Volume 6, Fall Issue, 1992

THE RESPONSIBILITY OF INTELLIGENT ARTIFACTS: TOWARD AN AUTOMATION JURISPRUDENCE

Leon E. Wein[*]

INTRODUCTION

We were warned about the dangers of automation. Over a hundred years ago, Dr. Frankenstein created and activated his monster, a conglomeration of human flesh and wire electrodes, the brainchild of a maniacal scientist. Can humanity, which considers itself the master, control its progeny? This terrifying embodiment of man's scientific genius grown too powerful was ultimately destroyed, but the symbol lives on. A feeling that somehow man has lost control of his creation pervades the modern psyche as machines seem to rule our lives. As technology spawns ever more competent and autonomous machines with which we must interact on a daily basis, humankind is being relegated to the status of a bystander, tremulously watching as our mechanized progeny come to dominate contemporary life. Must we subordinate ourselves to this twentieth-century Frankenstein? Are we still able to counteract its monstrous force? Within a whirlwind of technological advancement, one actor continues diligent in an effort to constrain the enormous power of automation that we have unleashed on ourselves. The law, which plays a dominant role in providing a framework within which human beings interact, also has responsibility for keeping technology within the bounds of human governance and control. What follows is an exploration of the role of our legal system as arbiter; the law is the instrument with which we seek to tame this "monster" and keep technology under control.

Every year, every day, perhaps every minute, someone is experiencing problems with a machine. A student loses quarters in the washing machine; a family of four is billed by computer for $7000 worth of

---

* Professor of Law, Brooklyn Law School. A.B., 1961, Brooklyn College; LL.B., 1966, New York University; Diploma in Law, 1966, Cambridge University. The author gratefully acknowledges a research stipend provided by Dean David G. Trager and Brooklyn Law School in support of this research. He is especially grateful to former research assistants, particularly Kimball Kenway and Madelaine Berg, as well as to Professor William Bechtel of Georgia State University's Department of Philosophy, Professor Michael Libonati of Temple University School of Law, Professor Lawrence Crocker of New York University School of Law, and Rifka Wein, each of whom read and commented upon different sections of the manuscript. This work was greatly improved by their suggestions.

electricity; a businessman deposits $1000 in an electronic night depository on a Sunday evening, and the following day his check bounces because the deposit was never recorded; a man inserts his bank card into a cash machine, but instead of giving him twenty dollars, the machine swallows his card. These people are only a few of thousands of victims of automation technology.[1]

These problems apparently are widespread. Cash machines are frustrating people frequently and randomly. Most often, the customer's recollection and testimony are no match for computerized banking records. And while each fifty-dollar error may not seem to be a problem of monumental proportion, the fact that so many people are victimized by banking automatons is worthy of notice. Indeed, the entanglements resulting from cash machine errors became very clear to one Mr. Stagg who deposited $608 with an automatic teller on the fourth of January, but failed to observe that his deposit slip had incorrectly dated the deposit receipt as the fourth of March. In early April, upon receiving an overdraft notice, Mr. Stagg informed the bank of its mistake, and ultimately demanded that the bank properly credit his account. One week later, the bank filed a complaint against him for felony theft, which resulted in his arrest and imprisonment for two days. When the incorrect deposit slip was discovered in May, Mr. Stagg refused to release the bank from civil liability for his arrest. He settled his case with the bank, the president of which contributed $19,000, and the vice-president an additional $1000, to Mr. Stagg's $50,000 recovery.[2]

Who has not experienced the frustration and helplessness of falling victim to electronic apparatus? Given the opportunity, most of us would gladly assault the offending device and squelch its electronic hum forever. Yet we refrain from doing so, for none of the problems caused by the mechanical foul-up would be solved by indulging our appetite for

---

1. *See, e.g.*, Judd v. Citibank, 435 N.Y.S.2d 210 (Queens County Civ. Ct. 1980). Dorothy Judd sued Citibank for $800, which she claimed was improperly deducted from her account. Citibank's computer records indicated that Ms. Judd had made two withdrawals from a Citibank cash machine between 2:00 and 2:30 P.M. on two separate afternoons. Ms. Judd, however, had been in her office on these afternoons, and obtained a letter from her employer attesting to that fact. At a trial in small claims court, Ms. Judd produced the letter from her employer, and she herself testified that she had neither revealed her personal identification number to anyone, nor ever let her access card leave her possession. The judge, apparently choosing to believe the human rather than the computer, awarded her the $800.

2. Stagg v. Bank of Breakenridge, 22 ATLA L. Rep. 269 (Colo. 1979); *see also* Young v. Bank of Am., 141 Cal. App. 3d 108 (1983) (billing error resulted in such severe stress, nervousness, and headaches that plaintiff entitled to receive $150,000 in damages).

vengeance. Ultimately, the human responsible for the machine's misbehavior is sought out, and when talking will not rectify matters, we are off to court. The student, no doubt, will finally recover the quarters from the washing machine company, the family of four will eventually persuade the utility company to answer for its miscalculation, and the bank may even one day locate the businessman's thousand-dollar deposit. Problems that arise from dealing with machines become ever more complex as technology advances and as we are made to depend increasingly on unattended intelligent machines that displace humans in everyday transactions.

The twenty-first century will confront the law with a world in which interaction with computers is the rule rather than the exception. The last decades of this century are witnessing unprecedented growth in the utilization of highly sophisticated, unattended artifacts in virtually every activity of our daily life, and replacement of man by machine is accelerating as technological expertise is increasingly devoted to this task. Today, it is part of our common experience to interact with machines that have become co-workers, bankers, teachers, and even physicians. As the promise of the computer age becomes a reality, the law, which sets standards of human conduct, will be required to legislate for machines as well. We should not be surprised to discover that judges, perhaps more than engineers, will dictate design specifications of emerging automation technologies.[3]

I. THE LEGAL PERSONALITY OF MACHINES

Determinations arrived at by computerized machinery are as prone to miscalculation and blunder as any human decision. Consequently, we expect the law to apportion responsibility for mechanical indiscretion among the owner of the computer, its designer, its programmer, and other possible defendants. But is a human master indispensable to assigning legal liability to an automated system?[4] Although there is

---

3. Rudimentary vending machines and similar present-day interactive machines lack the capacity to record images of such things as the product sold or the customer with whom a machine interfaced, although, from the standpoint of existing technology, it is not difficult to construct an interactive machine with this ability. Human bank employees are expected to be able to give an account of transactions in which they engage ("recountability"). *See, e.g.*, McEvans v. Citibank, 408 N.Y.S.2d 870 (N.Y. County Civ. Ct. 1978)(obligatory that a capability for "recountability" be retrofitted into an existing "automated teller system").

4. The conventional analytical view is that responsibility must be charged to the machine's

always someone who designs, manufactures, or to some degree controls every automaton sent out into society to replace discarded human beings, egocentric machines are coming to enjoy a new independence and individuality. Will we come to view autonomous machines as mere implements of their owners or instead as the lifeless and dependent legal agents of a human principal?

To think of a machine as suited to having unique legal duties is to suggest that unattended intelligent equipment is evolving to a point where some such apparatus has responsibilities. It is in this sense that we speak of such instrumentalities as "legal agents," a theme to which much of this analysis is devoted. While agency status is unquestionably lower than the autarchic status given to "principals" or "masters," the ascription of legal agency to a machine posits a significant degree of independence from the human creator.

Consideration of those situations in which we find machines standing in for humans is central to this analysis. Section II surveys and analyzes cases arising in circumstances in which transactions or tasks might comparably be performed by either an unattended intelligent system or a human being. This survey will demonstrate situations in which the machine performance is at greater jeopardy of legal liability than is a functionally comparable human performance, and situations that produce the contrary result. From the disparity in liability between cases involving mechanical transactions and those involving human behavior, we infer that liability generated by a human activity is not always

---

human proprietor, who serves as a lightning rod for liability. However, Yorick Wilks predicts that advances in artificial intelligence technology will make it ever more difficult to locate the responsible human actor:

> The difficulty can be avoided by always identifying humans, standing behind the machines and programs as it were, to carry the blame, in the sense in which there are always real humans standing behind agents and behind companies, which also have the legal status of non-human responsible entities ("anonymous persons" in much European law). . . . In most situations now imaginable, it will not be too hard to identify individuals, if there is a need to do so, behind programs and machines. However, things may become more tricky as time goes on, and the simple substitution of responsible people for errant machines harder to achieve.

Yorick Wilks, *Responsible Computers!*, PROC. NINTH INT'L JOINT CONF. ON ARTIFICIAL INTELLIGENCE 1279 (1985). Conversely, the position of this work is that there exist contexts in which legal responsibility may properly be attributed to some computerized systems themselves, systems which for lack of a better term may be regarded as "unattended."

reproduced by automated substitutes.

This Article argues that unattended intelligent artifacts should be subject to liability, independent of human masters, on consideration of universal concepts of accountability underlying legal systems throughout history. In analyzing the possibility of conferring independent legal agency status on the automaton, contemporary automatic defibrillator instrumentation (designed to monitor the medical condition of and minister to heart attack victims) is investigated. These emerging mechanical medics constitute a substratum on which the attribute of "duty" may properly be predicated. By attaching liability to risky behavior, the coercive force of the law influences the very texture of society. This Article therefore suggest that stand-in systems that generate such disparate liability may be evolving into emancipated entities responsible in some sense for their own decisions or actions.

Ascribing a "consummate" legal personality to a mechanized system, as is done in the case of a corporation, subsumes automatons, which are not only "duty-bound" or "responsible," but also enjoy rights and privileges. This merely means that attributing legal personality to an entity is predicated on giving "rights" and "duties" to it. In ascribing legal duties to an unattended intelligent device, it is sufficient that interactions with it entail idiosyncratic legal consequences and outcomes that deviate from those arising from analogous transactions accomplished without automated intermediaries. Because automated devices generate liability of a different order or degree than humans performing an equivalent task, they may more appropriately be viewed as "incomplete" legal persons. To demonstrate the existence of duty-bearing, unattended intelligent systems, examples in which a machine's performance is at greater or lesser jeopardy of legal liability than a functionally comparable human performance are suggested. Entities that engender idiosyncratic legal outcomes constitute "legal agents," notwithstanding their lack of rights, because they are to some degree held independently accountable for their actions.

One might object that the mere fact that automated transactions produce idiosyncratic legal outcomes is unexceptionable, since almost any entity could produce a different outcome. I would take a lead from John Dewey and respond as he did when calling to our attention that the term "legal person" in the context of a corporation denotes nothing more than a "right and duty bearing unit":

> From this point of view, the right-and-duty bearing unit, or subject, signifies whatever has consequences of a specified kind. The reason that molecules and trees are not juridical "subjects" is then clear; they do not display the specified consequences. The definition of a legal subject is thus a legitimate, and quite conceivably a practically important matter. But it is a matter of analysis of facts, not of search for inhering essence. The facts in question are whatever specific conse-quences flow from being right-and-duty bearing units . . . . The consequences must be social in character, and they must be *such* social consequences as are controlled and modified by being the bearer of rights and obligations, privileges and immunities. Molecules and trees certainly have social consequences; but these consequences are what they are irrespective of having rights and duties. Molecules and trees would continue to behave exactly as they do whether or not rights and duties were ascribed to them; their consequences would be what they are anyway.[5]

The conception that I superimpose on Dewey's, that there exists a range of possible legal statuses along which various types of automated instrumentation might be categorized, invites inquiry into whether computerized entities may eventually rise above the rank of "legal agent" and, by analogy to corporations,[6] become beings possessed of legal rights as well as duties. It seems clear enough that the safeguards of the Constitution or of civil rights laws could become applicable to automatons only at the upper reaches of the hierarchy of legal statuses. That legal systems have ascribed legal "rights" to artificial persons such as

---

5. John Dewey, *The Historic Background of Corporate Legal Personality*, 35 YALE L.J. 655, 655-61 (1926).

6. John Dewey described the logical process by which corporations came to be possessed of "rights" as follows:

> The contrast of "natural" and "artificial" persons got its point from the fact that "natural" connoted possession of inherent and inviolable rights. The dialectic of the courts, under the pressure of social facts, was equal to declaring that corporations, while artificial and fictitious, nevertheless had all the natural rights of an individual person, since after all they were legal persons.

