AAAS Annual Meeting
Atlanta
16-21 February 1995

AUTHOR: Marion Boykan Pour-El

AFFILIATION: University of Minnesota

SESSION TITLE: The Legacy of Alan Turing

PRESENTATION:

TIME: 2:30 p.m. **DATE:** Sun. 2/19/95

NOTE TO CORRESPONDENTS: Release time is the time of presentation, or time of news briefing, if scheduled.

Computability in Mathematics and Physical Theory

Marian Boykan Pour-El

The present computer age owes much to the contributions of Alan Turing. He is best known for the Turing machine, the most general digital computer. All digital computers - e.g. the MAC, the IBM-PC, etc. - are special cases of the Turing machine. Turing's work provides an answer to the question: what is a computable process?

Turing studied discrete processes. In recent years, research in computability has been extended to encompass continuous processes as well. These include the computability/noncomputability of physical phenomena (e.g. light, heat, etc.). Pour-El/Richards have proved the following. Wave propagation (e.g. the propagation of light) can be noncomputable even though the initial data which completely determines the wave propagation is computable. By contrast, heat dissipation is always computable if the initial data is. Actually, the results are quite general. They explore the gamut of processes in physics and mathematics from the viewpoint of computability. A brief summary will be given in the talk.

The results have implications for philosophy. For example, what is the relation between computability and measurement? There are also implications for artificial intelligence. Is the human mind merely a Turing machine, or is something else involved? Both of these topics will be discussed.

From: McGraw-Hill Yearbook of Science + Technology, 1994

other for transmit, in conjunction with the receive and transmit feeds and beam-forming networks, form four or five L-band spot beams plus the L-band global beams. As noted above, the transmit feed elements are connected to solid-state power amplifiers, whereas each of the receive feed elements is connected to a low-noise amplifier. The INMARSAT satellites are assigned locations on the geosynchronous orbit to cover the Atlantic Ocean at 26° and 55° west longitude, the Indian Ocean at 63° east longitude, and the Pacific Ocean at 180° east longitude. For optimum coverage of each region with any of the launched satellites, the beam positions and shape are readjustable to a limited extent by ground commands.

The radiated signals at C band and L band are circularly polarized, as shown in Fig. 3 by the designations X and Y for the left- and right-hand senses. The dual polarization applies only to the C-band transmissions to and from the shore stations. The L-band helix-type feeds do not lend themselves to dual polarization. Thus, with two sets of transponders, the dual-polarization capability of the C-band antennas, and the requirement that the L-band multibeam antenna system achieve a beam-to-beam isolation of greater than 18 dB, spectrum use and communications capacity are doubled.

Navigation payload. The U.S. Global Positioning System (GPS) satellites and the Russian GLONASS satellites allow fixed or mobile users to determine their positions to within 10 m (33 ft), and speed with a tolerance of ±0.1 m/s (±0.3 ft/s), when four satellites are visible and selective availability (purposeful accuracy degradation) is not activated. The GPS satellites are in an orbit inclined 63° to the Equator at a nominal altitude of 20,000 km (12,400 mi; 12-h orbit) and operate at a frequency of 1575.42 MHz. The GLONASS satellites operate at the same frequency but at an orbit altitude of 19,000 km (11,800 mi) and an inclination of 64.8°. The satellites transmit a highly accurate time signal, derived from on-board atomic clocks, and the satellite's ephemeris, periodically updated by a controlling earth station. The information allows the user's receiver to calculate position by measuring the time difference between the signals received from two or more satellites. The speed is calculated by observing the rate of change of position.

INMARSAT 3 carries a C-L-band transponder that will furnish similar signals, but the ephemeris and time signals are generated at an earth station and relayed through the satellite's global beam at the same frequency band as the GPS. The technique allows changing the information carried by the signals in real time, and it also simplifies the equipment needed on the satellite. The satellite transmission can also broadcast system status information.

