Reversible computation as a model for the quantum measurement process*

Karl Svozil†
Institute for Theoretical Physics, Vienna University of Technology,
Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria

One-to-one reversible automata are introduced. Their applicability to a modelling of the quantum mechanical measurement process is discussed.

The connection between information and physical entropy, in particular the entropy increase during computational steps corresponding to an irreversible loss of information—deletion or other many-to-one operations—has raised considerable attention in the physics community [1]. Figure 1 [2] depicts a flow diagram, illustrating the difference between one-to-one, many-to-one and one-to-many computation. Classical reversible computation [2–6] is characterized by a single-valued invertible (i.e., bijective or one-to-one) evolution function. In such cases it is always possible to "reverse the gear" of the evolution, and compute the input from the output, the initial state from the final state.

In irreversible computations, logical functions are performed which do not have a single-valued inverse, such as *and* or *or*; i.e., the input cannot be deduced from the output. Also deletion of information or other many (states)-to-one (state) operations are irreversible. This logical irreversibility is associated with physical irreversibility and requires a minimal heat generation of the computing machine and thus an entropy increase.

It is possible to embed any irreversible computation in an appropriate environment which makes it reversible. For instance, the computer could keep the inputs of previous calculations in successive order. It could save all the information it would otherwise throw away. Or, it could leave markers behind to identify its trail, the *Hänsel and Gretel* strategy described by Landauer [2]. That, of course, might amount to huge overheads in dynamical memory space (and time) and would merely postpone the problem of throwing away unwanted information. But, as has been pointed out by Bennett [4], for classical computations, in which copying and one-to-many operations are still allowed, this overhead could be circumvented by erasing all intermediate results, leaving behind only copies of the output and the original input. Bennett's trick is to perform a computation, copy the resulting output and then, with one output as input, run the computation backward. In order not to consume exceedingly large intermediate storage resources, this strategy could be applied after every single step. Notice that copying can be done reversible in classical physics if the memory used for the copy is initially considered to be blank.

Quantum mechanics, in particular quantum computing,

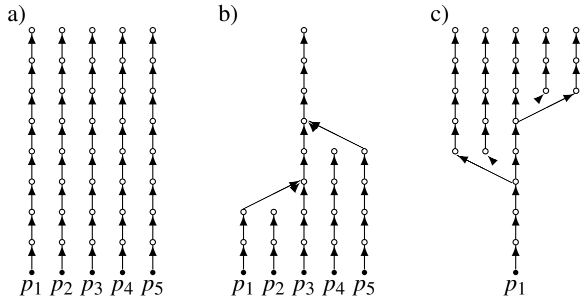

FIG. 1: In this flow diagram, the lowest "root" represents the initial state of the computer. Forward computation represents upwards motion through a sequence of states represented by open circles. Different symbols p_i correspond to different initial computer states. a) One-to-one computation. b) Many-to-one junction which is information discarding. Several computational paths, moving upwards, merge into one. c) One-to-many computation is allowed only if no information is created and discarded; e.g., in copy-type operations on blank memory. From Landauer [2].

teaches us to restrict ourselves even more and exclude any one-to-many operations, in particular copying, and to accept merely one-to-one computational operations corresponding to bijective mappings [cf. Figure 1a)]. This is due to the fact that the unitary evolution of the quantum mechanical state state (between two subsequent measurements) is strictly one-to-one. Per definition, the inverse of a unitary operator U representing a quantum mechanical time evolution always exists. It is again a unitary operator $U^{-1} = U^\dagger$ (where † represents the adjoint operator); i.e., $UU^\dagger = 1$. As a consequence, the *no-cloning theorem* [7–12] states that certain one-to-many operations are not allowed, in particular the copying of general (nonclassical) quantum bits of information.

