One-to-One

Reversible computation is a great metaphor for the foundations of physics

KARL SVOZIL
*Institut für Theoretische Physik, University of Technology Vienna,
Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria,
e-mail: svozil@tph.tuwien.ac.at*

Received July 9, 1998; accepted July 9, 1998

INTRODUCTION

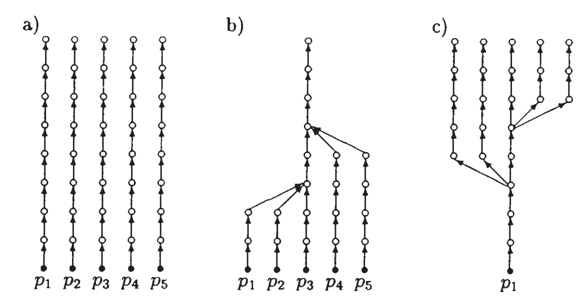
A reversible computation is a computation that can be reversed completely. That is, after insertion of an arbitrary input into a reversible computer, the reversible computer generates some output (if ever). In such a case, one may run the entire computation backward by inserting the output as new input into a reverse computation, and thereby obtaining the input one started with. The computation can flow back and forth an arbitrary number of times. The implicit time symmetry spoils the very notion of "result" since what is a valuable output is determined purely by the subjective meaning the observer associates with it and is devoid of any syntactic relevance.

In more formal terms, reversible computation can be characterized by one to-one operations, by a reversible, bijective evolution of the computer states onto themselves [1–6]. If only a finite number of such states are involved, this amounts to their permutation.

In such a scheme, not a single bit gets lost, and any piece of information (including the overhead) remains in the computer forever. That may be good news for the case of decay and loss of information, but it is bad news with respect to waste management. There is no way of trashing garbage bits other than cleverly compressing them and piling them "high and deep." Stated differently, in this restricted regime, many-to-one operations, such as deletion of bits, are not allowed.

In the strict sense of reversibility discussed here, one-to-many operations, such as copying, are forbidden as well. Any computation can be embedded into a reversible one. The trick is to provide markers in order to make backtracking possible, which amounts to memorizing the past states of the system. If no copying is allowed, this may amount to large space overheads as compared to irreversible computations.

FIGURE 1



In this flow diagram [5], the lowest "root" represents the initial state of the computer. Forward computation represents upward motion through a sequence of states represented by open circles. Different symbols p_i correspond to different computer states: (1) one-to-one computation; (2) many-to-one junction, which is information discarding—several computational paths, moving upward, merge into one; (3) one-to-many computation is allowed only if no information is created and discarded, for example, in copy-type operations on blank memory.

The flow diagram depicted in Figure 1 was introduced by Landauer [5]. It illustrates differences between one-to-one, many-to-one, and one-to-many information flows.

Classical continuum mechanics and electrodynamics are reversible "at heart." That means that all equations of motion are invariant with respect to reversing the arrow of time. Also, quantum theory postulates a unitary evolution of the state between (irreversible) measurements, which by definition is reversible. The no-copy feature of reversible computation is, for instance, reflected by the no-cloning theorem of quantum theory. Therefore, in quantum computations it is not possible to copy arbitrary quantum bits.

There is an undeniable potential technological advantage of reversible computers over irreversible ones. It lies in the fact that reversible computation is not necessarily associated with energy consumption and heat dissipation while the latter one is [6]. And since heat dissipation per computation step can be kept at arbitrary low levels, when "scaled up to very large sizes," reversible computation outperforms irreversible computation in that regime [7]. Moreover, after all, physics at very small scales *is* reversible. At this year's UMC 98 conference in Auckland, New Zealand, an MIT group headed by Thomas Knight presented silicon prototypes of reversible computers [8–10].

REVERSIBLE FINITE AUTOMATA

We shall concentrate on a particular class of reversible finite automata that were first discussed in the UMC 98 workshop [11]. Just as irreversible Mealy automata [12–13], reversible ones will be characterized by the following properties:

- A finite set S of states,
- A finite input alphabet I,
- A finite output alphabet O,
- Temporal evolution function $\delta: S \times I \rightarrow S$,
- Output function $\lambda: S \times I \rightarrow O$.

We additionally require one-to-one reversibility, which we interpret in this context as follows. Assume that the set of input and output symbols is identical, i.e.,

$$I = O.$$

Assume further that a reasonable formalization of reversibility is that the combined (state and output) temporal evolution is associated with a one-to-one (bijective) map

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

with $s \in S$ and $i \in I$. As will be discussed below, neither δ nor λ needs to be a bijection.

