Column: The Physics of Digital Information-Part 2[1]

Fred Cohen

In part 1 of this series (Cohen, 2011a), we discussed some of the basics of building a physics of digital information. Assuming, as we have, that science is about causality and that a scientific theory should require that cause(C) produces effect (E) via mechanism M (written $C \rightarrow^{M} E$), we explore that general theory of digital systems from the perspective of attributing effects (i.e., traces of activities in digital systems) to their causes. Full details of the current version of this physics are available online[2], and in this article, we explore a few more of them.

Previous results questioning consensus around common definitions for the field of digital forensics (Cohen, 2010) have led to additional study suggesting that definitions presented before discussion lead to substantial consensus (Cohen, 2012). Thus each item discussed will start with a loose definition and example.

Definition: A unique history is a single $C \rightarrow^{M} E$ chain that is the only consistent path from the cause to the effect.

For example, suppose we have an imitative copy[3] of an asserted electronic message sent from one party to another. Given that trace and a set of claims about the computers involved, a unique history would demonstrate that there is one and only one party who could have produced the resulting trace, using one and only one process, at one and only one time, from one and only one place.

Current state does not always imply unique history.

More generally there are two important rules that are almost always true for the DFE examiner:

Given initial state and inputs, later outputs and states are known.

Given final state and output, inputs and prior states are not unique.

Digital systems have a finite number of states (settings of the digital values across all of the stored values in the system). The mechanisms that manipulate digital data are commonly called finite state machines or automata (FSM),

1 This editorial piece is extracted and modified from Cohen (2011c).

2 http://infophys.com/

3 Imitative copy := A reproduction of both the form and content of a record. This is what is typically available and called an "exact", "bit image", or "forensically sound", copy in digital forensics. See Cohen (2011b).

often detailed in terms of Moore machines (Moore, 1956) or Mealy machines (Mealy, 1955). In such machines, current state and input lead to the next state and output of the machine in a unique way. That is, given the initial state and sequence of inputs, the final state and sequence of outputs are uniquely determined. Thus time transforms the artifice as it moves forward. But in digital forensics, we generally don't start with causes and try to predict effects. Rather, we start with effects and seek to identify their causes. In modern computers it is almost never possible to "run time backwards" given a set of traces, and identify a unique history that led to the traces found.

Definition: Convergence asserts that, as a mechanism transforms inputs and internal states into outputs and subsequent internal states over time, different inputs produce identical outputs. Divergence asserts the opposite, that for the same input, different outputs are detectable.

For example, if we test rolling a rock down a hill repeatedly and, no matter how tightly we control the process, there are slightly different outcomes each time, this would be divergence. But if we ran an FSM forward again and again with different inputs and initial states each time, and got identical outputs and final states, this would be convergence.

Digital space converges while physical space diverges with time.

The digital artifice over time is, in general, a many-to-one transform. Furthermore, inverting time in an FSM produces potentially enormous class sets of possible prior states and inputs, and determining them precisely is too complex to be done for nontrivial systems (Backes, Kopf, & Rybalchenko, 2009). This is at odds with the current model of the natural world, in that physical space is generally believed to have an essentially infinite number of possible states and to increase in entropy over time so that order is always reduced. No matter how tightly we control a physical experiment, there will always be a level of granularity at which outputs are differentiable. The difference between the digital and physical spaces is greatly influenced by the fact that the digital space has only finite granularity in time and space, as was discussed in the first article in this series.

Definition: Equivalent machines are, possibly different machines that, from an external perspective, behave identically with respect to a defined set of external data.

For example, different compilers may transform the same program into different binary executable codes that work slightly differently even though they produce the same outputs from the same inputs.

Many FSMs are equivalent.

An unlimited number of different FSMs may produce the same output

sequence from the same or different input sequences. For example, at the level of computer programs in common use, an editor, digital recorder, or user program, may produce the same outputs from different inputs. With incomplete traces, we cannot uniquely determine prior states and inputs. To the extent that traces are more or less complete, we may or may not be able to uniquely determine or bound the set of programs that might have produced the traces. We may not even be able to determine the extent of completeness of traces we have.

Definition: A lossy transform is a mapping from input to output that cannot be reversed to produce a unique input. That is, it is a many to one transform.

For example, JPEG files are often compressed using the JPEG lossy compression algorithm (Hamilton, 1992). The results trade off space for quality.

Hash functions and digital signatures as lossy and thus not unique.

Any transform that produces output space of a predefined size for an input space of a larger size is lossy and thus not unique. As an example, an MD5 or SHA hash of a file does not uniquely identify that file. There are in fact an unlimited number of other files that would produce that same hash value. Being careful, note that this is not an infinite number of files – only an unlimited number of them. To see this, suppose we generate file after file of length one bit more than the length of the hash. Since the length is one bit more, there are twice as many files of that length than there are hash values. If we create one after another of these files, eventually we will exhaust all of the possible values for the hash function, and as soon as we get to one more unique input file than the number of possible hash value, we are guaranteed that two different input files will have identical hash values. This does not make such hashes useless in digital forensics, but it does mean that they do not uniquely identify an input or certify that a produced file is unaltered from its initial creation.

A summary of properties

There are many other properties of digital systems and the physics of digital information. A summary extracted from the book chapter identified above is included here to expand thinking about these issues.