Id. at 669.

corporations without much difficulty is largely due to a perception that corporate organizations are but composites containing "human" constituents or ingredients. Beings such as slaves, who historically have come to be treated as complete persons in the sense that they possess innate "rights" as well as "duties," and other organisms that remain problematic in terms of legal personality (such as fetuses,[7] the dead, and the permanently unconscious), all share a common human essentiality.[8] In contrast, unattended intelligent machinery and computer-based systems, to which this Article attempts to attribute legal duties, have no human components.

The law is not likely to elevate artifacts to the status of beings with rights until they become very "person-like." Dewey puts forth a theory of legal agnosticism, which is the notion that denominating something a legal person means nothing more than saying that it is an entity to which rights and duties are ascribed. In seeking to apprehend a context in which intelligent artifacts might not only be held legally accountable, but also morally responsible, and in appraising advances in artificial intelligence technology from the standpoint of the possibility of someday constructing an artifact that would both possess rights and bear duties, we will in Section III discard the philosophical convenience of Dewey's legal agnosticism[9] and consider the sorts of simulated anthropomorphic attributes machines might come to exhibit. For the time being, as we limit our perspective to the question of whether machines provide an infrastructure on which "duties" are properly superimposed, we may temporarily adopt John Dewey's position of legal agnosticism.

From an historical perspective, the conceit of viewing an artifact as an

---

7. In Roe v. Wade, 410 U.S. 113, 158 (1973), the Supreme Court stated that the word "person" as used in the Fourteenth Amendment does not include the unborn. That case, which occasioned a continuing political melee over issues involving legal and moral personhood, speaks in terms of "potential life" and "viability," which the Supreme Court defined as commencing when the fetus had developed to a point where it has "the capability of meaningful life outside the mother's womb." *Id.* at 163.

8. *See* Marshal S. Willick, *Artificial Intelligence: Some Legal Approaches and Implications*, 4 AI MAG. 5 (1983). Willick considers these various categories of the human condition and the extent to which individuals in such states are accorded the status of a "legal person." He also seeks to draw analogies pertinent to the question of whether legal personality might ultimately come to be extended to intelligent machines.

9. *See* Dewey, *supra* note 5, at 655-60. For purposes of legal research, the idea of corporate legal personality merely "signifies what the law makes it signify." A corporation is a "person" because it constitutes a "right-and-duty bearing unit. . . . [W]hat 'person' signifies in popular speech, or in psychology, or in philosophy or in morals" is, for Dewey, inconsequential. *Id.* at 655.

agent of a human principal is rooted not in the nexus of people and inanimate objects, but in relationships among human beings. The industrial revolution and resultant mechanization of industry brought within humankind's grasp a capacity to multiply and disperse material wealth with continually diminishing human effort. The wealth of prior civilizations, on the other hand, was achieved by dividing the community into an elite or leisured class that arrogated to itself most of the material wealth produced by inferior laboring classes and wretched slaves.

Bondsmen, compelled to perform the most degrading labor, had onerous duties and negligible rights. Because a slave was merely chattel in the hands of his owner, his well-being, family, and even his life depended on the whim of his master. In early Roman law, a slave who caused damage or injuries to another was surrendered to the aggrieved party, who was privileged do with him what he or she pleased. Gradually, the master was permitted to avoid forfeiture by a payment of monetary compensation. The owner's responsibility for a slave's conduct is grounded in logic. Once the major premise, that a slave is but the property or tool of the master, is accepted, the middle term may be distributed with ease. The proprietor is responsible for damage caused by a slave in the same sense that we hold persons accountable for damage caused by their animate or inanimate property.

As serfdom and slavery became extinct, the doctrine of respondeat superior developed. Equality in legal status of employers and their employees made nonsense of the notion that a master should be held liable for the acts of a servant on grounds similar to those which held him responsible for losses caused by other categories of property. The doctrine is not grounded on a logical interconnection binding the wrongdoer to a loss he has brought about, but instead on a policy of providing compensation for loss, rather than imposing liability on financially incompetent parties. Consequently, employers are answerable for their employees' autonomous acts even though they neither immediately influenced nor participated in the wrongful behavior that occasioned the loss.

The respondeat superior doctrine treats an employee acting within the scope of his employment as a legal agent rather than as a mere instrument of the employer. The obligation to make restitution falls on the master as the financially responsible individual, rather than on the servant as the morally responsible party. If an employee injures a pedestrian with the company truck while en route to a delivery, he is viewed as the employ-

er's agent in contrast to the truck, which is merely a tool or instrument. While both employer and employee are said to be liable for the pedestrian's injuries, the truck is not liable. If one person hits another with a hammer, we hold the former liable for the latter's injury, and we consider it pointless to regard the hammer as liable from either a prudential or a moral point of view. The hammer is a tool under the control and mastery of a human tortfeasor. Recharacterizing the transaction by saying that the hammer "flew" out of one person's hand and "landed" on another is not plausible, for we do not consider a hammer capable of acting of its own accord.

Today, technological changes made possible by automation, computer technology, and artificial intelligence are creating a watershed in respondeat superior doctrine by bringing the conception of slavery back on the scene.[10] As employees who replaced slaves are themselves replaced by mechanical "slaves,"[11] the "employer" of a computerized system may once again be held liable for injury caused by his property in the same way that she would have been if the damage had been caused by a human slave. In holding the employer responsible for damage occasioned by an automated system, we need no longer resort to a policy of placing responsibility on the financially responsible party, as the logic of holding a mechanizer accountable for damage caused by his automaton remains as unshakable as in the days of slavery. Humans once again own chattels that perform thankless tasks for them, and automation technology is but slavery flourishing anew, without moral qualm.[12]

A. Liability of a Machine

Currently the law does not allow us to sue a machine, although it seems that some machines are beyond our control. We assign liability

---

10. *See* OLIVER W. HOLMES, JR., THE COMMON LAW 231 (1881) (describing the doctrine of respondeat superior as "the vanishing point of the servile status").

11. As Norbert Wiener has pointed out, "the automatic machine, whatever we may think of any feelings it may have or may not have, is the precise economic equivalent of slave labor." NORBERT WIENER, THE HUMAN USE OF HUMAN BEINGS 189 (1950).

12. Such an attitude is exemplified by the ritual Japanese ceremony of setting on fire worn out tools and utensils to give thanks for assiduous service. At a ceremony in which used transistors and burned-out micromotors were put to the flame, a participant remarked: "[T]o get this kind of quality we had to torture many of these parts. We boiled them and froze them, and ran them day and night. We engineers feel for our devices, and we want to give thanks." David E. Sanger, *For a Job Well Done, Japanese Enshrine the Chip*, N.Y. TIMES, Dec. 11, 1990, at A4.

exclusively to humans, not to quadrupedal or inanimate entities. If someone is run over by an automobile, we do not seek to destroy the car, but instead we penalize some human we hold responsible for the mischief. In 1981 the world press sensationalized the death of a Japanese worker crushed in a robotics accident, making it seem only a step removed from a full-scale robot uprising.[13] In actuality the victim had entered a prohibited area of the factory, and his misfortune was functionally equivalent to his having inadvertently stepped in front of a speeding automobile.[14] Still it was a first: a robot had killed a human, and the attention given to the incident reflects a persistent fear that we will ultimately lose control of our mechanized offspring. Mary Shelley's *Frankenstein*[15] is symptomatic of this unease about mankind's ostensible control of science and technology, an anxiety that things—animate things—might get out of hand. Symbols such as Frankenstein, who aspired to divine powers before his creation got away from him, or the sorcerer's apprentice, who was unable to control the wonderful magic of his master, have a way of reappearing in different forms.

The extent to which intelligent artifacts are evolving into legal agents rather than mere tools or instruments warrants attention. Now that the age of interactive computerized systems is upon us, machines are coming to possess the kind of autonomy which our ancestors assigned to them out of sheer superstition. What attitude ought the law to take toward dealings with anthropomorphic automated devices that interact with us,[16] and which to some extent function independently of those who own them? People interface with computer-based systems with ever greater frequency, and these interactions sometimes engender idiosyncratic legal

---

13. *See, e.g.*, Henry S. Stokes, *Japan's Love Affair with the Robot*, N.Y. TIMES, Jan. 10, 1982, Sec. 6 (Magazine) at 24, 62.

14. This misadventure led an observer of the artificial intelligence scene to suggest:

> It has been proven that computer-directed robots could accidentally cause the death of a human being. If such a death was deliberately caused by a self-programming AIequipped computer, the usual questions surrounding the "insanity defense" would appear applicable whether the machine itself or its original programmer was accused.

Willick, *supra* note 8, at 11.

15. MARY WOLLSTONECRAFT SHELLEY, FRANKENSTEIN (1869).

16. We are told that one of the great Talmudic law scholars of the preceding generation, Rabbi Aharon Kotler, made a point of utilizing manned rather than automatic toll booths. This example illustrates the attentiveness he brought to the ethical details of everyday life. "It's not *kovod habriyos* (respectful of humanity)," he would say, "to pass up a man for a machine." WILLIAM B. HELMREICH, THE WORLD OF THE YESHIVA 44-45 (1982).

outcomes that vary from those which would have arisen in an analogous transaction accomplished in the absence of the automated intermediary. The fact that mechanical transactions may generate a distinct species of liability indicates that the relationship between man and his machine may become so attenuated that artifacts assume a genuinely autonomous legal identity, accountable for their own misconduct.

A machine that functions independently of the human who set it in motion and that attains the status of "legal agent" is a good deal more than a mere article of commerce in which somebody has constitutionally protected property rights. In a case where a hospital is held liable because a procedure is undertaken by a nurse rather than a physician, the liability of the hospital is grounded on respondeat superior doctrine rather than on fault. Cases in which decisionmaking is delegated to intelligent machines are likely to result in an imposition of vicarious liability in situations where we would not consider it fitting to assign moral culpability.[17] Where a principal makes use of a human agent who blunders, there is always the possibility that responsibility will be assigned to the human agent. And if that human agent is an independent contractor untied to the employer, the independent contractor's accountability shields the principal from liability.[18] Where a human principal delegates decisionmaking to intelligent artifacts, however, our focus remains on the human as the locus of liability, and we are less inclined to attribute the mischief to the machine insofar as such a posture would insulate the human principal from liability.

B. Machines as Bearers of Duties

Legally responsible machines may seem extraordinary to us because we are accustomed to holding liable the *people* who implement the tortious machines. This fact demonstrates not only the extent to which modern tort law has become saturated with respondeat superior doctrine,[19] but

---

17. "While in the legal context of strict liability a causal role in the process may be sufficient for attributing responsibility, in most moral contexts it is not. We require in addition that it make sense to assign praise or blame for an action." William Bechtel, *Attributing Responsibility to Computer Systems*, 16 METAPHILOSOPHY 296, 298 (1985).

18. As a general rule, one who employs an "independent contractor" is not subject to "respondeat superior," a type of vicarious liability doctrine which imposes upon a master or other principal responsibility for the conduct of a servant or other agent even though the principal did not intend or direct such conduct. *See generally* RESTATEMENT (SECOND) OF AGENCY § 2 (1957).

19. In a classic criticism of the doctrine of respondeat superior, Thomas Baty argued:

also how the independent contractor exception has degenerated over time.[20]

We need to be able to sue machines directly because sometimes the (human) decision to program or design a certain way may be non-negligent, while the (computer) decision to behave a certain way in an individual situation *would* be considered negligent, if the computer could be sued. That decisionmaking processes programmed on a computer sometimes outperform competent human experts was recently demonstrated by a computer neural network with the ability to diagnose heart attacks more accurately than emergency room physicians.[21] These computerized systems may make correct decisions where a human specialist would not, and vice versa. For example, it is plausible that a person employing fallible human standards would make disastrous decisions in the circumstances of A and B, but manage situations of C, D, and E properly. A stand-in unattended system, on the other hand, might make more acceptable decisions in the circumstances of A and B, but a calamitous decision in situation E. In these circumstances, someone imaginably will become a victim who would not have been one had the decisionmaking

---

> One may venture, not improperly, to characterize the modern doctrine of vicarious liability as a veritable upas-tree. Unknown to the classical jurisprudence of Rome, unfamiliar to the mediaeval jurisprudence of England, it has attained its luxuriant growth through carelessness and false analogy, and it cannot but operate to check enterprise and to penalize commerce. . . . It certainly is not true that what your agents do you do yourself. Neither in law nor in morals are the unauthorized acts of employees attributable to the employer.

