. This introduces many complexities into the discussion. Suppose for example  $T$  is valid and maximally complete for all sentences in the expanded set. Then the sentence  $P(P(X))$  is printable and true which means that  $P(X)$  is printable and also true. This means informally that all paths containing  $P(P(X))$  also contain  $P(X)$  and all paths containing  $P(X)$

contain  $X$ . Of course  $P(X)$  may be present in paths not containing  $P(P(X))$ .

Application of the same argument to  $P(\sim P(X))$  gives the result that all paths containing this sentence must contain  $\sim P(X)$  and none of the paths containing  $\sim P(X)$  may contain  $X$ . In a similar fashion, none of the paths containing  $\sim P(P(X))$  may contain  $P(X)$  and all paths containing  $P(X)$  contain  $X$ ; for  $\sim P(\sim P(X))$  no path containing this sentence may contain  $\sim P(X)$  and no path containing  $\sim P(X)$  may contain  $X$ .

This is a complex set of requirements especially because each of the eight sentences involved is printable. For example, consistency means that  $P(P(X))$  and  $\sim P(P(X))$  have no paths in common. The same holds for the pair  $P(\sim P(X))$  and  $\sim P(\sim P(X))$ . However, in addition consistency demands that  $P(P(X))$  and  $P(\sim P(X))$  also have no paths in common. The reason is that any path containing both these sentences must contain both  $P(X)$  and  $\sim P(X)$  which is not possible. The same argument fails for the pair  $\sim P(P(X))$  and  $\sim P(\sim P(X))$  because validity and completeness mean that any path containing both these sentences must contain neither  $P(X)$  nor  $\sim P(X)$ . This is possible.

This shows how the complexity of the requirements of validity and completeness for  $T$  grows if one includes sentences of order greater than the first order, the atomic sentences, which is the set considered here. As is shown above the complexity is appreciable even for the eight types of second-order sentences described above.

This also shows quite forcefully that closed inductive definitions, which are used so much in mathematical logic, should be used here only with careful examination of the consequences. To see the problem, note that Smullyan's definition of sentences [16] restricted to sentences of the type  $P(X)$  and  $\sim P(X)$  can be given as (1) All expressions  $P(X)$  and  $\sim P(X)$  where  $X$  is not a sentence are sentences (the atomic sentences), (2) If  $X$  is a sentence so are  $P(X)$  and  $\sim P(X)$ , (3) Sentences are only as defined above.

The problem with this definition resides in the second requirement which expresses closure. Here a definition in terms of inductive orders is more suitable. For each  $k+1$  candidate sentences of order  $k+1$  are defined as those expressions of the form  $P(X)$  or  $\sim P(X)$  where  $X$  is a sentence of order  $k$ . For each  $k=1,2,\dots,T$  is  $k$ -valid and  $k$ -complete if it is valid and complete for all sentences of order  $\leq k$ . If and only if there exist  $T$  that are  $k$ -valid and  $k$ -complete should one consider expanding the set to include the candidate sentences of order  $k+1$  as sentences.

The reason for this is that as the sentence order increases the validity and completeness requirements become increasingly onerous. For example there may well be many  $T$  that are  $k$ -valid and  $k$ -complete but are not  $(k+1)$ -valid and  $(k+1)$ -complete. An example of a  $T$  that is valid and complete for the first-order (atomic) sentences but is not valid and complete for the second-order sentences would be any  $T$  that generates paths containing both  $P(P(X))$  and  $P(\sim P(X))$  and is valid and complete for the atomic sentences.

This situation is even more complex if sentences of the form  $PN(X)$  and  $\sim PN(X)$  are admitted (Sec. II) even if  $X$  is not a sentence. Not only must one deal with the fact that different sentences denote the same expression [e.g.,  $P(Y)$  and  $PN(X)$  where  $Y=X(X)$ ], but for some  $X$  these sentences

generate chains of sentences. A very simple chain of length 2 results from  $X=P$ . If  $PN(P)$  is printable and true so is  $P(P)$ . Since  $P(P)$  is a sentence it is also printable and true. The chain stops here as  $P$  is not a sentence.

Other  $X$  give longer chains. For  $X=PN(P)$  one has

$$PN(PN(P) \rightarrow PN(P(PN(P) \rightarrow P(P(PN(P(P(PN(P) \rightarrow P(PN(P(P(PN(P).$$

The chain terminates because the last expression is not a sentence [no terminal)]. The chain for  $X=PN\sim P$

$$PN(PN(\sim P) \rightarrow PN(\sim P(PN(\sim P) \rightarrow \sim P((PN(\sim P(\sim P(PN(\sim P)$$

stops one stage earlier than the chain for  $X=PN(P)$  because the last sentence, which is true by validity asserts the non-printability of an expression.