In what follows we shall consider a particular example of a one-to-one deterministic computation. Although tentative in its present form, this example may illustrate the conceptual strength of reversible computation. Our starting point are finite automata [13–17], but of a very particular, hitherto unknown sort. They are characterized by a finite set S of states, a finite input and output alphabet I and O, respectively. Like for Mealy automata, their temporal evolution and output functions are given by $\delta : S \times I \rightarrow S$, $\lambda : S \times I \rightarrow O$. We additionally require one-to-one reversibility, which we interpret in this context as follows. Let $I = O$, and let the combined (state and output) temporal evolution be associated with a bijective map

$$U : (s,i) \rightarrow (\delta(s,i), \lambda(s,i)), \tag{1}$$

with $s \in S$ and $i \in I$. The state and output symbol could be "fed

*published in "Cybernetics and Systems '98 Volume I, ed. by R. Trappl (Austrian Society for Cybernetic Studies, Vienna, 1998), pp. 102-106"
†Electronic address: svozil@tuwien.ac.at; URL: http://tph.tuwien.ac.at/~svozil

back" consecutively; such that N evolution steps correspond to $U^N = \underbrace{U \cdots U}_{N \text{ times}}$.

The elements of the Cartesian product $S \times I$ can be arranged as a linear list of length n corresponding to a vector. In this sense, U corresponds to a $n \times n$-matrix. Let Ψ_i be the i'th element in the vectorial representation of some (s,i), and let U_{ij} be the element of U in the i'th row and the j'th column. Due to determinism, uniqueness and invertibility,

- $U_{ij} \in \{0,1\}$;
- orthogonality: $U^{-1} = U^t$ (superscript t means transposition) and $(U^{-1})_{ij} = U_{ji}$;
- double stochasticity: the sum of each row and column is one; i.e., $\sum_{i=1}^n U_{ij} = \sum_{j=1}^n U_{ij} = 1$.

Since U is a square matrix whose elements are either one or zero and which has exactly one nonzero entry in each row and exactly one in each column, it is a *permutation matrix*. Let $\mathcal{P}_n$ denote the set of all $n \times n$ permutation matrices. $\mathcal{P}_n$ forms the *permutation group* (sometimes referred to as the *symmetric group*) of degree n. (The product of two permutation matrices is a permutation matrix, the inverse is the transpose and the identity $\mathbf{1}$ belongs to $\mathcal{P}_n$.) $\mathcal{P}_n$ has $n!$ elements. Furthermore, the set of all doubly stochastic matrices forms a convex polyhedron with the permutation matrices as vertices [18, page 82].

Let us be more specific. For $n = 1$, $\mathcal{P}_1 = \{1\}$. For $n = 2$, $\mathcal{P}_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$. For $n = 3$,

$$\mathcal{P}_3 = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \right.$$

$$\left. \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \right\}.$$

The correspondence between permutation matrices and reversible automata is straightforward. Per definition [cf. Equation (1)], every reversible automaton is representable by some permutation matrix. That every $n \times n$ permutation matrix corresponds to an automaton can be demonstrated by considering the simplest case of a one state automaton with n input/output symbols. There exist less trivial identifications. For example, let

$$U = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The transition and output functions of one associated reversible automaton is listed in table I. The associated flow diagram is drawn in Figure 2. Since U has a cycle 3; i.e.,

	δ		λ	
$S\backslash I$	1	2	1	2
s_1	s_2	s_1	1	2
s_2	s_2	s_1	2	1

TABLE I: Transition and output table of a reversible automaton with two states $S = \{s_1, s_2\}$ and two input/output symbols $I = \{1, 2\}$.

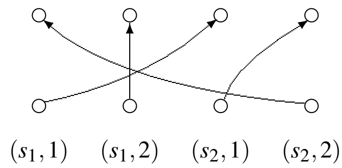

FIG. 2: Flow diagram of one evolution cycle of the reversible automaton listed in Table I.

$(U)^3 = \mathbf{1}$, irrespective of the initial state, the automaton is back at its initial state after three evolution steps. For example, $(s_1,1) \rightarrow (s_2,1) \rightarrow (s_2,2) \rightarrow (s_1,1)$.

The discrete temporal evolution (1) can, in matrix notation, be represented by

$$U\Psi(N) = \Psi(N+1) = U^{N+1}\Psi(0), \tag{2}$$

where again $N = 0,1,2,3,\ldots$ is a discrete time parameter.