THOMAS BATY, VICARIOUS LIABILITY 7 (1916).

20. There are many exceptions to the independent contractor rule, such that one who hires a competent contractor and does no negligent act himself may yet be held responsible for injuries occurring through the contractor's negligence. Among others, these now include situations where the thing contracted to be done is unlawful, where the acts performed create a public nuisance, where a duty has been imposed by statute or ordinance, where the principal is deemed to have assumed a duty by contract or by accepting a municipal permit imposing the duty as a condition thereto, where the services are accepted in the reasonable belief that they are being rendered by the hirer, and where the hirer is under some other duty which by decisional law has been denominated non-delegable. *See generally* RESTATEMENT (SECOND) OF TORTS §§ 416-29 (1965); WILLIAM L. PROSSER, TORTS § 71 (4th ed. 1974). Again, where the work performed is considered "inherently dangerous" or "intrinsically dangerous," a principal will not escape liability for the negligence of the independent contractor. RESTATEMENT, *supra*, §§ 416, 426, 427(a)-(b); PROSSER, *supra*, § 71.

21. For an evaluation of the proficiency of a neural network designed to diagnose acute myocardial infarction in emergency room patients, see William G. Baxt, M.D., *Use of an Artificial Neural Network for the Diagnosis of Myocardial Infarction*, 115 ANNALS INTERNAL MED. 843 (1991) (physicians found to have a diagnostic sensitivity of 77.7% and diagnostic specificity of 84.7%, while the artificial neural network had a sensitivity of 97.2% and specificity of 96.2%).

process been limited to humans.

Risk-utility analysis[22] of such disastrous outcomes suggests that superseding men by machines will not result in liability where the computer based system on the whole performs better than supplanted human procedures, even where someone becomes a casualty who would not have been one otherwise. This issue is similar to determining where to locate a dangerous gas storage facility. Once it has been decided that the hazardous facility must be placed somewhere, it is presumably best located in a place where the least number of people will be exposed to the perils it generates. If it has been located in a place where the least number of people are endangered, a decisionmaker, mindful that some few people will be injured in the event of misadventure, has nevertheless acted responsibly in the circumstances.

From the standpoint of vicarious liability, there is no reason why intelligent machines might not in appropriate circumstances be considered "independent contractors," even in cases where such a characterization insulates the human principal from liability. Intelligent machines employed to displace human activities will on occasion bring about idiosyncratic legal outcomes which vary from those which would have

---

22. Resolution of such conundrums requires use of Judge Learned Hand's risk utility analysis, which restates the negligence standard in terms of a risk-utility algorithm with three variables: the probability that the injury would result from an actor's conduct, the gravity of harm that might be expected to result should injury occur, and the burden of adequate precautions to avoid or minimize injury. Judge Learned Hand reduced his test to algebraic terms: "[I]f the probability be called P; the injury, L [loss]; and the burden, B [i.e., the burden of precaution to avoid the loss]; liability depends upon whether B less than L multiplied by P: i.e., whether B less than PL." United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947). Judge Hand himself recognized that the risk-utility formula he enunciated as an algebraic representation of liability did little more than sheath with a veneer of precision the commonplace assumption that reasonable people evaluate whether the advantages of a proposed course of action outweigh its risks. Judge Hand restated his algorithm as follows:

> It is indeed possible to state an equation for negligence in the form, C = P x D, in which the C is the care required to avoid risk, D, the possible injuries, and P, the probability that the injuries will occur if the requisite care is not taken. But of these factors care is the only one ever susceptible of quantitative estimate, and often that is not. The injuries are always a variable within limits, which do not admit of even approximate ascertainment; and, although probability might theoretically be estimated, if any statistics were available, they never are; and, besides, probability varies with the severity of the injuries. It follows that all such attempts are illusory, and, if serviceable at all, are only so to center attention upon which one of the factors may be determinative in any given situation.

Moisan v. Loftus, 178 F.2d 148, 149 (2d Cir. 1949).

resulted had the transaction been accomplished by a human. Thus, the ramifications of attributing legal agency to machines goes far beyond the question whether people may be held liable for relying on machines rather than on humans. In Section III, we consider computerized devices that recognize potentially fatal heart rhythms and automatically deliver electric jolts to the heart to restore normal heartbeat. This example will demonstrate why there is no reason that an intelligent machine might not be considered an independent contractor, even in circumstances where such a characterization would insulate the human principal from liability.[23]

Nothing about the current state of society suggests that intelligent artifacts will soon come to enjoy constitutional rights by analogy to corporations, which are abstract and artificial beings with the dignity and status of juridical persons and to which are attributed rights as well as liability. The idea that computer-based systems may come to possess human characteristics and assume social responsibilities has produced conundrums such as whether attributes like "blame" or even "punishment" might not be ascribed to them in appropriate circumstances.[24] The intelligent artifacts foreshadowed by "artificial life" technology, discussed in the final Section of this work, may be analogized to other categories of emancipated beings, such as blacks and women, who over time have achieved the status of beings with rights as well as duties.[25] The extent to which legal systems have in the past attributed legal "rights" as well as "duties" to non-anthropomorphic beings indicates the potential of computer-based systems to someday surmount the posture of "legal agent," and to achieve a rank now assigned to corporations. It seems unlikely that any machine will be deemed worthy of being considered a legal person possessed of inherent rights unless it is far more "personlike" and "intelligent" than the mundane systems that we presently encounter. A psychologically perceptive apparatus, possessed of a personality and the capacity to learn and adapt to its environment in different ways, will be suggested as the sort of person-like artifact on which the law might be willing to bestow such rights.

---

23. Section III considers in greater detail the liability for decisionmaking processes programmed on computers that fail to make the opportune decision in different cases than would the human specialist.

24. *See* Wilks, *supra* note 4.

25. *See* Marshall S. Willick, *Constitutional Law and Artificial Intelligence: The Potential Legal Recognition of Computers as "Persons."* 2 PROC. NINTH INT'L JOINT CONF. ON ARTIFICIAL INTELLIGENCE 1271, 1279 (1985).

The distinction between machine and man need not be wholly repudiated for us to recognize that both exist in a continuum. There has been a great deal of discussion in the philosophical literature concerning the applicability of various mental predicates to machines.[26] A sophisticated skeptic, representative of the conservative wing of this debate, is Hubert Dreyfus,[27] who dismisses a priori the notion that psychological predicates could ever properly be applied to machines, however much the science of artificial intelligence might advance technically. The other side of this debate is represented by the work of computer scientists such as Allen Newell and Herbert Simon.[28]

Whether the law will someday permit automatons to rise to a higher station in the hierarchy of legal personality such that they will come to be perceived as entitled to rights as well as burdened by duties will no doubt depend on the extent to which society comes to view future automatons as humanoid or person-like. From this standpoint it is immaterial whether skeptics such as Hubert Dreyfus are correct in their philosophical criticism of the possibility of *actual* artificial intelligence, or whether machine consciousness is in fact feasible.[29] What will be pivotal is whether self-regulatory automatons will become sufficiently person-like in popular estimation.

Societal perceptions are already moving in that direction, if colloquial references are any indication. Lay people commonly use language evoking human-like thought in describing machine activity, and it is likely that this will seem ever more natural as the technology of artificial intelligence progresses. Most of us cannot resist use of some such language when we talk about simple devices, such as chess-playing computer games. It seems natural to characterize a program that calculates a combination of moves as "thinking about" or "planning" its moves, although we do not surmise that the device will feel "embarrass-

---

26. Courts have not hesitated to attribute psychological predicates to machines, and this has sometimes confused the inquiry and entangled legal doctrine in perplexities. *See infra* notes 40-52 and accompanying text.

27. *See* HUBERT DREYFUS, WHAT COMPUTERS CAN'T DO (2d ed. 1979).

28. *See, e.g.*, Allen Newell & Herbert Simon, *Computer Science as Empirical Inquiry*, *in* MIND DESIGN 35 (John Haugeland ed., 1981).

29. Expectations occasioned by the science of artificial intelligence have led at least one contemporary philosopher to forebode: "[T]he issue of computer consciousness is not science fiction, arising from an overdose of movies and marketing hyperbole, but rather that conscious machines are conceptually and, in the end, technically feasible." Dan Lloyd, *Frankenstein's Children: Artificial Intelligence and Human Value*, 16 METAPHILOSOPHY 307, 308 (1985).

ment" if we succeed in checkmating it.

C. Machines as the Possessors of Rights

It should not escape our attention that legal systems have often treated nonhuman entities as if they possessed lives of their own, with the result that liability has been attributed to animals and even to insensate physical objects. Oliver Wendell Holmes, for example, called to our attention the extent to which common law inclinations to hold inanimate things responsible have been carried over into contemporary admiralty jurisprudence.[30] John Chipman Gray noted that temples in ancient Rome and church buildings in the middle ages enjoyed a separate corporate existence as legal persons.[31] Some legal systems even assign legal rights and duties to animals.[32] The common law of England ascribed criminal culpability to accursed things, inanimate objects called "deodands," and required that weapons, implements, or instrumentalities that caused a death be forfeited to the Crown. As late as 1842, a railroad locomotive

---

30. Justice Holmes wrote the following:

> It would seem that a similar form of words has been enough to satisfy the minds of great lawyers. The following is a passage from a judgment by Chief Justice Marshall, which is quoted with approval by Judge Story in giving the opinion of the Supreme Court of the United States: "This is not a proceeding against the owner; it is a proceeding against the vessel for an offence committed by the vessel; which is not the less an offence, and does not the less subject her to forfeiture, because it was committed without the authority and against the will of the owner. It is true that inanimate matter can commit no offence. But this body is animated and put in action by the crew, who are guided by the master. The vessel acts and speaks by the master. She reports herself by the master. It is, therefore, not unreasonable that the vessel should be affected by this report." And again Judge Story quotes from another case: "The thing is here primarily considered as the offender, or rather the offence is primarily attached to the thing."

HOLMES, *supra* note 10, at 29.

31. John Chipman Gray commented:

> [I]nanimate things have been regarded as the subjects of legal duties,—I was about to add in primitive times, but, as we shall see, the notion has persisted even to our own days. If there was a fiction here, it was not in attributing the real will of a human being to the thing, but in assuming that the thing had an intelligence of its own. It would seem, however, that there was often no conscious fiction, but some vaguely realized belief that the thing had a true intelligence and will.

JOHN CHIPMAN GRAY, THE NATURE AND SOURCES OF THE LAW, 46-47 (2d ed. 1921).

32. *Id.* at 43, 47.

was forfeited as a "deodand."[33] If judges and scholars of the common law, well aware that ships are no more alive than are stones, did not consider it absurd to treat inanimate objects as if they were alive for purposes of attributing liability to them, it is not far-fetched in the least to imagine a modern court assigning responsibility to computerized agents that are used in place of humans to perform sensitive tasks.

Man and machine are discontinuous conceptions. Computers are but conglomerations of electronic machinery programmed to perform tasks in designated ways in specific circumstances, while humans are thinking, feeling, spontaneous, conscious creatures capable of responding to their environment in what we like to consider a uniquely "human" manner. The crucial difference between a person and a mechanical alternative that attempts to impersonate or substitute for a human is the human ability to "comprehend," to relate emotionally, and to respond creatively to the environment.[34] The most technologically advanced present-day computers respond to their surroundings in stereotypically sterile ways. A computer cannot "think," at least in the conventional sense of the word, nor is it capable of responding on an emotional level to the decisions it makes. We assume that, regardless of how sophisticated automation technology becomes, mankind will continue to function on a different, and indeed higher, plane. There is an element of comfort in this belief, but perhaps we assume too much. Courts, as I will demonstrate, have long acknowledged circumstances in which even simple vending machines possess legal autonomy for some purposes. It is unnecessary to attribute

---

33. *Id.* at 47; *see also* HOLMES, *supra* note 10, at 25.