The global-coverage navigation beam transmits an effective isotropic radiated power (e.i.r.p.) of 27.5 decibels above 1 watt (dBW). A separate 0.7-m-diameter (2.1-ft) parabolic reflector antenna, on the Earth-facing side of the satellite, yields an edge-of-Earth gain of about 16 decibels above isotropic power (dBi). The power amplifiers are the same design as those in the communications transponders, yielding 17 W output power. The power assures that the signal at the mobile receiver is the same level as that obtained from the GPS, although the altitude of *INMARSAT 3* is approximately 16,000 km (10,000 mi) higher. To achieve an accuracy comparable to the coarse/acquisition GPS signal, which is available to civil (nonmilitary) users, the transponder bandwidth is 2.2 MHz. The composite signal delay through the transponder must be very constant over the range of operating temperature and is specified as 10 nanoseconds. In parallel with the C-band uplink to L-band downlink primary path, a low-power (0-dBW) C-to-C-band link is included. This will be used by control earth stations to adjust uplink timing.

For background information SEE COMMUNICATIONS SATELLITE; SATELLITE NAVIGATION SYSTEMS; SURFACE-ACOUSTIC-WAVE DEVICES in the McGraw-Hill Encyclopedia of Science & Technology.

Louis Pollack

Bibliography. American Institute of Aeronautics and Astronautics, *Proceedings of the 14th International Communications Satellite Systems Conference*, part 1, March 1992; S. Egami and M. Kawai, An adaptive multiple beam system concept, *IEEE J. Selected Areas Commun.*, SAC-5(4):630–636, 1987; D. W. Lipke et al., MARISAT: A maritime satellite communications system, *COMSAT Tech. Rev.*, 7(2):351–391, 1977; P. Wood, Mobile satellite services for travellers, *IEEE Commun. Mag.*, 29(11):32–35, 1991.

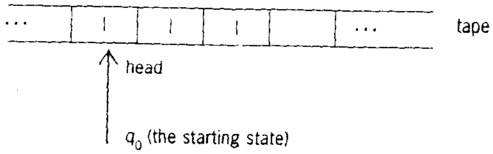
Computability (physics)

Does light propagate computably? Does heat dissipate computably? In order to answer these and related questions, an understanding of the nature of computability from the theoretical point of view is essential.

The study of the properties of an arbitrary computable process goes back at least to the 1930s, when several mathematical logicians—A. Turing, E. Post, J. Herbrand, K. Gödel, and A. Church formulated definitions of the term computable function. Although each of these definitions represented a different approach toward the subject, they were all shown to be equivalent. The concept of computability that emerged captures the intuitive notion quite well. Today, it provides the foundation for theoretical work in computer science.

In one respect, all the definitions agree. Both the domain and the range of a computable function can be encoded as subsets of the nonnegative integers N. This will be discussed below.

Turing-machine approach. It is instructive to consider Turing's work in some detail. Turing reasoned that the class of computable functions should consist precisely of those functions—with domain and range included in N—that can be computed by a digital computer. This led him to define the Turing machine, the most general type of digital computer.



Turing machine at the beginning of the computation discussed in the text.

The Turing machine consists of a read-write head and a tape divided into squares. The head can perform atomic acts of the types given in (1).

$$\begin{array}{l} \text{Print a letter of the alphabet} \\ \text{Move one square to the right: } R \\ \text{Move one square to the left: } L \end{array} \qquad (1)$$

It is assumed that the blank B is a letter of the alphabet. Thus, the instruction "Print B" is equivalent to "Erase." In addition, at any given instant, the machine is assumed to be in one of a finite number of internal states, given by (2).

$$\begin{array}{l} q_0 \text{ (the starting state)} \\ q_1, \ldots, q_{n-1} \\ q_n \text{ (the final state)} \end{array} \qquad (2)$$

To see how such a machine operates, it is instructive to consider a simple example. Suppose the function f is defined by (3). A Turing machine M adequate for

$$f : x \to x + 1 \qquad (3)$$

computing f can be designed as follows. The alphabet consists of two symbols, given by (4). The number 0

$$\begin{array}{c} | \text{ (stroke)} \\ B \text{ (blank)} \end{array} \qquad (4)$$

is represented by one stroke; the number 1 by two strokes, and so forth.

The computation of $f(2)$ will now be considered. The number 2 is placed on the tape by putting three strokes on consecutive squares. It is assumed that the machine is in state q_0 and that the head is scanning the leftmost square containing a stroke (see **illus.**). The computation is about to begin.