Digital World	Physical World
Finite time granularity (the clock)	Infinite time granularity
Finite space granularity (the bit)	Infinite space granularity

Digital World	Physical World
Observation without alteration	No observation w/out alteration
Exact copies, original intact	No exact copy, original changed
Theft without direct loss	Theft produces direct loss
Finite (fast) rate of movement	No locality (entanglement)
An artifice created by people	A reality regardless of people
Finite State Machines (FSMs)	Physics and field equations
Homing sequences may exist	No perfect repeatability
Forward time perfect prediction	Forward time non-unique
Backward time non-unique	Backward time unique
Digital space converges in time	Physical space diverges in time
The results are always bits	The results are always continua
Results are always "Exact"	Results never perfectly known
Time is a partial ordering	Time is real(location)
Errors accumulate	Errors are local
Representation limits accuracy	Reality is what it is
Precision may exceed accuracy	Precision is potentially infinite
Forgery can be perfect	Forgery cannot be perfect
DFE is almost always latent	Some evidence is latent
DFE is trace but not transfer	Traces comes from transfers
DFE is circumstantial	Evidence is circumstantial
DFE is hearsay	Evidence is physical

Digital World	Physical World
DFE cannot place a person at a place at a time	Evidence may put an individual at a place at a time
DFE can show consistency or inconsistency only	Evidence can show more than just consistency
Probability is dubious	Probability is often usable
Content has information density	No defined density limits
Content density variable	Content density not controlled
Content perfectly compressible	No perfect compression
Digital signatures, fingerprints, etc. generated from content	Body (phenome) generated from DNA (genome)
Content meaning is dictated by context	No universal theory of meaning but physicality exists regardless
Context tends to be global and dramatically changes meaning	Context tends to be local and incrementally changes meaning
FSMs come to a conclusion	Eats shoots and leaves
Cognitive limits from program	Cognitive limits from physiology
Hardware fault models from computer engineering	Hardware fault models from physics
Time and space tradeoffs known	Tradeoffs unclear
Near perfect virtualization and simulation possible	No virtualization
Many nearly or equivalent FSMs	The uncertainty principal
Undecidable problems	Nothing known as "unthinkable"
Computational complexity limits computations	No well understood limits on new ideas

Digital World	Physical World
Everything is decidable	Many things are not decidable
Consistency is guaranteed	Consistency is possible
Completeness is guaranteed	Completeness is possible
Consistency AND completeness	Consistency OR completeness
Time limits on achievable results	Time limits unknown
Complexity-based designs	Complexity not determinant
Fault tolerance by design	Normally not fault tolerant
Accidental assumption violations	Assumptions non-violable
Intentional assumption violations	Assumptions non-violable
Discontinuous space	Continuous space
Discontinuous time	Continuous time
Minor differences amplified near discontinuities	Differences retain fidelity
Major differences suppressed away from discontinuities	Differences retain fidelity
Identical use of an interface may produce different results	No such thing as identical, each thing is unique
Ordering may be reversed	Ordering subject to light time
Value sorts may be reversed	Value sorts remain consistent
Actuate-sensors loop errors	Interference based errors
Sensors/actuators limited in physical properties	All physical properties present

Table 1 – Summary of Information Physics

A final comment

There is a lot to learn about the physics of digital information, and from the perspective of digital forensics, this is the sort of knowledge that is increasingly necessary to understanding what you are doing when you undertake to testify about such matters.

I urge you to review the details of the physics in its full richness and with its current limitations, and to draw your own conclusions. Read the chapter cited above, comment on it, prove it is wrong if and where it is, show its limits, and move the field forward.

And I urge you to challenge yourself and others to up your game. In case after case, we encounter self-identified experts who don't understand the basics of how things work and end up testifying with inadequate basis. In many cases their conclusions may be right, but their presentation and the facts they provide may not support them. In other cases, their conclusions are not right at all. At the heart of it all is the lack of attention to the basics of the science that underlies digital forensics. This is a problem we hope to continue to address in this series and this publication.

References

Backes, M., Kopf, B., & Rybalchenko, A. (2009). Automatic Discovery of Quantification of Information Leaks. In *Proceedings of the 30th IEEE Symposium on Security and Privacy*, May 17-20, 2009, Berkeley, CA. Los Alamitos, CA: IEEE.

Cohen, F. (2010). Toward a Science of Digital Forensic Evidence Examination. IFIP TC11.8 International Conference on Digital Forensics, Hong Kong, China, January 4, 2010. In K.-P. Chow and S. Shenoi (eds.), *Advances in Digital Forensics VI*, Springer, pp. 17-36.

Cohen, F. (2011a). *Digital Forensic Evidence Examination*. Livermore, CA: ASP Press.

Cohen, F. (2011b). Putting the Science in Digital Forensics. *Journal of Digital Forensics, Security and Law*, *6*(1), 7-14.

Cohen, F. (2011c). The Physics of Digital Information. *Journal of Digital Forensics, Security and Law*, *6*(3), 11-16.

Cohen, F. (2012). Update on the State of the Science of Digital Evidence Examination. Submission to the 2012 Conference of Digital Forensics, Security and Law.

Hamilton, E. (1992, September 1). *JPEG File Interchange Format, Version 1.02*. Milpitas, CA: C-Cube Microsystems. Retrieved from http://www.jpeg.org/public/jfif.pdf

Mealy, G. (1955). A Method for Synthesizing Sequential Circuits. *Bell Systems Technical Journal*, *34*, 1045–1079.

Moore, E.F. (1956). Gedanken experiments on sequential machines. In *Automata Studies*. Princeton, N. J.: Princeton University Press, pp. 129-153.