There is even a nonterminating chain. To see this set  $X=PN(PN)$ . The first few terms are

$$PN(PN(PN) \rightarrow PN(PN(PN(PN) \rightarrow PN(PN(PN(PN(PN(PN) \rightarrow \dots$$

It is clear that the number of symbols in the successive sentences grows exponentially with increasing position in the chain. If  $N_n$  denotes the number of repetitions of  $PN$  in the  $n$ th position then one sees that  $N_{n+1}=2(N_n-1)$ , where  $N_1=3$ .

## VI. DISCUSSION

It is important to reemphasize that the choice of which expressions are sentences and the particular interpretation assumed for the sentences is external to the quantum enumeration system. It is imposed from the outside. As such the restrictions that validity and completeness place on the dynamics are relative to this interpretation. The quantum system is completely silent on which expressions, if any, are sentences and how they are to be interpreted. This is the case even for  $T$  that are valid and maximally complete. It is a very long way from valid and complete  $T$  as described here to the dynamics of quantum systems that describe to the maximum extent possible their own validity and completeness.

Nevertheless the example described here has aspects that may be useful for a description of systems that describe their own validity and completeness. It is expected, for instance that the definitions of validity, completeness, and possibly consistency will remain. So will some aspects of the definition of truth. It also may be that the notion of printability will remain in more general systems. This follows from the fact that any quantum system that is telling us something about some other quantum system has to print or enumerate strings of symbols as some type of sequence of physical signals. It also has to tell us which groups of signals (i.e., expressions) are meaningful (sentences) and which are meaningless noise. If the system cannot print or enumerate anything it cannot tell us anything.

It should also be noted that there are many other interpretations of the sentences in addition to the interpretation de-

scribed here. A very simple one defines the printability of expressions as is done here but defines the truth of sentences differently. That is  $P(X)$  is true [false] if  $\lim_{n \rightarrow \infty} \langle \Psi(n) | Q_X^h | \Psi(n) \rangle > 0$  [=0] and  $\sim P(X)$  is true [false] if  $P(X)$  is false [true]. The same definitions of validity and completeness can be used to show that consistency is a consequence of validity and that the system is incomplete if the sentences  $PN(\sim PN), \sim PN(\sim PN)$  are included.

This interpretation is much simpler than the one examined in this paper as truth is defined everywhere and  $\sim P(X)$  is true [false] if  $P(X)$  is false [true]. However, it makes no use of the quantum-mechanical nature of the enumeration system. Also, the path sum description of the evolution plays no role in this interpretation as there is no path connection between the occurrence of  $P(X)$  or  $\sim P(X)$  and  $X$ .

The existence of different interpretations or models of the sentences is well known in mathematical logic in that consistent axiom systems have many different inequivalent models [15,17]. For some axiom systems some models are more useful than others. An example is arithmetic where the standard model is almost universally used. However there also exist many nonstandard models of arithmetic which may be useful for some purposes.

In quantum mechanics the freedom of choice of interpretations is much greater than that in classical mechanics. For example, there are many linear combinations of the five basis states  $|P\rangle, |(\cdot), (\cdot)\rangle, |\sim\rangle, |0\rangle$  that also can be used to represent the five symbols. In addition the choice of linear combinations can be different at different lattice sites.

This freedom is similar to the gauge freedom that exists in quantum field theory in that many different gauge choices are possible [35]. This similarity may be quite important in future developments of the ideas presented here.

In spite of the specialized nature of the example, it does serve to introduce the use of mathematical logical concepts such as truth, validity, consistency, and completeness into physics in a fashion similar to how they are defined and used in mathematical logic [15,17]. It is strongly suspected that the definitions of these concepts and their use as restrictions on the generators of the dynamics of a sentence generating system is quite general and applies to the situation where the system that enumerates sentences is distinct from the system to which the sentences refer. In this context the existence question is very important.

It is also suspected that the classical limit of these systems may play an important role. However, the classical limit must be taken so that at each step in the limiting process the dynamics generator refers to its own validity and completeness to the maximum extent possible. It is speculated that this type of limit of quantum robots interacting with environ-

ments may have some important and significant characteristics of intelligence. As such these limit systems may be quite different from classical computers which are also limits of quantum systems.

## ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy, Nuclear Physics Division, under Contract No. W-31-109-ENG-38.

## APPENDIX

In the main text  $n, m$ -truth and  $n, m$ -validity were defined. The value  $m=0$  was chosen and the limit  $n \rightarrow \infty$  of  $n, 0$ -truth was used to define truth and validity. Another way to generate limit definitions is to start with  $n, m$ -truth and take the limit  $m \rightarrow \infty$  to define  $n$ -truth and  $n$ -validity. Truth and validity are then defined by taking the limit  $n \rightarrow \infty$ . The independence of the limit definitions on the choice of  $m$  follows if it can be proved that the two limit definitions are the same. That is one must prove that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \langle \Psi(n) | (T^\dagger)^m Q_{X_S}^h T^m Q_S^h | \Psi(n) \rangle \\ &= \lim_{n \rightarrow \infty} \langle \Psi(n) | Q_{X_S}^h Q_S^h | \Psi(n) \rangle. \end{aligned} \quad (A1)$$

To see that this is the case, note that the following relations hold:

$$\begin{aligned} & \langle \Psi(n) | (T^\dagger)^m Q_{X_S}^h T^m Q_S^h | \Psi(n) \rangle \\ &= \langle \Psi(n) | (T^\dagger)^m Q_{X_S}^h Q_{S, [0, n-2], m}^h T^m | \Psi(n) \rangle \\ &\leq \langle \Psi(n+m) | Q_{X_S}^h Q_S^h | \Psi(n+m) \rangle, \end{aligned}$$

where Eqs. (9) and (6) have been used along with the commutativity of  $Q_{X_S}^h$  and  $Q_S^h$ . Since this is true for each  $n, m$  the left-hand limit is  $\leq$  the right-hand limit.

Conversely for each  $n$  the unitarity of  $T$  and the above noted commutativity and referenced equations give

$$\begin{aligned} & \langle \Psi(n) | Q_{X_S}^h Q_S^h | \Psi(n) \rangle \\ &= \langle \Psi(n) | (T^\dagger)^m Q_{X_S, [0, n-2], m}^h T^m Q_S^h | \Psi(n) \rangle \\ &\leq \langle \Psi(n) | Q_S^h (T^\dagger)^m Q_{X_S}^h T^m Q_S^h | \Psi(n) \rangle, \end{aligned}$$

which completes the proof. It follows from this that Eq. A1 also holds if  $Q_{X_S}^h$  is replaced by  $Q_{\sim X_S}^h$ .

- 
- [1] P. Shor, in *Proceedings, 35th Annual Symposium on the Foundations of Computer Science*, edited by S. Goldwasser (IEEE Computer Society Press, Los Alamitos, CA, 1994), pp. 124–134; *SIAM J. Comput.* **26**, 1481 (1997).  
 [2] L.K. Grover, *Phys. Rev. Lett.* **79**, 325 (1997); G. Brassard,

*Science* **275**, 627 (1997); L.K. Grover, *Phys. Rev. Lett.* **80**, 4329 (1998).

- [3] D.A. Lidar and H. Wang, e-print quant-ph/9807009.  
 [4] D.S. Abrams and S. Lloyd, e-print quant-ph/9807070.  
 [5] N.A. Gershenfeld and I.L. Chuang, *Science* **275**, 350 (1997);