The elements of the Cartesian product $S \times I$ can be arranged as a linear list of length n, just like a vector. In this sense, U corresponds to an $n \times n$ matrix. In some analogy to quantum theory, we shall call this matrix U *transition matrix*. Let ψ_i be the ith element in the vectorial representation of some (s, i), and let U_{ij} be the element of U in the ith row and the jth column. Equation (1) can in be rewritten as

$$\Psi_i\left(t+1\right)=U_{ij}\,\Psi_j\,(t),\ \text{or just}\ \Psi\left(t+1\right)=U\,\Psi\,(t). \tag{2}$$

t is a discrete time parameter. Thus in general, the discrete temporal evolution (1) can, in matrix notation, be represented by

$$\Psi\left(t+1\right)=U\,\Psi\,(t)=U^{N+1}\,\Psi\left(0\right), \tag{3}$$

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

Now we shall specify the form of the transition matrix U. Due to determinism, uniqueness, and invertibility, we require

- $U_{ij} \in \{0,1\}$,
- Orthogonality: $U^{-1} = U^t$ (superscript t means transposition) and $(U^{-1})_i = U_{ji}$
- Double stochasticity: The sum of each row and column is one. That is,

$$\sum_{i=1}^{n} U_{ij} = \sum_{j=1}^{n} U_{ij} = 1$$

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 P_n denote the set of all $n \times n$ permutation matrices. P_n forms the *permutation group* (sometimes referred to as the *symmetric group*) of degree n [17]. (The product of two permutation matrices is a permutation matrix, the inverse is the transpose, and the identity 1 belongs to P_n) P_n has $n!$ elements.

The simplest case is $n = 1$. It is just the identity $P_1 = \{1\}$.

The first nontrivial case is $n = 2$. The permutation matrices of

$$p_2=\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix}\right\},\left\{\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\}$$

correspond to the identity and the *not*-gate.

Before we consider more examples, let us mention that the correspondence between permutation matrices and reversible automata is straightforward. (Indeed, by taking the pairs $(s,i) \in S \times I$ as states of a new finite automaton (with empty output), the permutation matrix is just the adjacency matrix of the transition diagram of this automaton [18–20].) 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. In this particular but rather trivial case, the transition function is many-to-one (in fact, n-to-one), but the output function is one-to-one (in fact, n-to-n).

There exist less trivial identifications. For example, let

$$U_1 = \mathbf{I} = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix},\ U_2 = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix},$$

$$U_3 = \begin{pmatrix} 0\,1\,0\,0 \\ 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix},\ U_4 = U_2 U_3 = \begin{pmatrix} 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix}$$

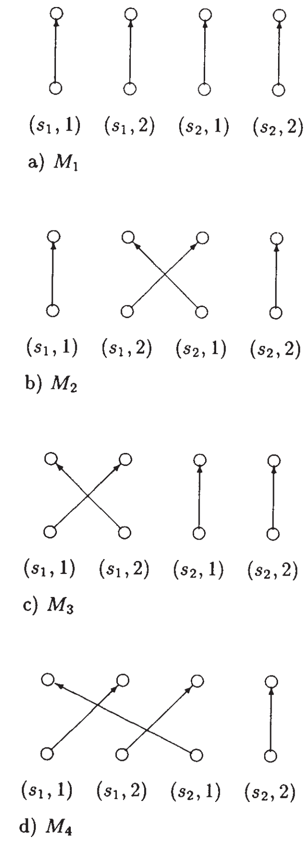
The transition and output functions of the four corresponding reversible automa are listed in Table 1. The associated flow diagrams are drawn in Figure 2.

TABLE 1

S / I	δ 1	δ 2	λ 1	λ 2
M_1				
S_1	S_1	S_1	1	2
S_2	S_2	S_2	1	2
M_2				
S_1	S_1	S_2	1	1
S_2	S_1	S_2	2	2
M_3				
S_1	S_1	S_1	2	1
S_2	S_2	S_2	1	2
M_4				
S_1	S_1	S_2	2	1
S_2	S_1	S_2	1	2

Transition and output table of four reversible automata M_1, M_2, M_3 and M_4 with two states $S = \{S_1, S_2\}$ and two input/output symbols $I = \{1,2\}$. For M_1, the transition as well as the output function is one-to-one. For M_2, the transition function is many-to-one, but the output function is one-to-one. For M_3, the transition function is one-to-one but the output function is many-to one. M_4 is a concatenation of M_3 and M_2. Both its transition function as well as its output function are many-to-one.