The instructions for the operation of the machine are given in (5). They permit the head to move past

$$\begin{array}{ll} q_0|Rq_0: & \text{in state } q_0\text{, scan `}|\text{', move one} \\ & \text{square to the right, go into} \\ & \text{state } q_0 \\ q_0B|q_1: & \text{in state } q_0\text{, scan `}B\text{', print `}|\text{',} \\ & \text{go to state } q_1 \\ q_1|Rq_2: & \text{in state } q_1\text{, scan `}|\text{', move one} \\ & \text{square to the right, go into} \\ & \text{state } q_2 \text{ (the halting state)} \end{array} \qquad (5)$$

all squares containing a |, until it reaches the first blank square. The head then prints a | and halts. Thus, since there are three strokes on the tape at the beginning, the machine will now have four strokes, indicating that the answer is 3.

The computation for $f(2)$ can be carried out by performing the operations in (6) in the order indicated.

$$\begin{array}{c} q_0|Rq_0 \\ q_0|Rq_0 \\ q_0|Rq_0 \\ q_0B|q_1 \\ q_1|Rq_2 \end{array} \qquad (6)$$

Actually, it is possible to go further. It can be shown not only that a Turing machine computes a function whose domain and range are subsets of the nonnegative integers N but also that the whole structure of the machine can be encoded in N by using a device called Gödel numbering. For example, the odd integers can be assigned to the basic symbols |, B, R, L, q_i as in (7).

Symbol	*Assignment*	
|	3	
B	5	(7)
R	7	
L	9	
q_i	$2i + 11$	$i = 0, 1, \ldots$

Then sequence (8) can be represented by the se-

$$q_0|Rq_0 \qquad (8)$$

quence of nonnegative integers 11, 3, 7, 11. By using prime factorization this representation can be converted to a nonnegative integer $2^{11}3^3 5^7 7^{11}$, which can be denoted by the number r_1. Here, the numbers 2, 3, 5, and 7 are the successive prime numbers. (A positive integer n is prime if $n \neq 1$, and its only divisors are 1 and itself.) Similarly, the second and third instructions can be coded into the numbers r_2 and r_3, respectively. Then the code for the instructions for the machine M is $2^{r_1} \cdot 3^{r_2} \cdot 5^{r_3}$.

In a similar fashion, computation (6) can also be encoded. First, each line of the computation is encoded. Recall that (6) has five lines. If the number associated with the ith line of the computation is s_i, then $2^{s_1} \cdot 3^{s_2} \ldots p_5^{s_5}$ is the number of the computation. (Here p_i is the ith prime in order of magnitude beginning with 2.)

Similar observations can be made in connection with other equivalent definitions of computability. Thus, it is seen that not only is the basic notion of computability discrete but so is the syntax describing this notion.

Two results in the history of computability are worth mentioning. The first states a universal Turing machine exists. A universal Turing machine is a Turing machine that can simulate all other Turing machines. The second, called Church's thesis, was formulated by Church in 1936. It implies that any computable function mapping a subset of N into N can be computed by a Turing machine.

The notion of computability can be, and has been, extended. Thus, for example, there are definitions for a computable real number, a computable complex number, and a computable continuous function of either a real or a complex variable. More generally, the concept of computability on a Banach space is defined. This latter definition actually includes the definition of computability for continuous functions just mentioned. It is these extended definitions that play a role in studying the computability of physical phenomena.

Computability and physical theory. What is the relation between computability and physical theory? This question has been investigated by M. B. Pour-El and J. I. Richards. Two specific examples of this question were given at the beginning of the article, namely the computability or noncomputability of wave propagation and heat dissipation.

