International Journal of Theoretical Physics, Vol. 39, No. 3, 2000

Logic of Reversible Automata[†]

K. Svozil[1]

Received December 8, 1999

The empirical logic of reversible automata is studied.

1. GENERAL DISCUSSION

Reversible computation is a computation which can be reversed completely: one may run the entire computation backward by inserting the output as new input, thereby obtaining the input one started with.

In more formal terms, reversible computation can be characterized by one-to-one operations; i.e., by a reversible, bijective evolution of the computer states onto themselves [Lan61, Ben73, FT82, Ben82, Lan94, LH90]. If only a finite number of such states are involved, this amounts to their permutation. For such a restricted regime, many-to-one operations such as deletion of bits or one-to-many operations such as copying are not allowed.

The flow diagram depicted in Fig. 1 was introduced by Landauer [Lan94]. It illustrates differences between one-to-one, many-to-one, and one-to-many information flows.

We shall concentrate on a particular class of reversible *finite* automata which were first discussed in the UMC'98 workshop in Auckland [Svo98b]. Just like irreversible Mealy automata [HU79, Bra84], 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 transition function $\delta\colon S \times I \to S$

[†] This paper is dedicated to the memory of Prof. Gottfried T. Rüttimann.

[1] Institut für Theoretische Physik, University of Technology Vienna, A-1040 Vienna, Austria; e-mail: svozil@tph.tuwien.ac.at.

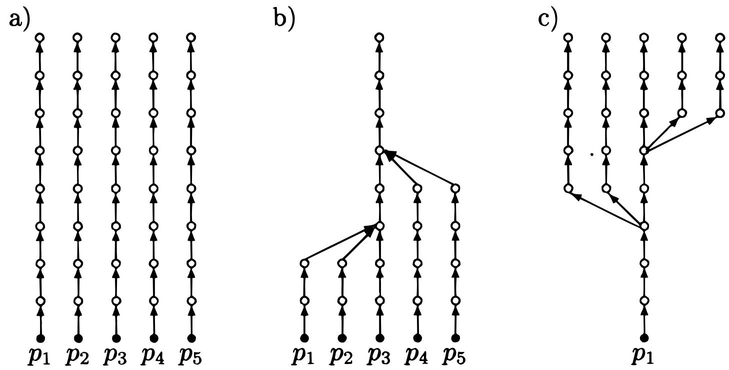

Fig. 1. In this flow diagram, the lowest "root" represents the initial state of the computer. Forward computation represents upward motion through a sequence of states represented by open circles. Different symbols correspond to different computer states. (a) One-to-one computation. (b) Many-to-one junction, which is information discarding. Several computational paths, moving upward, merge into one. (c) One-to-many computation is allowed only if no information is created and discarded; e.g., in copy operations on blank memory. From Landauer [Lan94].

- output function $\lambda: S \times I \to O$

We additionally require one-to-one reversibility and assume that the set of input and output symbols is identical, i.e.,

$$I = O.$$

Further we shall require that the combined (state and output) temporal evolution is associated with a one-to-one (bijective) map

$$u: (s, i) \to (\delta(s, i), \lambda(s, i)) \tag{1}$$

with $s \in S$ and $i \in I$. As will be discussed below, neither δ nor λ needs to be a bijection. Note that the temporal evolution is characterized by the transition and output function *combined*.

The elements of the Cartesian product $S \times I$ can be arranged as a linear list of length n, just like a vector Ψ; i.e., Ψ_j is the jth element in the vectorial representation of (s, i). In this sense, u can be identified with an $n \times n$-matrix which will be denoted by U. In analogy to quantum theory, we shall call this matrix U *the evolution matrix*. Let U_{jk} be the element of U in the jth row and the kth column.

We shall specify the form of the evolution matrix U next. Due to conditions of determinism, uniqueness, and invertibility, we require the following:

- $U_{jk} = \begin{cases} 1 & \text{iff the image of the } j\text{th element is the } k\text{th element} \\ 0 & \text{else} \end{cases}$

That is, for a discrete time evolution labeled by $T \in \mathbb{Z}$, $\Psi(T+1) = U\Psi(n)$.