- D.G. Cory, A.F. Fahmy, and T.F. Havel, Proc. Natl. Acad. Sci. USA **94**, 1634 (1997).
- [6] J.I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1994); C. Monroe, D.M. Meekhof, B.E. King, W.M. Itano, and D.J. Wineland, *ibid.* **75**, 4714 (1995); P. Domokos, J.M. Raimond, M. Brune, and S. Haroche, Phys. Rev. A **52**, 3554 (1995); C. Monroe, D. Leibfried, B.E. King, D.M. Meekhof, W.M. Itano, and D.J. Wineland, *ibid.* **55**, R2489 (1997); J.F. Poyatos, J.I. Cirac, and P. Zoller, Phys. Rev. Lett. **81**, 1322 (1998).
- [7] R. Laflamme, C. Miquel, J.P. Paz, and W. Zurek, Phys. Rev. Lett. **77**, 198 (1996); D.P. DiVincenzo and P.W. Shor, *ibid.* **77**, 3260 (1996); E.M. Raines, R.H. Hardin, P.W. Shor, and N.J.A. Sloane, *ibid.* **79**, 953 (1997); E. Knill, R. Laflamme, and W. Zurek, Science **279**, 342 (1998).
- [8] D. Cory, M. Price, W. Maas, E. Knill, R. Laflamme, W. Zurek, T. Havel, and S. Somaroo, Phys. Rev. Lett. **81**, 2152 (1998).
- [9] D.A. Lidar, I.L. Chuang, and K.B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).
- [10] W.H. Zurek, Phys. Rev. D **24**, 1516 (1981); **26**, 1862 (1982); E. Joos and H.D. Zeh, Z. Phys. B **59**, 223 (1985); J.R. Anglin and W.H. Zurek, Phys. Rev. D **53**, 7327 (1996).
- [11] R. Landauer, Phys. Lett. A **217**, 188 (1996); Philos. Trans. R. Soc. London, Ser. A **353**, 367 (1995); Phys. Today **44** (5), 23 (1991); IEEE Trans. Electron Devices **43**, 1637 (1996); S. Haroche and J.-M. Raimond, Phys. Today **49** (8), 51 (1996).
- [12] P. Benioff, Phys. Rev. A **58**, 893 (1998); Superlattices Microstruct. **23**, 407 (1998).
- [13] C.H. Bennett, G. Brassard, E. Bernstein, and U. Vazirani, SIAM J. Comput. **26**, 1510 (1997); W. Van Dam, e-print quant-ph/9805006.
- [14] A. Peres and W.H. Zurek, Am. J. Phys. **50**, 807 (1982).
- [15] J.R. Shoenfield, *Mathematical Logic* (Addison-Wesley, Reading, MA, 1967).
- [16] R. Smullyan, *Gödel's Incompleteness Theorems* (Oxford University Press, Oxford, 1992).
- [17] A. A. Frankel, Y. Bar-hillel, A. Levy, and D. van Dalen, *Foundations of Set Theory*, 2nd revised ed., Vol. 67 of *Studies in Logic and the Foundations of Mathematics* (North-Holland, Amsterdam, 1973).
- [18] Y. Ozhigov, e-print quant-ph/9809015.
- [19] H. Buhrman, R. Cleve, and A. Wigderson, e-print quant-ph/9802040.
- [20] P. Benioff, J. Math. Phys. **17**, 618 (1976); **17**, 629 (1976).
- [21] D. Finkelstein, Int. J. Theor. Phys. **31**, 1627 (1992); in *Current Issues in Quantum Logic*, edited by E. Beltrametti and B. van Frassen (Plenum, New York, 1981), pp. 323–333; *Quantum Relativity* (Springer-Verlag, New York, 1994).
- [22] K. Svozil, Found. Phys. **25**, 1541 (1995).
- [23] V. Krishnamurthy, Int. J. Mod. Phys. C **9**, 213 (1998); K. Svozil, Ann. (N.Y.) Acad. Sci. **755**, 834 (1995); G. Takeuti, in *Current Issues in Quantum Logic*, (Ref. [21]), pp. 303–322; H. Nishimura, Int. J. Theor. Phys. **32**, 1293 (1993); **43**, 229 (1995).
- [24] H. P. Stapp, *Mind, Matter, and Quantum Mechanics* (Springer-Verlag, Berlin, 1993).
- [25] E. Squires, *Conscious Mind in the Physical World* (IOP, Bristol, England, 1990).
- [26] D. Albert, Phys. Lett. **98A**, 249 (1983); Philos. Sci. **54**, 577 (1987); in *Complexity, Entropy and the Physics of Information*, Vol. VIII of SFI Studies in the Sciences of Complexity, edited by W.H. Zurek (Addison-Wesley, Reading, MA, 1990), pp. 471–476.
- [27] T. Breuer, Philos. Sci. **62**, 197 (1995).
- [28] M. Davis, *Computability and Unsolvability* (McGraw-Hill, New York, 1958); C. H. Papadimitriou, *Computational Complexity* (Addison-Wesley, Reading, MA, 1994).
- [29] R. Penrose, *The Emperor's New Mind* (Penguin, New York, 1991).
- [30] R.P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [31] A. Peres, Phys. Lett. **101A**, 249 (1984).
- [32] R.B. Griffiths, Phys. Rev. A **54**, 2759 (1996); Phys. Rev. Lett. **70**, 2201 (1993).
- [33] M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy, and the Physics of Information* (Ref. [26]), Vol. VIII, pp. 425–458; Phys. Rev. D **47**, 3345 (1993).
- [34] R. Omnes, in *Complexity, Entropy, and the Physics of Information* (Ref. [26]), Vol. VIII, pp. 495–512; J. Stat. Phys. **53**, 893 (1988); **53**, 933 (1988); **53**, 957 (1988).
- [35] I. Montavay and G. Münster, *Quantum Fields on a Lattice* (Cambridge University Press, Cambridge, 1994).