34. The question of whether a machine might ever constitute a functional substitute for a human in the sense that it could be said to think led the mathematician, Alan Mayne Turing, to propose what is now referred to as the "Turing Test." To apply the Turing Test, we place a human being in a room with two computer terminals, one connected to the computer in question and the other connected to another human being. The person in the room would not be told which terminal was linked to the computer, and would endeavor to carry on "conversations" utilizing written messages sent and received through both terminals. By use of these "conversations," he would seek to determine which of the terminals was linked to a computer and which to another human. We might expect that he could soon determine whether he was communicating with a machine, but conceivably, if a computer had been programmed to converse in a sufficiently clever, amusing, and sophisticated way and had a sufficiently broad store of knowledge, the participant in this test might not find it possible to ascertain whether he had been conversing with a person or with a machine. According to Turing, such a computer for all practical purposes may be viewed as capable of thinking. A.M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433, 433-68 (1950). Within already existing legal contexts, issues involving the status of a Turing machine as an independent actor will almost surely be affected by the development of computers which cannot be identified as machines by the unsuspecting humans who interact with them.

an attenuated consciousness to machines, or even to accept the fact that they have a capacity to think as humans do, to warrant exploration of the extent to which unattended computerized systems will be viewed as emancipated entities.

Courts have seldom balked at conferring on machines a legal identity separate from their proprietors. Consider, for example, *McCaughn v. American Meter Co.*,[35] in which the issue of machine autonomy arose in the context of how purchases from coin-operated gas meters should be categorized for tax purposes. The court in *McCaughn* held that prepayment gas meters were capable of autonomously conducting sales transactions with purchasers. The court found that the "contract, sale, delivery and payment" were all "effected by mechanism, automatically and without any human agency," and viewed the machine as more than a mere extension of the proprietor's hand.[36] From the standpoint of taxation, at least, the machine was characterized as an independent participant, indeed a party to the sales transaction.

Courts that have been persuaded that a vending machine might autonomously enter into a contract, however, have not been inclined to attribute similar independence to machines in a criminal law context. *People on Complaint of Nicoletti v. Gargivlo*[37] illustrates that the criminal liability arising out of a contract effected autonomously by machine cleaves nonetheless to the proprietor. In this case, a police officer entered the defendant's place of business, inserted coins in a pinball machine, and achieved a score that entitled him to "points" under a schedule posted on the machine. Officer Nicoletti called his score to the proprietor's attention, and was given ten cents. Had the proprietor himself offered the "points," he would explicitly have entered into an illegal gambling contract, and his guilt would have been manifest. The defendant, however, argued that the statutory requisite of an "express" agreement had not been alleged and could not be proved, because a pinball machine is not suited to entering into "express" agreements. Although the crime of gambling was legislatively defined in terms of human activity, the court found that

> the card on the machine constituted an offer by the defendant

35. 1 F. Supp. 753 (E.D. Pa. 1932), *rev'd*, 67 F.2d 148 (3d Cir. 1933).
36. 1 F. Supp. at 754.
37. 298 N.Y.S. 951 (Magis. Ct. 1937).

> to the player that if he succeeded in obtaining a certain prescribed score he would be entitled to certain "points." This offer was accepted when the police officer inserted a coin in the machine and the play began. The agreement was fully executed when the defendant gave the police officer 10 cents for the score he made.[38]

Implicitly, the court's position was that "acceptance" of the police officer's coin, on the proprietor's behalf, made the machine an intermediary in the process of contract formation. Because the question of machine autonomy emerged in a criminal context, though, the court considered it more fitting that the human, rather than his machine, should be punished.

Notions of autonomous machinery have thus far presented few challenges to the law. It is reasonably evident to an individual what may be expected when his counterpart is a machine. The sequence of events the machine is programmed to accomplish is usually known and unambiguous. But what ought to happen when the machinery malfunctions or when for some other reason events fail to transpire as anticipated? Computerization will challenge courts of the future with these issues. For example, a sequence of significant events may depend on random or computer-dependent variables such as alphabetic sequences or identification numbers that otherwise would not be considered legally significant. It may well be that yet unknown categories of legal analysis will be devised and that lawyers will have to learn the language of computers. An attempt has been made to seek out contract as well as tort theories under which liability may be predicated to machines. As we engage in transactions with artifacts of increasing sophistication, machines that react in ways that we would expect persons to behave, the notion of autonomous liability becomes increasingly worthy of investigation.

II. LEGAL RESPONSIBILITY HISTORICALLY TO UNATTENDED INTELLIGENT SYSTEMS

This Article examines disparities in liability exposure based on whether the claim-producing activity is accomplished by a person or by a

---

38. *Id.* at 953.

machine. The differences in treatment suggest that "unattended" systems may be evolving into legally distinct entities. James Moor admonishes:

> It is important to understand computer activity in some contexts as decision making not only because it is so, but because to see it otherwise tends to minimize our appreciation for the potential impact of computers on society. To delegate decision making is to delegate control. Ultimately, the issue is what aspects of our lives, if any, computers should control.[39]

That unattended "agents" may be held legally accountable for the decisions they make suggests that a new body of law is emerging in response to problems precipitated when humans interact with machines.

It should not escape our attention that when computerized systems make decisions and act on them, they operate in a manner identical to that of humans. Both machines and people acquire and apprehend information to which they then apply some criteria or rule of decision which resolves what is to be done in the circumstances. Whether decisions are made by a person or by a computerized system, the process is the same. However, courts do not impose comparable liability when these decisions result in a legal claim.

Conventionally, we conceive of computer machinery as a tool used by people. However, under existing case law, applyng this approach to cases dealing with automated machines results in a drastic oversimplification. The "autonomous" nature of some unattended machinery has long been recognized in court decisions. Whether an interactive machine is characterized as a mere tool or as an agent depends on the degree of autonomy it exhibits, as well as the circumstances in which it is utilized. The following Part will explore cases in which automated tasks produce less legal liability than a comparable task accomplished by a person.

---

39. James H. Moor, *Are There Decisions Computers Should Never Make?*, 1 NATURE & SYS. 217, 219 (1979).

A. Machines Subjected to a Lower Standard of Accountability

Situations presented by the case law in which the performance of an interactive machine is at less jeopardy of liability than a functionally comparable human performance. The cases provide evidence for the inference that a range of possible legal statuses exists along which various categories of automated instrumentation might be systematized. The range extends from machines being characterized as mere tools to attaining the status of "legal agents." We are arriving at a point at which "unattended" systems have evolved into emancipated entities which are, from a legal point of view, accountable as "agents" for decisions they make.

Let us first consider a category of cases in which people have been displaced by mechanical devices. Some transactions permitted to unattended intelligent machinery are prohibited to their human counterparts, while other transactions, by contrast, are allowed to humans but forbidden to unattended machinery.[40] The seeming inequity of such rules has not gone unchallenged. The *New York Times* reported that Mr. Reader, the president of First Stamford Bank & Trust Co., inveighed against the use of automatic teller machines by his competitors.[41] Such machines, he argued, compete unfairly with human tellers, because they permitted his competitors to provide banking services on Sunday when Connecticut law forbade him from opening his bank. Mr. Reader claimed that automatic teller machines should be compelled to observe the

---

40. *See* People v. Andor Corp., 206 N.Y.S.2d 89 (Westchester County Ct. 1960) (holding that operation of a coin-operated, self-service laundry establishment on Sunday did not violate the Sabbath Law if neither the owner nor the employees were on the premises); Dellwood Dairy Co. v. City of New Rochelle, 165 N.E.2d 566 (N.Y. 1960); City of Newark v. Daly, 205 A.2d 459 (N.J. Super. Ct. 1964) (holding that vending milk by machine in apartment buildings does not violate zoning regulations as business conducted in a residential premise), *aff'd*, 214 A.2d 410 (N.J. 1965).

41. Diane Henry; *Blue Law vs. Automatic Teller*, N.Y. TIMES, Nov. 10, 1978, at B3. Ever more ubiquitous, prodigious, and prententious interactive machines continue to insinuate themselves into legally significant human transactions. One such machine, called "Victor," similar in design to an automated bank teller, is employed as a court clerk in the Municipal Court of Long Beach, California. Victor not only instructs motorists how to pay traffic and parking tickets, but offers advice on how to prosecute a claim in the small-claims court as well. It accepts guilty pleas, accepts payment of the fine it imposes by check or credit card, and, in cases where a user wishes to enter a plea of not guilty, Victor will in appropriate cases grant adjournments of scheduled hearings, offer alternate trial dates, and where necessary, even calculate bail. *See* Debra C. Moss & Mark Hansen, *Autoclerk*, 78 A.B.A. J. 34 (1992).

same legal prohibitions as humans. The existence of other transactions permitted to human actors but forbidden to their mechanical replacements stands in rebuttal to Mr. Reader's contentions and illustrates that there are times when it pays to be human![42] Vending machines that sell condoms or cigarettes are examples of interactive machines that have in some cases been prevented from standing in for people.[43]

Cases involving the automation of lockers, parking facilities, and other receptacles constitute another category of decisions that exemplifies the differences in the liability engendered by human actors and their mechanical alternatives. In this area, courts have applied psychological concepts such as "consent" or "promise" to increasingly ubiquitous and autonomous unattended systems. Some of these decisions analyze the legal impact generated by such automated facilities almost ritualistically, by classifying the category of activity within existing legal paradigms. These courts assess the extent to which the automated deposit of goods fits within a legal framework which denominates the legal relationship in terms of whether it constitutes a "sale" or "bailment," a "lease" or "license," or some other existing rule.

1. Bailment Liability Imposed on Machines

Notwithstanding numerous decisions holding that coin-operated apparatus may effectuate a "sale,"[44] use of coin-operated lockers does not, as a rule, constitute a "bailment." Courts have not found a mecha-

---

42. For example, the American Bar Association has taken the position that the utilization by lay persons of analytical or advisory legal information and research systems is tantamount to unauthorized practice of law. *See* George G. Lorinczi, *When Does the Computer Engage in Unauthorized Practice?*, 54 A.B.A. J. 379 (1968).

43. The issue that arose in the context of vending machines selling condoms and that reemerges today in the context of cigarette vending is illustrated by Cavalier Vending Corp. v. State Bd. of Pharmacy, 79 S.E.2d 636 (Va. 1954). In that case the proprietors of the vending machines argued that regulations that prohibited the sale of condoms by machines but did not bar sales by persons were unconstitutional. The court rejected this constitutional challenge and emphasized that vendors of condoms were bound to a strict standard of care restricting condom sales to the purpose of disease prevention and forbidding sales to minors. In concluding that the regulations in question did not discriminate against the plaintiffs, the court commented: "There is nothing in the statute which prevents appellants from selling the devices. . . . They are simply prohibited, like all other persons, from selling them through vending machines." *Id.* at 640; *accord* Sanitary Vendors, Inc. v. Byrne, 190 A.2d 870 (N.J. 1963).

44. *See* American Meter Co. v. McCaughn, 1 F. Supp. 753 (E.D. Pa. 1932) (holding that a coin-operated gas meter autonomously entered into a contract with the purchaser, a contract "effected by mechanism, automatically and without any human agency"), *rev'd*, 67 F.2d 148 (3d Cir. 1933).

nized facility capable of "consenting" to bailment liability on behalf of its human principal.[45] The courts' decisions are not entirely consistent, as is evidenced by the fact that courts do impose bailment liability in other functionally equivalent automated receptacle transactions, such as the use of safe-deposit boxes.