Modeling of physical phenomena requires functions of a real or complex variable, rather than functions with domain and range included in N. Hence, the first task is to extend the definition of computability to functions of a real or complex variable. On the basis of this extension, it can be proved that, although heat dissipates computably, waves do not propagate computably. More specifically, wave propagation is governed by the wave equation (9), subject to the initial

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - \frac{\partial^2 u}{\partial t^2} = 0 \qquad (9)$$

conditions of Eqs. (10).

$$\begin{aligned} u(x, y, z, 0) &= f(x, y, z) \\ \frac{\partial u}{\partial t}(x, y, z, 0) &= g(x, y, z) \end{aligned} \qquad (10)$$

It is possible to construct continuous functions of a real variable f and g (with continuous derivatives) so that both of these initial conditions are computable, but the unique solution u of the wave equation, although continuous, is not computable. Physical phenomena can take computable initial data into a noncomputable solution.

There are two ways in which these results have been generalized. The first concerns the equations of classical physics other than the heat and wave equations. The second goes beyond classical physics to theoretical constructs adequate for quantum mechanics. The appropriate model for both is computability on a Banach space. Hilbert space, a special kind of Banach space, plays an important role in quantum theory.

The procedure is as follows. First, the concept of a computability structure on a Banach space is defined axiomatically. It is then shown that the definition is intrinsic. Now, physical phenomena can be associated with linear operators on a Banach space. (For example, on a suitable Banach space, the mapping associated with the wave equation that sends f and g into the solution u is a linear operator.) It is then proved that, under very general conditions, bounded linear operators take computable data into a computable solution and unbounded linear operators do not. The linear operator associated with the heat equation is bounded; the one associated with the wave equation is not. Thus, the results stated earlier concerning heat and waves follow from this general theorem. Of course, this theorem can be, and has been, applied to a host of phenomena of classical physics other than heat dissipation and wave propagation. It has also been applied to the mathematical constructs useful in modeling physical phenomena, such as Fourier series and Fourier transforms.

Quantum theory is concerned with linear operators on Hilbert space, in particular with the so-called self-adjoint operators. Associated with these operators are certain real numbers, the eigenvalues, which are the quantities that are measured. There are two questions dealing with the computability of eigenvalues. Is each eigenvalue a computable real number? (A real number a is computable if there is a Turing machine that can generate a sequence of rationals $\{r_n\}$ so that $|a - r_n| < 1/2^n$.) Is it possible to arrange the eigenvalues in a sequence $\lambda_1, \lambda_2, \lambda_3, \ldots$, so that there is a computer M that, on input n, effectively approximates λ_n? Under very general conditions it can be proved that the answer to the first question is yes, but the answer to the second is no.

AI, the mind-body problem, and physics. The pervasiveness of computers in society—in business, in industry, for research, for recreational games, and in the home—and the speed with which older machines become outdated and are replaced by newer and better ones—suggest a reconsideration of some well-known questions. Will computers become so much better that they are not merely mechanical servants of human beings but intelligent beings themselves? What should the criteria be for judging whether or not a computer is, in fact, intelligent?

The above questions have provided motivation for many specialists in artificial intelligence (AI). There are two subfields of particular interest, robotics and expert systems. Robotics is concerned with the needs of industry for sophisticated devices, that is, machines that can carry out complicated procedures that, at present, require the control or intervention of a human operator. In expert systems, the aim is to code all of the basic knowledge of a specific profession (such as medicine) into a computer package. *SEE INTELLIGENT MACHINES.*

Some of the accomplishments are quite impressive. For example, there are computers that can play chess so well that they have defeated a grandmaster. Nevertheless, there does not appear to be any device that can be said to simulate human intelligence.

In 1950, Turing wrote a philosophical article on computing machinery and intelligence, in which he presented what is now referred to as the Turing test. The test represents an operational approach to the relation between computers and intelligence. Very roughly, if a machine M interacts with a person so that its behavior is indistinguishable from human behavior, it will be said to be intelligent. More specifically, Turing envisioned a game played between two people and a machine. One of the persons is the interrogator I, who is placed in a room apart from the other two. There is no direct contact between I and either of the others, who are known to I only as X and Y. The interaction between I and either X or Y is limited to questions that I asks one of them and that are answered by the one to whom each question is addressed. The questions are asked and answered in an anonymous way, for example, with a computer as an intermediary. The human responds to questions truthfully, but the machine is programmed to lie so as to persuade the interrogator that it is, indeed, a person. The interrogator's task is to determine which is the machine and which is the person. Turing stated his belief that, by about the year 2000, it would be possible to program computers so that there would not be more than a 70% chance for an average person, I, to make the correct identification after 5 min of questioning. Although this goal will probably not be reached by the end of the century, Turing's work is still of great importance.