FIGURE 2



Flow diagram of one evolution cycle of the reversible automata listed in Table 1.

Note that, for wxample, M_4 is a serial composition of M_2 amd M_3. The final example is based on the permutation matrix

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

It can be realized by a reversible automaton, which is represented in Table 2. Neither its evolution function nor its transition function is one-to-one since, for example, $\delta(s_1, 3) = \delta(s_2, 1) = s_2$, and $\lambda(s_1, 2) = \lambda(s_1, 3) = 2$. Its flow diagram throughout five evolution steps is depicted in Figure 3.

MEASUREMENTS

Let us now attempt to model the measurement process within a system whose states evolve according to a one-to-one evo-

TABLE 2

$S\backslash I$	δ 1	2	3	λ 1	2	3
s_1	s_1	s_1	s_2	1	2	2
s_2	s_2	s_2	s_1	1	3	3

Transition and output table of a reversible automaton with two states $S = \{s_1, s_2\}$ and three input/output symbols $I = \{1, 2, 3\}$. Neither its transition nor its output function is one-to-one.

lution. This is distinct from the orthodox quantum mechanical conception of an irreversible measurement associated with the reduction of the state vector or with the notorious “wave function collapse.”

In what follows we shall artificially divide a reversible system into an “inside” and an “outside” region [21–29]. This can be suitably represented by introducing a black box that 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. This establishes a (arguably artificial) *cut* between the observer and the observed.

Throughout temporal evolution, not only is information transformed one to-one (bijectively, isomorphically) 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 strategy [27]: 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 with no computational complementarity [30].

The scenario is drastically changed, however, if we assume a one-to-one evolution also for the environment at and outside the black box. That is, one deals with a homogeneous and uniform one-to-one evolution “inside” and “out side” 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 invest 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. This is similar to the opening, closing, and reopening of Schrödinger’s catalog of expectation values [31–33].

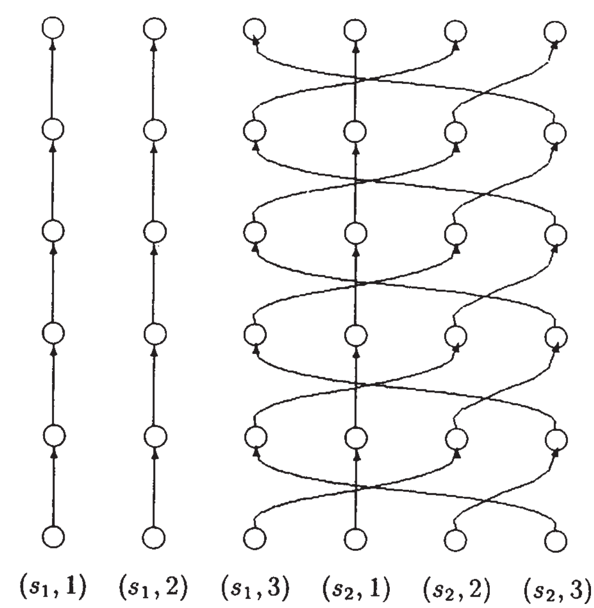
As a result, the representation of measurement results in one-to-one reversible systems may cause a sort of complementarity due to the impossibility of measuring all variants of the representation at once.

AFTERTHOUGHTS

One algorithmic aspect of reversibility seems disturbing. If reversible computation is just a rephrasing, a permutation of the input, then what use is it anyway? The “garbage in/garbage out” metaphor is particularly pressing here.

This issue seems to be somewhat related to an old question in proof theory, in which sense a proof of a statement is “better” than the mere knowledge that this statement is true. It may be that the term “informative” can only be given a subjective, idealistic meaning devoid of any formal rigor.

FIGURE 3



Flow diagram of five evolution cycles of the reversible automaton listed in Table 2.