Automated receptacle cases as a category illustrate the extent to which decisions involving automated facilities have fixed on controlling legal pigeonholes. These transactions have been analyzed in terms of whether they ought to be considered bailments, or alternatively leases, and not in terms of what would seem to be a more significant determinant: whether the automated substitute in a particular case is arguably a functional equivalent of the related human transaction. There is little analytical consistency to be found in these cases. Some functionally analogous transactions, such as the parking of vehicles, which ordinarily are categorized as bailments when occurring between humans, are regarded as a lease of space or a license when accomplished automatically. Other functionally equivalent automated receptacle transactions, however, such as those involving bank safe-deposit boxes or night-depository systems, have tended to be classified as bailments.[46]

Another area in which these issues commonly arise is in night-depository banking. In an ordinary banking transaction, the person places money into the hands of a human bank teller. The relationship thus created is one of debtor and creditor, requiring the consent of both parties to the transaction. However, this relationship is altered if the depositor places money with a machine. In *Bernstein v. Northwestern National Bank in Philadelphia*,[47] a bank claimed it never found money which the plaintiff alleged she had placed in a night depository. The relationship intended by the parties was that of debtor and creditor, but placing the money in the machine could not create that relationship because there was no "consent" given on the part of the bank. The court held the use of a night depository constitutes a mere tender for the purpose of making a deposit.[48] Some unequivocal act by the bank is necessary to consummate a deposit and to establish a debtor-creditor relationship. An

45. *See, e.g.*, Marsh v. American Locker Co., 72 A.2d 343 (N.J. Super. Ct. 1950) (holding that an automated locker was not capable of the acquiescence required to participate in a bailment transaction, thereby exposing the proprietor to less liability than if the lockers had human attendants).

46. *See, e.g.*, Carples v. Cumberland Coal & Iron Co., 148 N.E. 185 (N.Y. 1925).

47. 41 A.2d 440 (Pa. Super. Ct. 1945).

48. *Id.* at 441.

inanimate depository cannot provide the requisite act of conscious reception, and therefore is incapable of entering into a consensual relationship on behalf of the bank. Night depositories, though unable to consummate a creditor-debtor relationship, give rise to a mutual benefit bailment which requires the bank to exercise ordinary care and diligence.

Note how the court, by reasoning that an inanimate object is incapable of a consensual relationship, gives rise to contradictory outcomes. In *Marsh v. American Locker Co.*,[49] the inability of the coin-operated locker to "consent" to an assumption of bailment liability led the court to discard the possibility that the transaction in question was a bailment. In *Bernstein*, however, the court held that the inadequacy of the mechanical night depository which prevented it from entering into a consensual relationship prevented it from accepting a deposit, but not from accepting bailment liability.[50] That court held it proper to instruct the jury that once it was established that the money was in fact placed in the night depository, the bank has the burden of demonstrating that it used reasonable care, and will be held liable for an unexplained loss.[51] Perhaps the *Bernstein* court considered the acquiescence required to accept responsibility as a "debtor" greater than the assent required for bailment liability and that in the night-depository contexts, even the tenuous assent of a machine may be sufficient to bring about a bailment.[52] Little attention has been paid to such inconsistencies, and a doctrine within which the discordant decisions may be systematized is not readily discernible.

2. *License Liability Imposed on Machines*

Cases involving parking lots also typify a disparity in the liability of comparable parking transactions. The burden of responsibility in parking transactions has depended on whether functionally comparable incidents of loss are denominated a "lease," as is usually the case when the lot is automated, or as a conventional "bailment," as is more often the case when the loss transpire in a manned facility. The reason automated

---

49. *Supra* note 45.

50. *Bernstein*, 41 A.2d at 442.

51. *Id.* at 443.

52. *See generally Contracts—Requisites & Validity—Whether or Not a Bank May, by Agreement with Its Customer, Completely Exonerate Itself from Liability for Its Negligence in Maintaining a Night Depository*, Note, 29 CHI.-KENT L. REV. 334 (1951); Gerald O. Dykstra, *The User of a Bank's Night Depository Facilities*, 70 BANKING L.J. 121 (1953).

parking transactions are commonly characterized as a "lease" or "license" rather than as a "bailment" is that courts have been reluctant to apply the psychological concept of consent to inanimate objects. This results in the attenuation of responsibility for the losses transpiring in automated lots. Thus, proprietors of automated garages can usually avoid the added liability found in the bailment, which is imposed on otherwise indistinguishable conventional parking facilities.

Consider, for example, the following excerpt from the dissent in *Ellish v. Airport Parking Co. of America*,[53] a case in which the proprietor of a parking lot controlled by an automatic machine was absolved of any liability for the loss of a vehicle in his lot:

> In our view, the shift to lessened personal contact between the lot operator's employees and his patrons in no way changes the basic nature of the relationship between the lot operator and his patrons so long as he retains dominion over the cars parked in his lot and can withhold their return until he is paid the full fee for the parking. It would be ironic if the operator's use of additional cost-saving devices were read as lessening his responsibility to use due care in protecting cars parked with him from theft.[54]

---

53. 345 N.Y.S.2d 650 (N.Y. App. Div. 1973), *aff'd*, 316 N.E.2d 715 (N.Y. 1974).

54. 345 N.Y.S.2d at 658. It is frequently held that claim tickets dispensed by automatic machines at the entrance to a parking lot do not qualify as bailment transactions which would impose on the proprietor a duty to safeguard the vehicle. Characterizing a parking transaction as a bailment rather then as a lease or license has the effect of imposing on the owner of the lot this duty of vigilance to protect the vehicle. Courts consider the patronage of automated facilities as a lease or license to use the parking space rather than as a bailment, which is used to define the liability customarily imposed when automobiles are stolen from manned parking facilities. A majority of the Appellate Division held that in automated facilities, the risk of loss should fall on the customer, rather than on the operator of the lot. *Id.* at 655. Justice Shapiro remonstrated in dissent that displacement of human parking lot attendants by automated facilities effected an unwarranted diminution of liability and shifted the risk of loss from the proprietor to the consumer. *Id.* at 657-58. The facts of the case are interesting. Ms. Ellish had driven her car into the enclosure of an airport parking facility after removing a ticket from an automated machine stationed at the entrance by the proprietor. Removing the ticket from the machine, which stamped it with the time and date, activated a gate and permitted entry into the lot. Once inside, the vehicle could only exit through another gate attended by human employees, who required that the ticket be surrendered. If a driver seeking to exit was unable to produce a ticket, the human attendants would not release the automobile without proof that the driver owned the vehicle. When Ms. Ellish returned to retrieve her car, she found that it had been stolen from the lot. Apparently, the machine had dispensed a ticket to a larcenist pedestrian who paid a parking fee to the gate attendant and absconded with Ms. Ellish's vehicle. The machine that supervised entry into the lot was not designed to discern whether a person removing a ticket

By characterizing automated parking lots as effectuating license or lease contracts rather than bailment contracts, courts regularly sidestep the question of a proprietor's negligence. In other types of cases, defendants have been held duty-bound to safeguard others from criminal activity which they can reasonably expect to occur, including situations involving the lease of a space.[55] Certainly, the law should encourage use of the

---

proffered by the machine had in fact brought a car into the lot. The circumstances required prudence on the part of the automatic gate, but the design of that mechanism was woefully inadequate. Unmindful of whether vehicles were driven into the lot, it carelessly dispensed tickets to any passerby. The stratagem for this thievery is explained in Makower v. Kinney Sys., 318 N.Y.S.2d 515, 518 (N.Y. County Civ. Ct. 1971), as follows:

> It is true that the use of a machine instead of an attendant to hand out tickets adds a little spice to the situation. It creates the possibility that the person presenting himself at the exit may be driving a different car than the one in which he entered, or, indeed, if he is clever and of a more larcenous bent of mend [sic], he may even have come in afoot. The use of the machine, however, is not dictated for the convenience of the customers. It is dictated rather by the desire to obtain the savings in manpower made possible by modern technology. It is a calculated risk the operator is taking. But just because it makes later theft easier does not affect the question of whether a bailment is created when a car enters the lot.

Id. at 518. Case law on the precise point of whether it is negligent for a machine to be designed so as to dispense a ticket on entry into a parking lot without first ascertaining that a vehicle has been brought into the lot is scanty because *Ellish*- type machines known as "ticket-spitters" have for the most part been replaced with the familiar automated gates which do not dispense a ticket unless a car is driven into the lot. This up-to-date technology is more competent in the sense that it exercises judgment to a greater degree than the ticket-spitter it replaced. The fact that the type of machine commonly utilized by automated parking facilities has evolved in this way is suggestive of the proposition that interactive machinery must be designed so as to employ the best available technology. Although contemporary ticket dispensers are not "foolproof"—a criminal could drive a stolen vehicle into an automated lot, abandon the car with which he entered and use the ticket he obtained to steal another more valuable vehicle—use of the more sophisticated automated gate nevertheless considerably curtails the risk of thievery.

55. *See* RESTATEMENT (SECOND) OF TORTS § 302B cmt. e (1965), which states:

> There are, however, situations in which the actor, as a reasonable man, is required to anticipate and guard against the intentional, or even criminal misconduct of others. In general, these situations arise where the actor is under a special responsibility toward the one who suffers the harm, which includes the duty to protect him against such intentional misconduct; or where the actor's own affirmative act has created or exposed the other to a recognizable high degree of risk of harm through such misconduct, which a reasonable man would take into account.

See also Sparrow v. Airport Parking Co. of Am., 289 A.2d 87 (Pa. Super. Ct. 1972) (holding that even if an automated parking arrangement gives rise to a lease or license rather than a bailment, a duty to protect the customer's property may still be inferred from the contractual agreement); Garlock v. Multiple Parking Servs., Inc., 427 N.Y.S. 670 (Buffalo City Ct. 1980) (holding that a determination that the parking transaction did not constitute

best available technology so that automated systems will be operated in a manner which protects the public and prevents loss.[56] Yet dogmatic judgments arrived at by pigeonholing automated transactions into particular legal paradigms highlight the inadequacy of rules that apply psychological concepts such as "consent" or "promise" to cases involving automated facilities. This approach fails to take into account the entire substratum of human behavior, temperaments, perversity, and other consequential anthropological attributes that are eliminated in an automated process. It also minimizes the importance of the expectations of the parties in automated transactions. Most importantly, this technique of analysis invites absurd and capricious results.

When insinuating that an interactive machine increases, diminishes, or relieves its owner of liability as compared to that imposed on functionally analogous non-automated transactions, courts grant the automated facility a significant degree of legal autonomy. This willingness to accept automatons as having some degree of legal autonomy indicates that courts have not repudiated as absurd the notion that mechanical intermediaries might function as responsible "agents" for their proprietors. The extent to which courts will continue to treat ever more autonomous and accomplished mechanical stand-ins as agents remains to be seen.

B. Machines Subjected to a Higher Standard of Accountability

In contrast to the "low-scrutiny" decisions which employ formalistic analysis, the liability occasioned when a human activity has been displaced by an automated system sometimes depends on whether an autonomous machine was "competent," in the sense of adequate to the task of standing in for a person. Cases that arise out of rudimentary automation technologies exhibit a twofold theme: first, whether the particular automated system is sufficiently competent to avoid untoward hazards in the displacement of the human actor in a transaction, and second, how losses generated by the lack of human attributes appropriate to the transaction should be distributed.[57] The touchstone in this type of

---

a bailment of itself was not conclusive on questions of parking lot liability).

56. *See, e.g.*, The T. J. Hooper, 60 F.2d 737 (1930).

57. *Compare* Ellish v. Airport Parking Co. of Am., 345 N.Y.S.2d 650 (N.Y. App. Div. 1973) (holding that a flawed automated parking lot gate, which dispensed tickets to persons whether or not they entered the lot with an automobile, and which therefore facilitated

case is whether the machine should be permitted to replace a human actor and, if so, whether a greater burden of liability ought to be imposed on the automated routine. Where machines satisfactorily substitute for humans, the legal relationship between consumer and machine should mirror that which the consumer formerly had with the human agent.[58] Yet, there are instances in which a greater burden of liability is assigned to an automated process because the process lacks an adequate analogue for some critical human response indispensable to the proper functioning of the system.[59]

Consider the unlikely case of a coin-operated, gun-vending machine as an example of a contrivance quintessentially unfitted to substitute for a

---

removal of the victim's automobile by a thief, did not create an unreasonable hazard) *with* New York v. Citibank, 537 F. Supp. 1192 (S.D.N.Y. 1982). Although the latter case, concerning automatic teller fraud, settled, the Assistant Attorney General, Stephen Mindell, alleged that the automated teller machines utilized by Citibank did not make use of the best available technology. Indeed, the settlement agreement required that the bank file with the court under seal a description of the security measures it agreed to implement to prevent a recurrence of the scheme. Settlement Agreement, New York v. Citibank, 537 F. Supp. 1192 (S.D.N.Y. 1982) (No. 81-7273). Ongoing refinements of automation technology will require courts not only to approve or disapprove automation of a transaction by the imposition of liability for the losses it occasions, but also will directly involve courts in the establishment of constraints on the design of automated facilities.