Let us come back to our original issue of modelling the measurement process within a system whose states evolve according to a one-to-one evolution. Let us artificially divide such a system into an "inside" and an "outside" region. This can be suitably represented by introducing a black box which contains the "inside" region—the subsystem to be measured, whereas the remaining "outside" region is interpreted as the measurement apparatus. An input and an output interface mediate all interactions of the "inside" with the "outside," of the "observed" and the "observer" by symbolic exchange. Let us assume that, despite such symbolic exchanges via the interfaces (for all practical purposes), to an outside observer what happens inside the black box is a hidden, inaccessible arena. The observed system is like the "black box" drawn in Figure 3.

Throughout temporal evolution, not only is information transformed one-to-one (bijectively, homomorphically) inside the black box, but this information is handled one-to-one *after* it appeared on the black box interfaces. It might seem evident at first glance that the symbols appearing on the interfaces should be treated as classical information. That is, they could in principle be copied. The possibility to copy the experiment (input and output) enables the application of Bennett's argument: in such a case, one keeps the experimental finding by copying it, reverts the system evolution and starts with a "fresh" black box system in its original initial state. The result is a classical Boolean calculus.

The scenario is drastically changed, however, if we assume a one-to-one evolution also for the environment at and outside of the black box. That is, one deals with a homogeneous

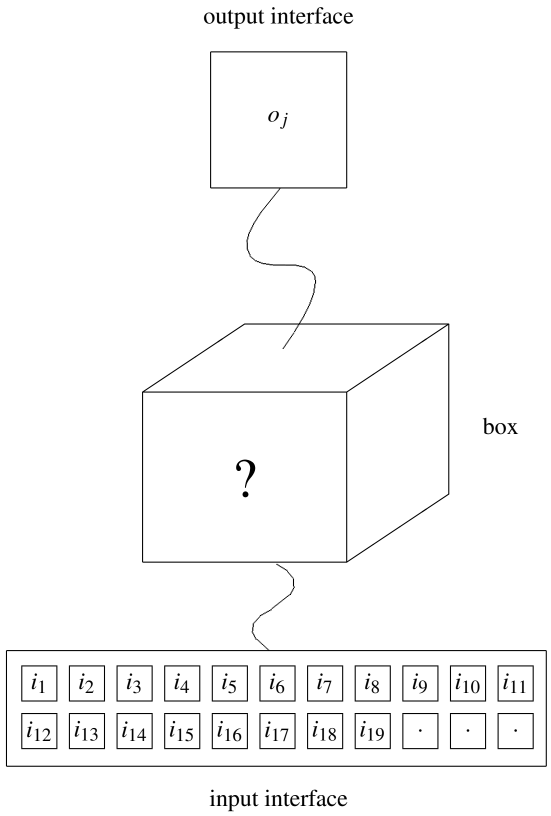

FIG. 3: A system in a black box with an input interface and an output interface.

and uniform one-to-one evolution "inside" and "outside" of the black box, thereby assuming that the experimenter also evolves one-to-one and not classically. In our toy automaton model, this could for instance be realized by some automaton corresponding to a permutation operator U inside the black box, and another reversible automaton corresponding to another U' outside of it. Conventionally, U and U' correspond to the measured system and the measurement device, respectively.

In such a case, as there is no copying due to one-to-one evolution, in order to set back the system to its original initial state, the experimenter would have to erase all knowledge bits of information acquired so far. The experiment would have to evolve back to the initial state of the measurement device and the measured system prior to the measurement. As a result, the representation of measurement results in one-to-one reversible systems may cause a sort of complementarity due to the impossibility to measure all variants of the representation at once.

Let us give a brief example. Consider the 6×6 permutation

	δ		λ	
$S\backslash I$	1	2	1	2
s_1	s_1	s_3	2	2
s_2	s_2	s_1	1	1
s_3	s_3	s_2	1	2

TABLE II: Transition and output table of a reversible automaton with three states $S = \{s_1, s_2, s_3\}$ and two input/output symbols $I = \{1, 2\}$.