- Orthogonality: $U^{-1} = U^t$ (superscript t means transposition); i.e., $(U^{-1})_{jk} = U_{kj}$.
- Doubly stochasticity: the sum of each row and column is one. That is, $\sum_{j=1}^{n} U_{jk} = \sum_{k=1}^{n} U_{jk} = 1$ [Lan73, Per93, Lou97].

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*.

Before we consider examples, let us mention the connection between permutation matrices and reversible automata. In fact, the correspondence between permutation matrices and reversible automata is straightforward.[2] 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

$$U_1 = \mathbb{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$U_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad U_4 = U_2U_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The transition and output functions of the four corresponding reversible automata are listed in Table I. Here, we have made the identifications

$$\Psi = ((s_1, \lambda_1), (s_1, \lambda_2), (s_2, \lambda_1), (s_2, \lambda_2))$$

The associated flow diagrams are drawn in Fig. 2.

Let us now attempt to model the measurement process within a system whose states evolve according to a one-to-one evolution. This is distinct from the orthodox quantum mechanical conception of an irreversible measurement

[2] 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 [LM95, Ei174, Big93].

Table I. 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 $I = \{1, 2\}$[a]

	δ		λ	
$S \backslash I$	1	2	1	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

[a] 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.

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 [Bos55, Tof78, Svo83, Svo86a, Svo86b, Rös87, Rös92, GW92, Svo93, Chapter 6]. 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. This establishes an (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 appears 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, which could in principle be copied.

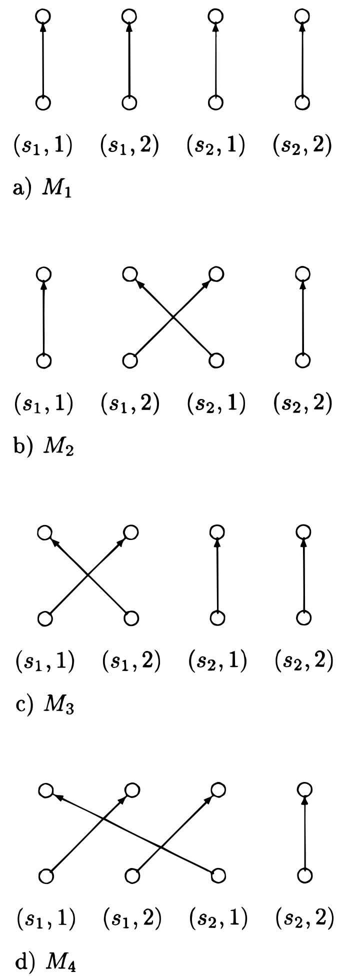

Fig. 2. Flow diagram of one evolution cycle of the reversible automata listed in Table I.

The possibility to copy the experiment (input and output) enables the application of Bennett's strategy [Ben73]: in such a case, one keeps the experimental finding by copying it, reverses 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 [CCŞ98].

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 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, and 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 catalogue of expectation values [Sch35, p. 823; GY89, HKWZ95]).

As a result, the representation of measurement results in one-to-one reversible systems may cause a particular class [Svo93, SS96, DPS95, Svo98a] of complementarity due to the impossibility of measuring all variants of the representation at once.

REFERENCES

[Ben73] Bennett, C. H., Logical reversibility of computation, *IBM Journal of Research and Development*, **17** (1973), 525–532 [Reprinted in LR90, pp. 197–204].

[Ben82] Bennett, C. H., The thermodynamics of computation—A review, *International Journal of Theoretical Physics*, (1982), 905–940 [Reprinted in LR90, pp. 213–248].

[Big93] Biggs, N., *Algebraic Graph Theory*, 2nd ed., Cambridge University Press, Cambridge, 1993.

[Bos55] Boskovich, R. J., *De spacio et tempore, ut a nobis cognoscuntur*, Vienna, 1755 [English translation in Bos66].

[Bos66] Boskovich, R. J., *A Theory of Natural Philosophy*, MIT Press, Cambridge, Massachusetts, 1966 pp. 203–205.

[Bra84] Brauer, W., *Automatentheorie*, Teubner, Stuttgart, 1984.

[CCŞ98] Calude, C., Calude, E., and Ştefănescu, C., Computational complementarity for Mealy automata, CDMTCS preprint, 1998.

[BPS95] Dvurečenskij, A., Pulmannová, S., and Svozil, K., Partition logics, orthoalgebras and automata, *Helvetica Physica Acta*, **68** (1995), 407–428.