58. *See, e.g.*, State v. Arnold, 258 N.W. 843 (Wis. 1935) (holding that a condom machine was not capable of meeting the statutory standard of care that permitted the sale of condoms only to married persons for the prevention of disease, but not to prevent pregnancy). The court in *Arnold* refused to establish a lesser standard for machines than that imposed by the statute on human sellers, and imposed criminal liability on the proprietor.

59. See, for example, McEvens v. Citibank, 408 N.Y.S.2d 870 (N.Y. County Civ. Ct. 1978), which held a bank responsible for money allegedly lost during an automatic teller transaction. Judge Nardelli stated:

> [T]he bank could have better protected itself and more importantly, its customer, by following . . . some form of recording surveillance device in the teller's cage which could, at a later time, show and corroborate every step of the transaction from the opening of the lockbox and the unsealing of the envelopes to the making of the actual count and crediting of the account.

The judge's remarks suggest a legal trend toward demanding that factors essential to the adequate performance of an interactive machine be built into the system and that machines that fail to incorporate essential human factors lost in the process of automation will have to be redesigned.

For example, categorizing transactions involving safe deposit facilities as bailments exposes banks to fraudulent claims in circumstances where they lack knowledge of facts which would enable them to refute spurious fabrications. In Veihelmann v. Manufacturers Safe Deposit Co., 104 N.E.2d 888 (N.Y. 1952), a case involving an unexplained loss from a safe deposit box, the New York Court of Appeals regarded the transaction as a bailment and concluded that the bank might be held liable for an unexplained loss of goods merely alleged to have been deposited in the box.

human actor. Society relies on the good judgment of human gun dealers who are expected to limit sales of their wares to sane adults without criminal records. A gun-vending machine lacks the capacity to exercise such judgment and would sell a gun to anyone with the appropriate payment. Cases involving sales by unattended machines raise provocative questions regarding the extent to which mechanical analogues of prototypical human attributes, such as judgment, must be embodied in interactive machinery.

Another human factor the law may require of a machine is the ability to give an account of the transactions in which it participated, or "recountability." For example, human gun dealers may be called on to give an account of the transaction and to provide information about the gun sold, as well as a description of the purchaser. Present-day vending machines lack the capacity to record images of customers with whom they deal, although it is not difficult to construct vending machines that do so. Omission of this surveillance capacity gives rise to legal questions and judicial holdings that affect the design of mechanical surrogates. Indeed, it is becoming obligatory in the banking industry to endow automated teller machines with this capability.[60]

Not only "incomplete" interactive machines, incapable of a full range of human capabilities, generate hazards calling for legal control. Ordinary computerized systems sometimes introduce hazards precisely because of their ability to perform a particular human task with superhuman proficiency. One such situation reached the Supreme Court in *Whalen v. Roe*.[61] In this case, a state statute required physicians to identify patients using prescription drugs, such as cocaine and methadone, so that this confidential data could be kept on file in a central computer database. Patients who regularly received prescriptions for the drugs in question for their legitimate medical treatment and physicians who prescribed these medications brought an action challenging the constitutionality of this statutory requirement. The Supreme Court rejected the plaintiffs' argument that a potential risk of unwarranted disclosure was a sufficient ground to strike down the statute.[62] The court was equally unresponsive to plaintiffs' contentions that the fact that "the information

60. *See generally* BANK ADMINISTRATION INSTITUTE TASK FORCE, ATM SECURITY HANDBOOK 69 (2d ed. 1988). This is also well illustrated by Judge Nardelli's remarks in *McEvens* , 408 N.Y.S.2d at 870.

61. 429 U.S. 589 (1977).

62. *Id.* at 600.

is readily available in a computerized file creates a genuine concern that causes some persons to decline needed medication."[63] The Court's analysis in *Whalen v. Roe* implicitly analogizes the colossal computerized system there in issue to a human record keeping system. It concluded that if humans may maintain drug prescription records, no legal impediment exists to prevent machines from keeping those same records.

The contemporary legal issue is not whether a replacement computerized system approaches the aptitude of a human actor, but whether the machine has a basic competence adequately to perform the essential components of the transaction or process. Consider for a moment those courteous contrivances, such as automated teller machines that say "good morning." Although this machine provides but a flimsy approximation of a human greeting, it is difficult to imagine that the legal quandaries arising from this shortcoming are sufficiently serious or threatening to require a judicial response. However, there are cases in which the capacity of a computerized system to interact sympathetically with humankind is so compelling that courts have decreed that critical human attributes such as courtesy be integrated into the system.[64]

In all of these cases, what separates human and machine is that machines conventionally are not considered responsible when they cause injury. When mischief emanates from a machine, we believe that a human being must have acted irresponsibly; injured parties do not look to the machine for satisfaction, but instead to a human in charge of the machine. Still, disparities in the liability relative to whether the task was accomplished by human or machine suggest that "unattended" systems are evolving into emancipated entities that are held accountable for the decisions they make. In this sense they function as agents.

---

63. *Id.* at 602-03.

64. *E.g.*, Palmer v. Columbia Gas of Ohio, Inc., 479 F.2d 153 (6th Cir. 1973) (holding that absence of civility in an automated system detrimentally affecting the lives of thousands of consumers could not be permitted, and ordering the gas company to interpose human intermediaries who would be empowered to take a more responsive and accommodating attitude in dealing with billing adjustments than could an inadequate computer system). *See also* ERNEST W. KENT, THE BRAINS OF MEN AND MACHINES 271-72 (1981) (suggesting that we will be inclined to construct mechanical analogues of emotions in complex computerized systems to the extent advances in technology permit our doing so); *see generally* Theodor D. Sterling, *Guidelines for Humanizing Computerized Information Systems: A Report from Stanley House*, 17 COMM. ASS'N FOR COMPUTING MACHINERY 609 (1974).

C. Analysis of Case Law

Though most of the cases mentioned thus far have been tort cases, questions of automated system liability and consideration of psychological motifs such as negotiation arise in contracts as well. The law has long considered that vending machines make offers, which humans accept. It may seem odd to find decisions applying the language of offer and acceptance to circumstances in which people deal with machines, for in its ordinary sense "negotiation" is an attribute of human behavior. We are accustomed to thinking that parties to an agreement "bargain," in the sense of employing psychological strategies in order to arrive at an understanding that each considers sufficiently advantageous to formalize into a contract. But time and again "sticky" situations have arisen in which courts have been called on to decide whether a particular interactive machine enjoyed sufficient autonomy to bind its principal to contract terms that were not initially intended or that transformed the essence of the contemplated arrangement. While such cases are not yet numerous enough to permit a systematic critique of human-machine contracts, they do indicate some of the concerns courts have identified as important. From the standpoint of our analysis, they also exemplify disparities in the contractual obligations that arise when one party to an understanding is a machine, as compared to contract transactions between humans.

Where contracts proffered by a machine are complex, or include obscure terms that undermine the ordinary expectations of a purchaser, the courts have often rewritten the contract in terms unfavorable to the seller. Anyone who has spent an afternoon attempting to read the simplest insurance policy knows how complex these contracts can be. When an insurance policy is purchased from a human salesperson, there is at least the possibility that the particulars of the agreement have been explained. Purchasers who have some understanding of the terms in a proffered policy, and who signed the application or paid the premium, may be deemed to have assented even to the contract's most cryptic terms. However, courts may respond differently when the insurance policy is proffered by a machine that neither displays the details of its terms nor provides the vendee an opportunity to obtain an explanation of problematic terms. It has been held that consumers are not bound by express terms of the insurance contract at variance with their reasonable expectations, at least in situations where the consumer was not aware of

the express limitation prior to purchasing the policy from the machine.[65]

Use of the reasonable expectation of a purchaser as a criterion for determining the validity of the terms of a contract purchased from a machine may also require that the mechanizers of the contract transaction be bound by a machine's perverse conduct. Consider the following anecdote, which was reported in the *New Yorker Magazine*:

> Fred Finn Mazanek, a one-year-old guppy, died recently, leaving an estate of $5,000. Stan Mazanek, twenty-four, a student at the University of Arizona, had filled out an insurance form he received in his mailbox marked "Occupant," entering the fish as the insured party. No fraud was involved in the policy. The guppy's age was listed as six months, his weight as thirty centigrams, and his height as three centimeters.
>
> The Glove Life & Accident Insurance Co. apparently issued Policy No. 3261057 in Fred Finn's name through a computer error. When Mazanek filed a claim following the guppy's demise, they sent a sales representative to see him to find out if he was the sort of person who would take advantage of a clerical error.
>
> He was. The company settled out of court for $650.[66]

The reason Glove Life & Accident Insurance Co. was inclined to make an offer of settlement to Mr. Mazanek was not that the company had a sense of humor or was in a particularly generous mood that day. Rather it is because courts are inclined to hold insurance companies bound by the eccentric behavior of their automated agents to a greater extent than for mistakes caused by human agents.[67]

---

65. *See* Lachs v. Fidelity & Casualty Co. of N.Y., 118 N.E.2d 555 (N.Y. 1954) (holding insurance company liable to pay on its airline insurance policy notwithstanding the fact that the claim arose out of a crash excluded from coverage). The reasonable expectation of people served by a mechanical device substituting for human agency is the pivotal issue in this class of cases; *see also* Steven v. Fidelity & Casualty Co., 377 P.2d 284 (Cal. 1962); Allen v. Beneficial Fin. Co., 531 F.2d 797 (7th Cir. 1976) (finding the defendant's computer-generated disclosure statement did not comply with a regulation promulgated under the Truth in Lending Act, requiring a statement of information containing the terms of the loan and the cost of credit to be presented in a logical and sequential order which the ordinary borrower can be expected to comprehend).

66. *The Talk of the Town*, NEW YORKER MAG., Apr. 4, 1977, at 31.

67. The propensity to impose a greater onus on insurance sales that take place without the intervention of human agents is also demonstrated by Fritz v. Old Am. Ins. Co., 354 F.

State Farm Mutual Automobile Insurance Co. v. Bockhorst[68] also suggests that courts may impose a higher standard of accountability on computerized agents and may be less forgiving of a computer error than of the missteps of a human agent. In this case, a motorist who had permitted his policy to lapse for lack of payment was later involved in an accident in which he caused a fatality. He belatedly tendered his check for the arrears to an insurance agent and brazenly requested that his policy be reinstated retroactively. The agent called the relevant circumstances to his company's attention, particularly the fact that the insured had killed a pedestrian after the policy had been allowed to lapse. However, because State Farm Mutual's computer had not been programmed to deal with this unusual sequence of events, it perversely generated and dispatched a notice to the effect that the policy was retroactively reinstated. The court took note of "the traditional rule that a waiver occurs only when a party in full possession of the facts intentionally relinquishes a known right."[69] The court held that the computer system constituted a competent agent capable of binding its principal in circumstances where a similar decision by a human agent might not amount to intentional relinquishment of a known right.[70]

D. The Future Direction of Case Law

Preceding Parts have considered instances in which deployment of an interactive machine to perform a task increased liability exposure over the level present during a functionally comparable human activity. These Parts also examined instances where the trend was precisely the reverse. Decisions about simple vending machines, automated receptacles, and

---

Supp. 514 (S.D. Tex. 1973), a case involving a consumer killed after mailing his application to the insurance company, but prior to the time the application was approved. The *Fritz* court equated insurance sales effected by mail to sales by machine and concluded that "[w]here the non-human device is used, it is logical to require the company to satisfy the applicant's reasonable expectations which were generated by the media chosen by the company." *Id.* at 518; *see* Riordan v. Automobile Club of N.Y., Inc., 422 N.Y.S.2d 811 (N.Y. Sup. Ct. 1979).