matrix

$$U = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

corresponding to a reversible 3-state automaton with two input/output symbols $1, 2$ listed in table II. The evolution is

$$\begin{pmatrix} (s_1,1) \\ (s_1,2) \\ (s_2,1) \\ (s_2,2) \\ (s_3,1) \\ (s_3,2) \end{pmatrix} \xrightarrow{U} \begin{pmatrix} (s_1,2) \\ (s_3,2) \\ (s_2,1) \\ (s_1,1) \\ (s_3,1) \\ (s_2,2) \end{pmatrix} \xrightarrow{U} \begin{pmatrix} (s_3,2) \\ (s_2,2) \\ (s_2,1) \\ (s_1,2) \\ (s_3,1) \\ (s_1,1) \end{pmatrix}$$

$$\xrightarrow{U} \begin{pmatrix} (s_2,2) \\ (s_1,1) \\ (s_2,1) \\ (s_3,2) \\ (s_3,1) \\ (s_1,2) \end{pmatrix} \xrightarrow{U} \begin{pmatrix} (s_1,1) \\ (s_1,2) \\ (s_2,1) \\ (s_2,2) \\ (s_3,1) \\ (s_3,2) \end{pmatrix}.$$

The associated flow diagram is drawn in Figure 4. Thus after the input of just one symbol, the automaton states can be grouped into experimental equivalence classes [19]

$$v(1) = \{\{1\}, \{2, 3\}\}, \quad v(2) = \{\{1, 3\}, \{2\}\}.$$

The associated partition logic corresponds to a non Boolean (nondistributive) partition logic isomorphic to MO_2. Of course, if one develops the automaton further, then, for instance, $v(2222) = \{\{1\}, \{2\}, \{3\}\}$, and the classical case is recovered [notice that this is not the case for $v(\dot{1}) = v(1)$]. Yet, if one assumes that the output is channelled away into the interface after only a single evolution step (and that afterwards the evolution is via another U'), the nonclassical feature pertains despite the bijective character of the evolution.

In this epistemic model, the interface symbolizes the *cut* between the observer and the observed. The cut appears somewhat arbitrary in a computational universe which is assumed to be uniformly reversible.

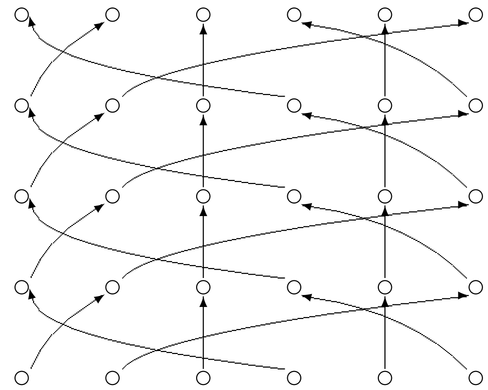

FIG. 4: Flow diagram of four evolution cycles of the reversible automaton listed in Table II.

What has been discussed above is very similar to the opening, closing and reopening of Schrödinger's catalogue of expectation values [20, p. 53]: At least up to a certain magnitude of complexity—any measurement can be "undone" by a proper reconstruction of the wave-function. A necessary condition for this to happen is that *all* information about the original measurement is lost. In Schrödinger's terms, the prediction catalog (the wave function) can be opened only at one particular page. We may close the prediction catalog before reading this page. Then we can open the prediction catalog at another, complementary, page again. By no way we can open the prediction catalog at one page, read and (irreversible) memorize the page, close it; then open it at another, complementary, page. (Two noncomplementary pages which correspond to two co-measurable observables can be read simultaneously.)

From this point of view, it appears that, strictly speaking, irreversibility may turn out to be an inappropriate concept both in computational universes generated by one-to-one evolution as well as for quantum measurement theory. Indeed, irreversibility may have been imposed upon the measurement process rather heuristically and artificially to express the huge practical difficulties associated with any backward evolution, with "reversing the gear", or with reconstructing a coherent state. To quote Landauer [21, section 2],

> *"What is measurement? If it is simply information transfer, that is done all the time inside the computer, and can be done with arbitrary little dissipation."*

Let us conclude with a metaphysical speculation. In a one-to-one invertible universe, any evolution, any step of computation, any single measurement act reminds us of a permanent permutation, reformulation and reiteration of one and the same "message"—a "message" that was there already at the beginning of the universe, which gets transformed but is neither destroyed nor renewed. This thought might be very close to what Schrödinger had in mind when contemplating about Vedic philosophy [22].