[Eil74] Eilenberg, S., *Automata, Languages, and Machines*, Vol. A, Academic Press, New York, 1974.

[FT82] Fredkin, E., and Toffoli, T., Conservative logic, *International Journal of Theoretical Physics*, **21** (1982), 219–253.

[GW92] Gerbel, K., and Weibel, P. (eds.), *Die Welt von Innen—Endo & Nano (the world from within—endo & nano)*, PVS, Linz, Austria, 1992.

[GY89] Greenberger, D. B., and YaSin, A., "Haunted" measurements in quantum theory, *Foundation of Physics*, **19** (1989), 679–704.

[HKWZ95] Herzog, T. J., Kwiat, P. G., Weinfurter, H., and Zeilinger, A. Complementarity and the quantum eraser, *Physical Review Letters*, **75** (1995), 3034–3037.

[HU79] Hopcroft, J. E., and Ullman, J. D., *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, Massachusetts, 1979.

[Lan61] Landauer, R., Irreversibility and heat generation in the computing process, *IBM Journal of Research and Development*, **3** (1961), 183–191 [Reprinted in LR90, pp. 188–196].

[Lan73] Landé, A., *Quantum Mechanics in a New Key,* Exposition Press, New York, 1973.

[Lan94] Landauer, R., Zig-zag path to understanding, in *Proceedings of the Workshop on Physics and Computation PHYSCOMP '94*, IEEE Computer Society Press, 1994; pp. 54–59.

[LM95] Lind, D., and Marcus, B., *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, Cambridge, 1995.

[Lom59] Lomont, J. S., *Applications of Finite Groups*, Academic Press, New York, 1959.

[Lou97] Louck, J. D., Doubly stochastic matrices in quantum mechanics, *Foundations of Physics*, **27** (1997), 1085–1104.

[LR90] Leff, H. S., and Rex, A. F., *Maxwell's Demon*, Princeton University Press, Princeton, New Jersey, 1990.

[PER93] Peres, A., *Quantum Theory: Concepts and Methods*, Kluwer, Dordrecht, 1993.

[Rös87] Rössler, O. E., Endophysics, in *Real Brains, Artificial Minds*, Casti, J. L., and Karlquist, A., eds., North-Holland, New York, 1987, p. 25.

[Rös92] Rössler, O. E., *Endophysics. Die Welt des inneren Beobachters*, Merwe Verlag, Berlin, 1992.

[Sch35] Schrödinger, E., Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften*, **23** (1935), 807–812, 823–828, 844–849 [English translation in Tri80 and WZ83, pp. 152–167].

[SS96] Schaller, M., and Svozil, K., Automaton logic, *International Journal of Theoretical Physics*, **35** (1996), 911–940.

[Svo83] Svozil, K., On the setting of scales for space and time in arbitrary quantized media, Lawrence Berkeley Laboratory preprint LBL-16097 (1983).

[Svo86a] Svozil, K., Connections between deviations from Lorentz transformation and relativistic energy-momentum relation, *Europhysics Letters*, **2** (1986), 83–85.

[Svo86b] Svozil, K., Operational perception of space-time coordinates in a quantum medium, *Nuovo Cimento*, **96B** (1986), 127–139.

[Svo93] Svozil, K., *Randomness and Undecidability in Physics*, World Scientific, Singapore, 1993.

[Svo98a] Svozil, K., *Quantum Logic*, Springer, Singapore, 1998.

[Svo98b] Svozil, K., The Church–Turing thesis as a guiding principle for physics, in *Unconventional Models of Computation*, Calude, C. S., Casti, J., and Ninneen, M. J., eds., Springer, 1998, pp. 371–385.

[Tof78] Toffoli, T., The role of the observer in uniform systems, in *Applied General Systems Research*, Klir, G., ed., Plenum Press, New York, 1978.

[Tri80] Trimmer, J. D., The present situation in quantum mechanics: A translation of Schrödinger's "cat paradox," *Proc. Am. Phil. Soc.* **124** (1980), 323–338 [Reprinted in WZ83, pp. 152–167].

[WZ83] Wheeler, J. A., and Zurek, W. H., *Quantum Theory and Measurement*, Princeton University Press, Princeton, New Jersey 1983.