68. 453 F.2d 533 (10th Cir. 1972).

69. *Id.* at 536.

70. Courts have also found liability in analogous cases. *See* Bank of New Orleans v. Western Union, 27 La. Ann. 45 (1875) (holding that transmission by machine of erroneous information is actionable); Cracknell v. Long Island Lighting Co., 285 N.Y.S. 13 (N.Y. App. Div. 1936), *aff'd*, 4 N.E.2d 815 (1936) (holding that a utility company must prove the ongoing accuracy of its metering device to prevail on a contract for electricity measured by meter).

parking facilities, all of which are prototypes for forthcoming unattended intelligent systems, provide a basis from which inferences may be drawn about the future course of automation jurisprudence. The contrast between the liability of unattended automated agents and that of human agents invites a rethinking of assumptions about the nature of liability and presages how forthcoming intelligent artifacts and supermachines may fit into the evolving doctrine. The challenge presented is to unravel and systematize legal doctrine with respect to sophisticated computer-based systems and clever automatons.

"Artificial intelligence," as defined by Marvin Minsky, is the "science of making machines do things that would require intelligence if done by men."[71] But artificial intelligence bears upon much more than the present-day applications of automation technology. It comprehends such things as machines that can diagnose illness and provide medical care, robot workers, or even machines capable of recognizing speech and conversing with people. Yet existing case law, for the most part, has dealt only with simple vending devices and passive unattended mechanisms. It might be doubted that there is much we can surmise about how judges will treat the intellective machines of the future on the basis of how they have regarded primitive automatons such as coin operated lockers, ticket dispensers, and vending machines. Courts will either ignore the presence of computer-based systems and the special problems they present, and simply apply existing doctrine, or they will revamp concepts of liability and loss allocation as applied to an automated environment.

Thematically, the question of whether there is much that the case law of simple and passive machines of yesterday can tell us about the legal status of tomorrow's intelligent artifacts does not require an adjustment to the central thesis of this work. Differences between the liabilities generated by human and automated transactions have been demonstrated, and the inquiry into whether it is the mechanical or the human performance that is exposed to a greater burden of responsibility remains pertinent. Many are disinclined to think of unattended machines that participate in commonplace transactions as "participants" or "actors." But when disparities in liability are found such that the liability generated by a human is not simply replicated in the computerized substitute,

---

71. Marvin L. Minsky, *Introduction* to SEMANTIC INFORMATION PROCESSING at v (Marvin L. Minsky ed., 1968).

unattended systems that supplant human activity may be evolving into more or less emancipated entities responsible for their own decisions and actions.

There is good reason to believe that cases involving the simple and passive machines of yesterday can tell us something about the process by which the law will keep tomorrow's intelligent artifacts in check. Oliver Wendell Holmes began *The Common Law* with the oft quoted teaching: "The life of the law has not been logic: it has been experience."[72] In making an estimation of the jurisprudential aftermath of awaited breakthroughs in automation technology, we ought not to rely on "logic" alone, for what is significant is not the intellective capacity of machines, but the scope and impact of the machines' interaction with people. What is suggestive is not the acumen of "intelligent" systems or any lack thereof, but the impact an automation process has on society. Therefore, the automated devices with which we now interact are legally significant because they have engendered case law that anticipates the legal principles that may come to govern displacement of human activity by intelligent artifacts.[73]

---

72. HOLMES, *supra* note 10, at 1. Many years later Justice Holmes enshrined his admonition in the jurisprudence of the Supreme Court with the formulation, "a page of history is worth a volume of logic." New York Trust Co. v. Eisner, 256 U.S. 345, 349 (1921).

73. When the manner in which traditional doctrines, foundational to negligence or contract, might function in a thoroughly automated environment is examined, the noteworthy repercussions of computerization and their effect on the law will require scrutiny of such questions as whether or not an automated procedure has tampered with the actuality of hazards engendered in the transaction and whether the essence of decisions that induced people to engage in the activity in question has been undermined by the automated process. Automated procedures occasionally circumvent the reasonable expectations of the parties. By and large, however, automation processes are not deliberately designed to accomplish perverse outcomes. If the predominant objective of automation is to promote efficiency and productivity, the import of accomplishing routine transactions with supernatural efficacy may nonetheless transform community judgments concerning basic assumptions: the degree of care required in negligence, or, in contract, whether persons ought to be considered bound by characteristics indigenous to complex machinery that they could not possibly have appreciated. *See, e.g.*, Burnett v. Westminster Bank, [1965] 3 All E.R. 81 (Q.B.). It will not be doubted that if a forthcoming computer-based intermediary does not subvert the reasonable expectations of persons who interact with it, or occasion other untoward side effects, automated transactions are likely to be assimilated within the same legal infrastructure as has evolved in the primordial transactions with automated lockers, ticket dispensers, and vending machines. Otherwise, where anticipated automated procedures or artificial intelligence technologies are apt to bring about a perilous discontinuity in the social substratum, we would expect the law to pursue its historic role of keeping technology within the bounds of human governance by outlawing the mischievous device or procedure.

Consider a spectrum between computerbased systems that do not occasion dire consequences or disturb the reasonable expectations of those interacting with them, and automated processes that bring about such a socially harmful maleficence that they are likely

The empirical "experience" provided by the accumulating case law involving rudimentary unattended machinery, as timeworn as many of these decisions may seem to us today, is not out of harmony with more recent decisions arising out of interactions with advanced computer-based systems. Nothing in this body of case law, which is circumscribed by the fact that the vast majority of the decisions have involved primitive systems, portends that courts lean toward a wholesale revamping of concepts of liability and loss allocation as applied to unattended systems.[74] Law may one day come to reassess its position and take greater

---

to be entirely proscribed. Only at an intermediate position along this spectrum should any question arise of the extent to which an appealing normative theory accounting for the legal status of advanced intelligent machinery may be derived from our experience with more primitive interactive processes. Anticipating how the law will ultimately come to treat sophisticated computerized contrivances that introduce social discontinuities or objectionable aftereffects, in a degree not perceived as sufficiently menacing to call for their outlawing, subsumes conjecture rather than foreknowledge. Tort and contract law will in all likelihood continue to speak in the centuries-old language of "fault," "negligence," "reliance," and "waiver," even when we have arrived at a thoroughly automated environment. Where employing an advanced computational artifact or routine does not affect basic social judgments and relationships, where it does not modify promissory obligations or burdens of responsibility such as those requiring an exercise of due care, and where it does not otherwise generate significant unwanted hazards, we do not foresee creation of novel principles of liability that deviate from conventional doctrines historically associated with such transactions. Alternatively, where an automated process transpires against a societal backdrop in which some of the determinants underlying human decisionmaking are frustrated, to disregard the role of the computer is to overlook a salient anomaly which should affect liability.

74. Such attitudes are exhibited by courts which have had occasion to consider nationwide data base systems. Misuse of information technology has the potential of introducing social discontinuities to an extent that transmutes inoffensive computer machinery into something truly ominous and menacing. Courts nevertheless have been disinclined to reassess traditional doctrine in light of the altered environment and recognized perils created by such systems. Illustrative are cases involving pernicious computerized database schemes such as that which came before the Supreme Court in Whalen v. Roe, 429 U.S. 589 (1977). Machines, particularly sophisticated computers, are so efficient at compiling related data that their very efficiency has come to pose real threats to society. The danger of political oppression and imaginable infringements of privacy is moderated by human inadequacy, the sheer impracticality of creating an enormous omnipotent bureaucracy to inquire into and maintain files on everyone's activities. But computer technology makes it possible to document the behavior of millions of people and to conduct surveillance on a grand scale. Existing computerized systems not only might scrutinize each move we make, but are capable of merging records of our "activities" by processing information contained in the numerous records of our activities. Tax records, police files, educational transcripts, social security data, employment records—information located any computerized file—may be accessed, analyzed, and stored. The technology to construct such an awesome universal database is readily available, and that "monster" apparatus would be far more sinister than Orwell's "Big Brother." The Court has rejected the notion that the fact that computerization increases the potential for unwarranted disclosure and is otherwise oppressive constitutes sufficient grounds for striking down a statutory policy. The Court did not consider the social discontinuity introduced by computerization to be crucial, and emphasized instead that disclosures mandated by the legislation were, "not significantly different from those . . .

account of the distinctive manner in which duties and obligations are discharged in an automated environment. But accumulating evidence indicative of doctrinal inertia is compelling.[75] To the extent that standards of care and degrees of risk arising from an automated activity are compatible with the expectations of similarly situated persons functioning without interactive machinery, we have no reason to predict a sweeping reconstruction of traditional doctrine. The advent of machines that operate not merely to increase the productivity of human labor, but to interact with humankind as if they were themselves human, has prompted the legal system to personify and ascribe legal responsibility to unattended intelligent machinery even to the point of considering intelligent artifacts to be "responsible agents."

As legal systems continue to assert authority over the implementation of automation technology, humane constraints will gradually but deliberately shape the assimilation of computer-based systems into the mainstream of commercial activity. It has long been argued that the judicial and legislative responses to social transformations brought about

---

required under the prior law." *Id.* at 602. In the years since *Whalen,* no court has considered it necessary to place an effective restraint on such technology. *See, e.g.*, Doe v. Axelrod, 545 N.Y.S.2d 490 (N.Y. Sup. Ct. 1989); Volkman v. Miller, 363 N.E.2d 355 (N.Y. 1977); Louise B. v. Coluatti, 606 F.2d 392 (3d Cir. 1979); Perky v. California Dept. of Motor Vehicles, 150 Cal. App. 3d 74 (1983); Kansas v. Harder, 641 P.2d 366 (Kan. 1982); Minnesota Medical Ass'n v. Minnesota, 274 N.W.2d 84 (Minn. 1978).

75. Apprehension about political tyranny and privacy illustrated in cases such as Whalen v. Roe has presaged a parallel social transmutation in the workplace. A United Airlines reservation clerk was penalized for some trivial observations she communicated to an associate at work, which were occasioned by a three-minute telephone conversation with an obnoxious customer. Her conversation had been monitored by management which was of the view that although she was courteous to the customer and conducted herself well, she still should not have indicated her displeasure to a co-worker. The causerie led to her being placed on probation, told to see the company's psychiatrist, and eventually fired. Gary T. Marx, *The Case of the Omniscient Organization*, HARV. BUS. REV., Mar.-Apr. 1990, at 12, 30. Shoshana Zuboff explained the transformation in the workplace brought about by information technology as follows:

> Earlier generations of machines were designed to do essentially what human bodies could, more reliably, and at less cost. With machines, work required less human intervention and, overall, fewer human skills. This process has come to be known as automation. The ideal of automation is the self-diagnosing, self-correcting machine system that runs perfectly without human assistance. Information technology can be used to automate all kinds of work in factories and offices. But unlike other tools of automation, information technology simultaneously registers data about the conversion processes it governs.

Id. at 22. *See generally* SHOSHANA ZUBOFF, IN THE AGE OF THE SMART MACHINE (1988).

by machinery are inadequate for this task.[76] It should not be suggested, however, that technological innovation functions unconstrained by public control. Courts and legislatures are coming to appreciate that the sophistication of computer systems requires that relevant policy decisions be treated as crucial legal issues rather than as inconsequential technical questions. In order to examine the doctrinal differences posed by autonomous machines in a specific context, we next turn our attention to an area in which decisions made by autonomous systems in their performance of sensitive tasks generates a quantum of liability different from the persons employing these systems.

III. ATTRIBUTING LEGAL DUTIES TO INTELLECTIVE ARTIFACTS

We have noted that legal responsibility for "transgressions" is invariably assigned to human actors and that losses caused by intelligent machinery are conventionally ascribed to the owner of the equipment or whoever permitted the computerized system to make a deleterious decision. Searching for accountable human participants conditions us to ignore the possibility that plausible contexts may exist in which intelligent artifacts themselves might be considered answerable.