REFERENCES

1. R. Landauer: Irreversibility and heat generation in the computing process. IBM Journal of Research and Development. 3: pp. 183-191, 1961.
2. C. H. Bennett: Logical reversibility of computation. IBM Journal of Research and Development 17: pp. 525-532, 1973.
3. E. Fredkin, and T. Toffoli: Conservative logic. International Journal of Theoretical Physics. 21: pp. 219-253, 1982.
4. C. H. Bennett: The thermodynamics of computation—a review. International Journal of Theoretical Physics. pp. 905-940.
5. R. Landauer: Zig-zag path to understanding. Proceedings of the Workshop on Physics and Computation. PHYSCOMP '94 (Los Alamitos, CA), IEEE Computer Society Press, 1994.
6. H. S. Leff and A. F. Rex: Maxwell's demon. Princeton, N.J.: Princeton University Press, 1990.
7. M. Frank, T. Knight, and N. H. Margolus: Reversibility in optimally scalable computer architectures. In: Unconventional models of computation. C.S. Calude, J. Casti, and M.J. Dinneen, (Eds.) Singapore: Springer, 1998. pp. 165-182.
8. M. Frank, C. Vieri, J. Ammer, N. Love, N. H. Margolus, and T. Knight: A scalable reversible computer in silicon. In: Unconventional models of computation.
9. T. Knight, and G. J. Sussman: Cellular gate technology. In: Unconventional models of computation. C.S. Calude, S. Cris tian, J. Casti, and M. J. Dinneen, (eds.), Singapore: Springer, 1998.
10. C. Vieri, J. Ammer, A. Wakefield, L. Svensson, W. Athas, and T. Knight: Designing reversible memory. In: Unconventional models of computation. C.S. Calude, J. Casti, and M.J. Dinneen (Eds.), Singapore: Springer, 1998.
11. K. Svozil. The Church-Turing thesis as a guiding principle for physics. In: Unconventional models of computation. C.S. Calude, J. Casti, and M.J. Dinneen (Eds.), Singapore: Springer, 1998.
12. J. E. Hopcroft, and J. D. Ullman: Introduction to automata theory, languages, and computation. Reading, MA: Addison-Wesley, 1979.
13. W. Brauer: Automatentheorie. Stuttgart: Teubner, 1984.
14. A. Landé: Quantum mechanics in a new key. New York: Exposition Press, 1973.
15. A. Peres, Quantum theory: concepts and methods. Dordrecht: Kluwer Academic Publishers, 1993.
16. J.D. Louck: Doubly stochastic matrices in quantum mechanics. Foundations of Physics. 27: pp. 1085-1104; 1997.
17. J. S. Lomont, Applications of finite groups. New York: Academic Press, 1959.
18. D. Lind, and B. Marcus: An introduction to symbolic dynamics and coding. Cambridge, MA: Cambridge University Press, 1995.
19. S. Eilenberg: Automata, languages, and machines. Vol. A, New York: Academic Press, 1974.
20. N. Biggs: Algebraic graph theory, Second ed. Cambridge, MA: Cambridge University Press, 1993.
21. R. J. Boskovich: De spacio et tempore, ut a nobis cognoscuntur. In: A theory of natural philosophy. J. M. Child (Ed.) Cambridge, MA: MIT Press, 1966.
22. T. Toffoli: The role of the observer in uniform systems. In: Applied general systems research. G. Klir, (Ed.), New York: Plenum Press, 1978.
23. K. Svozil: On the setting of scales for space and time in arbitrary quantized media. Lawrence Berkeley Laboratory preprint LBL-16097, 1983.
24. K. Svozil: Connections between deviations from lorentz transfor mation and relativistic energy-momentum relation. Europhysics Letters. 2: pp. 83-85, 1986.
25. K. Svozil: Operational perception of space-time coordinates in a quantum medium. Il Nuovo Cimento. 96B: pp. 127-139, 1986.
26. O.E. Rössler: Endophysics. Real brains, artificial minds. J. Casti and A. Karlquist (Eds.) New York: North-Holland. 1987.
27. O.E. Rössler: Endophysics, die Welt des inneren Beobachters. Berlin: Merwe Verlag, 1992.
28. O.E. Rössler: Endophysics: the world as an interface. Singapore: World Scientific, 1998.
29. K. Svozil: Randomness & undecidability in physics. Singapore: World Scientific, 1993.
30. C. Calude, E. Calude, and C. Stefanescu: Computational complementarity for Mealy automata. CDMTCS Research Report. 83, April 1998. C.S. Calude, J. Casti, and M.J. Dinneen (Eds.) Singapore: Springer, 1998, pp. 183-200.
31. J. D. Trimmer: The present situation in quantum mechanics: A translation of Schrödinger's "cat paradox." Proc. Am. Phil. Soc. 124: pp. 323-338, 1980.
32. D.B. Greenberger and A. YaSin: "Haunted" measurements in quantum theory. Foundation of Physics. 19 (6): pp. 679-704, 1989.
33. T.J. Herzog, P.G. Kwiat, H. Weinfurter, and A. Zeilinger: Complementarity and the quantum eraser. Physical Review Letters. 75 (17): pp. 3034-3037, 1995.