Turing's test attempts to determine whether mental processes involve only the execution of a sophisticated algorithm. Perhaps there is a nonalgorithmic element. Much has been written about this controversial question. R. Penrose discusses this and similar matters in a recent book. In brief, he thinks that there is a nonalgorithmic element to mental processes and that present-day physics does not suffice to answer the question. Nevertheless, he believes that physical theory of the future may provide some insight.

The study of the relation between computability, mental processes, and physical theory is only at its beginning. Hopefully, future generations will have much more to say about the topic.

For background information SEE ARTIFICIAL INTELLIGENCE; AUTOMATA THEORY; CONDUCTION (HEAT); EXPERT SYSTEMS; LOGIC; NONRELATIVISTIC QUANTUM THEORY; OPERATOR THEORY; ROBOTICS; WAVE EQUATION in the McGraw-Hill Encyclopedia of Science & Technology.

Marian Boykan Pour-El

Bibliography. R. Penrose, *The Emperor's New Mind*, 1989; M. B. Pour-El and J. I. Richards, *Computability in Analysis and Physics*, 1989; J. R. Searle, Minds, brains and programs, *Behav. Brain Sci.*, 3:417–457, 1980; A. M. Turing, Computing machinery and intelligence, *Mind*, 59:433–460, 1950.

Computer graphics

Many disciplines in science, technology, and medicine generate massive multidimensional data sets. The data sets result from laboratory experiments, satellite observations, or direct numerical simulations carried out with high-performance (massively parallel) computers. Techniques to visualize and quantify this increasing amount of information have become crucial to understanding, analysis, and model formulation. In this article, the visiometric (visualization and quantification) approach will be discussed as applied to some fundamental problems in computational fluid dynamics.

Understanding and visualization. The goal of experiments, observations, and numerical simulations in science and technology is to obtain a quantitative and mathematical understanding of the natural and artificial environment. Examples includ numerical simulations of fluid and chemical d of the ozone hole (and its health consequence aircraft and automotive flow patterns (for ir design and performance); and satellite or Dopp image interpretation for weather prediction. amounts of data are generated or acquired and interpreted to obtain a proper understanding design, prediction, and control processes.

Visualization is the process of converting n data into a geometric or graphic representatior ample, one-dimensional waveforms, two-dim contour diagrams, and three-dimensional soli such as a three-dimensional model of the he a series of magnetic resonance imaging (MR The application of computer graphics techn these large data sets is a first step in the vis process. Visualization, as an endeavor in sci engineering, has been growing rapidly since pearance of the National Science Foundatio report *Visualization in Scientific Computing*, w fined and elaborated the activity in the co modern computer technology.

To gain an understanding from the visu it must be possible to recognize areas of i activity. The persistence of a noted effect (p coherent region of high or low intensity suc ozone depletion) under variations and perturb a signal of the presence of something new. T the different disciplines is to study the evolu essential dynamics and interactions of these ef describe them for a finite time period, thus ob partial solution or gaining a new understandi original problem. For example, it is possible the evolution of a storm front, the change in t levels, or the flow of air over a maneuvering

The quantification process includes identifyi sifying, tracking, and projecting evolving an phenomena. Ultimately, it will lead to improv ematical models of the evolution and intera amorphous objects that are observed in time-de data sets.

In this article, some of the procedures ne extract and identify the interacting coherent (regions, objects, and effects) are highlighted, related to ongoing research in computational f namics. The focus is on the evolution and int of coherent scalar, vector, and tensor field st and topologies.

Visiometrics. Visiometrics involves thr cipal steps: identification, quantification, and standing or mathematization.

Identification can be subdivided into visual or rendering an image; feature extraction, wh volves isolating coherent noted regions; and classification, which involves identifying such a as either a known or a possibly new pheno Examples include eddies, blast waves, rings, tu spikes.

Quantification entails measurements to det how the regions have changed, where they moved, and how they interact with other regi