[1] H. S. Leff and A. F. Rex, *Maxwell's Demon* (Princeton University Press, Princeton, 1990).
[2] R. Landauer, "Zig-Zag Path to Understanding," in *Proceedings of the Workshop on Physics and Computation PHYSCOMP '94* pp. 54–59 (1994).
[3] R. Landauer, "Irreversibility and Heat Generation in the Computing Process," IBM Journal of Research and Development **3**, 183–191 (1961), reprinted in [1, pp. 188-196].
[4] C. H. Bennett, "Logical Reversibility of Computation," IBM Journal of Research and Development **17**, 525–532 (1973), reprinted in [1, pp. 197-204].
[5] E. Fredkin and T. Toffoli, "Conservative Logic," International Journal of Theoretical Physics **21**, 219–253 (1982), reprinted in [23, Part I, Chapter 3]. http://dx.doi.org/10.1007/BF01857727
[6] C. H. Bennett, "The Thermodynamics of Computation—A Review," in *International Journal of Theoretical Physics* [1], pp. 905–940, reprinted in [1, pp. 213-248].
[7] N. Herbert, "FLASH—A superluminal communicator based upon a new kind of quantum measurement," Foundation of Physics **12**, 1171–1179 (1982).
[8] W. K. Wooters and W. H. Zurek, "A single quantum cannot be cloned," Nature **299**, 802–803 (1982).
[9] L. Mandel, "Is a photon amplifier always polarization dependent?" Nature **304**, 188 (1983).
[10] P. W. Milonni and M. L. Hardies, "Photons cannot always be replicated," Physics Letters **92A**, 321–322 (1982).
[11] R. J. Glauber, "Amplifiers, Attenuators and the Quantum Theory of Measurement," in *Frontiers in Quantum Optics*, E. R. Pikes and S. Sarkar, eds. (Adam Hilger, Bristol, 1986).
[12] C. M. Caves, "Quantum limits on noise in linear amplifiers," Physical Review D **26**, 1817–1839 (1982).
[13] E. F. Moore, "Gedanken-Experiments on Sequential Machines," in *Automata Studies*, C. E. Shannon and J. McCarthy, eds. (Princeton University Press, Princeton, 1956), pp. 129–153.
[14] J. H. Conway, *Regular Algebra and Finite Machines* (Chapman and Hall Ltd., London, 1971).
[15] W. Brauer, *Automatentheorie* (Teubner, Stuttgart, 1984).
[16] M. Schaller and K. Svozil, "Automaton logic," International Journal of Theoretical Physics **35**, 911–940 (1996).
[17] C. Calude, E. Calude, K. Svozil, and S. Yu, "Physical versus Computational Complementarity I," International Journal of Theoretical Physics **36**, 1495–1523 (1997).
[18] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* (Academic Press, New York, 1979).
[19] K. Svozil, *Randomness & Undecidability in Physics* (World Scientific, Singapore, 1993).
[20] E. Schrödinger, "Die gegenwärtige Situation in der Quantenmechanik," Naturwissenschaften **23**, 807–812, 823–828, 844–849 (1935), English translation in [24] and [25, pp. 152-167]. http://dx.doi.org/10.1007/BF01491891, http://dx.doi.org/10.1007/BF01491914, http://dx.doi.org/10.1007/BF01491987

[21] R. Landauer, "Computation, Measurement, Communication and Energy Dissipation," in *Selected Topics in Signal Processing*, S. Haykin, ed. (Prentice Hall, Englewood Cliffs, NJ, 1989), p. 18.

[22] E. Schrödinger, *Mein Leben, meine Weltansicht* (Zsolnay, Vienna, 1985), English translation: My view of the world, Cambridge University Press, Cambridge, 1964.

[23] A. Adamatzky, *Collision-based computing* (Springer, London, 2002).

[24] J. D. Trimmer, "The present situation in quantum mechanics: a translation of Schrödinger's "cat paradox"," Proceedings of the American Philosophical Society **124**, 323–338 (1980), reprinted in [25, pp. 152-167]. http://www.tu-harburg.de/rzt/rzt/it/QM/cat.html

[25] J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, Princeton, 1983).