Elaborating on this theme, Professor John Snapper asserts that there are situations in which we should consider holding a machine itself legally responsible, and he has set forth as an example a hypothetical mechanical medic.[77] Snapper's mechanical medic would have the ability to diagnose a patient's ailment and provide actual medical care as well, such as dispensing medication or administering some other suitable treatment. Although such devices would probably not be as proficient as the best

---

76. The notion that twentieth-century technology has taken over, and that man no longer controls his creation but instead serves it, is not a new idea. Where previously man was subject only to natural forces, such as earthquakes and volcanic eruptions threatening his well being and defying his illusion of power, man now lives in fear of the technology he himself has spawned. Supreme Court Justice Douglas expressed this sentiment: "The search . . . today is for ways and means to make the machine . . . the servant of man. That is the revolution that is coming." WILLIAM O. DOUGLAS, POINTS OF REBELLION 96 (1970). These sentiments have been with us for well over a century. Consider, for example, George Moore, who wrote in 1886 that "[t]he world is dying of machinery; that is the great disease, that is the plague that will sweep away and destroy civilization; man will have to rise up against it sooner or later." GEORGE MOORE, CONFESSIONS OF A YOUNG MAN 124 (1959).

77. *See* John W. Snapper, *Responsibility for Computer Based Errors*, 16 METAPHILOSOPHY 289 (1985).

human practitioners, the ever increasing expense of ministering to numerous patients around the clock with human attendants might necessitate the use of mechanical medics, at least in cases where imprudent diagnosis and treatment would be improbable. In dealing with intellective machinery such as mechanical medics, Professor Snapper argues that we should consider holding the machine itself legally responsible for furnishing proper medical care. His contention is that liability should be imposed on a no-fault basis for harms occasioned by computers used for sensitive matters, and that the law should require that such machines be specifically insured for that purpose.[78]

Professor Snapper's whimsical notion of holding the computer itself accountable for the decisions it makes was provoked by the work of the philosopher James H. Moor, who demonstrated that computerized systems do in fact make "autonomous" decisions.[79] Addressing the arguments of those who believe that computerized devices lack decision-making capabilities, Moor persuasively demonstrates that intelligent machines make decisions in all the philosophically significant connotations of the term. Moor's conception is that when computers make decisions, they operate in a manner identical to that in which humans resolve comparable difficulties. Both machine and human obtain and apprehend information to which some criterion or rule is applied in order to decide what should be done. However, unlike Snapper, Moor rejects any suggestion that a computer might be legally responsible for its decisions. "The kind of computers under discussion are not persons, and although they are causally responsible for their decisions, they are not legally or morally responsible for their decisions."[80]

Critics of Snapper's proposal may also assert that it is pointless to speak of the intelligent artifact being held accountable inasmuch as humans rather than machines will be asked to pay monetary damages. However, Professor Snapper correctly points out that the ability to pay damages is not an indispensible factor in assigning legal responsibility.[81] For example, hospital physicians found blameworthy are seldom themselves called on to make reparations. Physicians typically refer the summons in malpractice actions brought against them to their employer's malpractice carrier, and bear no greater pecuniary jeopardy on a personal

78. *Id.* at 290-91.
79. *See* Moor, *supra* note 39, at 217.
80. *Id.* at 227-28.
81. *Id.* at 290-91.

level than does a computer.

Fiscal concerns aside, there still remains one important distinction, involving a separate component of responsibility, between Snapper's mechanical medic and human practitioners; unlike the mechanical medic, the malfeasant physician bears the moral responsibility for having caused the injuries. Professor William Bechtel has suggested that if responsibility is to be shared with computer-based systems, we should consider whether moral responsibility might be assigned to appropriately programmed computerized machinery.[82]

From an ethical standpoint, we properly assign moral responsibility to an instrumentality that played a merely causal role in a transaction only in a case where it makes sense to praise or blame the act at issue. We consider voluntariness and emotional response from a moral point of view because these matters excuse otherwise culpable acts, decisions, or behavior. However, intellectual artifacts lack a capacity for voluntariness and emotional response. Moreover, distinctions between acts that are voluntary and those that are not are absurd and have no meaning in the context of computer decisionmaking.[83] Professor Snapper has rejected the notion that moral responsibility need be assigned in order to hold a computerized device legally responsible for its actions. In contradistinction to moral responsibility, legal liability is not conditioned on an activity's voluntaristic quality.[84] Nor is it indispensable to legal responsibility that an agent be possessed of an attitude toward the relevant decision in the sense of having had a capacity to respond to it emotionally. "When concentrating on material objects as a cause of harm and a source of liability, we tend to ignore the issues of intent and mental

---

82. William Bechtel, *Attributing Responsibility to Computer Systems*, 16 METAPHILOSOPHY 296, 297 (1985).

83. Snapper, *supra* note 77, at 294. Professor Snapper sets out the argument as follows:

> Deliberate choice is treated by Aristotle as a sort of voluntary act, but the distinction between voluntary, involuntary, and nonvoluntary makes no sense in the context of computer activity. These distinctions turn on the actor's mental attitude towards his actions, in particular on his desires and regrets. I now argue that although a computer may make a deliberate choice, it cannot take an attitude towards that choice. An Aristotelian emphasis on voluntariness as a prerequisite for responsibility, then, takes computer decision entirely out of the realm of acts for which one acquires responsibility. It makes nonsense out of an attempt to treat a computer as an agent.

Id.

84. While voluntary action is not an indispensable component of civil liability, it is almost always a necessary component of criminal liability.

attitude that characterize our disputes over responsibility for harm caused by human action."[85] This led him to take the position that doctrines of strict liability would have to be utilized if a legal context for attribution of responsibility to intelligent artifacts is to be found. In the legal context of strict liability, it is sufficient that the responsible entity has played a causal role in the process that occasioned a loss.

A. Existing Intelligent Devices

Recent technological advances have led to the creation of computerized medical devices that come quite close to Snapper's hypothetical mechanical medic. While the legal and moral quandries surrounding these devices remain largely unaddressed, the existence of these devices moves Snapper's proposal closer toward practical reality. One such device is the computerized automated defibrillator, which has received widespread attention from the medical community.

Each year, as many as 400,000 people in the United States die from ventricular fibrillation, a category of heart attacks in which a victim's heart beat becomes erratic and ineffectual. Many lives are saved by emergency room specialists who use a device called a "defibrillator" to administer electric shocks to the heart. Correctly "jump starting" the heart with an appropriate dosage of electrical shock restores its normal rhythm. Time is the critical factor affecting the success of defibrillation.[86] Therefore, the task of administering this delicate medical treatment has not been confined to physicians, as are most other intricate medical procedures, but may also be performed by emergency medical technicians and skilled laymen such as policemen and firefighters.

Recently, autonomous computerized systems have been adapted to provide this lifesaving therapy. Computerized automated defibrillators designed for unskilled operators were developed to make immediate defibrillation practicable in emergency situations in which persons trained to respond to cardiac arrest are not instantly available.[87] The lifesaving potential of these devices has been compared to such medical break-

---

85. Snapper, *supra* note 77, at 295.

86. This has led to a campaign to encourage women over the age of fifty, who are seen as those most likely to witness a cardiac arrest, to telephone the emergency medical service prior to beginning CPR in an endeavor to reduce the average time from a patient's collapse to EMS notification. *See* Kenneth R. Stults, *Phone First!—Making Rapid Defibrillation Programs More Rapid*, J. EMERGENCY MED. SERVICE, Sept. 1987, at 28.

87. Cummins, *infra* note 94, at 1246.

throughs as vaccines and antibiotics.[88]

This machinery is engineered with various safety features that prevent shocking a heart that does not have an abnormal rhythm.[89] Two pads placed on a patient's chest enable the computerized apparatus to monitor the electric activity of the heart. The patient's medical condition is evaluated by the automatic mechanism, and if the mechanism decides defibrillation is necessary, it administers the precise number and strength of electric shocks that it judges necessary for proper treatment.[90] Such decisions previously required the skills of a physician or a trained emergency medical technician. One randomized clinical study, which compared the effectiveness of automated defibrillators to conventional manual defibrillators utilized by emergency medical technicians, indicates that fully automated equipment performs as well as conventional manual devices.[91] In light of the evidence showing the overwhelming utility of such devices, the Food & Drug Administration approved the use of

---

88. *See generally* Joseph J. Bocka, *Automatic External Defibrillation*, 18 ANNALS EMERGENCY MED. 1264, 1267 (1989).

89. This description of the way such units operate is based on sales brochures available from Laerdal Medical Corp., One Labriola Court, Armonk, N.Y. 10504, which claims a 70% share of the sales of automated defibrillators over the past two years.

90. A designer of this technology described its development as follows:

> A decision algorithm was developed for a semiautomatic defibrillator. The function of the algorithm is to evaluate the ECG of a patient and determine whether a defibrillation shock should be delivered. The development process included establishment of defibrillation criteria, creation of ECG databases, algorithm design, development of test protocols, and clinical testing. The result was an algorithm with sensitivity and specificity sufficiently accurate to allow a defibrillation shock to be delivered safely outside the hospital.

D.C. Edwards, *The Development of a Decision Algorithm for a Semiautomatic Defibrillator*, 18 ANNALS EMERGENCY MED. 1276 (1989).

91. Richard O. Cummins et al., *Automatic External Defibrillators Used by Emergency Medical Technicians: A Controlled Clinical Trial*, 257 JAMA 1605 (1987). This study concluded that the automated defibrillators have advantages over conventional devices in training, skill retention, and speed of operation. With respect to performance of the two types of mechanisms, it was reported: "We observed no significant differences between the capability of Automatic External Defibrillators and trained Emergency Medical Technicians to identify and shock clinically significant ventricular fibrillation correctly." *Id.* at 1609. Not all recent reports are as encouraging. Dr. Ian Stiell of Ottawa's Civic and General Hospitals has argued that merely making this automatic technology available to emergency medical technicians does not increase the survival rate. On the basis of his three-year study of automatic heart defibrillators in ambulances, he concluded that the introduction of automatic defibrillators did not improve the survival rate. In his view, attendants need full paramedic training in order to increase the odds of saving heart attack victims. *See* Jochen Kessel, *Ambulance Attendants Have Been Using Heart-Starting Equipment to Save More Lives, But a Study Shows It Isn't Enough of a Shock to the Heart*, OTTAWA CITIZEN, Oct. 22, 1991, at B1.

completely automatic defibrillators designed for home use.[92]

The medical community is already grappling with the dilemma of where to place ultimate responsibility for misadventure in circumstan-ces where medical treatment is provided by computerized equipment.[93] Exigencies that require expeditious treatment of dangerously ill heart attack victims create a potential for mistreatment. There is growing recognition in the medical community that displacement of human medical providers by automated computerized apparatus will have important repercussions on the present system of liability for medical malpractice.[94] Clarification of these conundrums requires that what constitutes "mistake," "misjudgment," or "negligence" in an automated context be unraveled. There is always the implication that a decision resulting in patient injury has been bungled, but even sensible procedures appropriately applied sometimes produce deplorable outcomes. If the human or machine acted responsibly under the circumstances, we should consider them both equally "blameless" and conclude that no "mistake" had been made.

Indisputably, the human medical provider is exposed to liability for consequences of a mishap in a far greater measure than is a similarly

---

92. *See* Paul SerVaas, *How To Jump-Start Your Husband's Heart*, SATURDAY EVENING POST, Apr. 1988, at 60. Research findings supporting widespread use of automatic defibrillators are presented in W. Douglas Weaver et al., *Use of the Automatic External Defibrillator in the Management of Out-of-Hospital Cardiac Arrest*, 319 NEW ENG. J. MED. 661, 661-66 (1988). This study demonstrated that use of fully automatic defibrillators by firefighters in Seattle on 1287 cardiac arrest victims increased the rate of survival to 30% from 19%. *See generally* Doug Podolsky, *Keeping the Bed Alive*, U.S. NEWS & WORLD REP., July 22, 1991, at 54 (discussing automated defribrillators).

93. *See* Richard O. Cummins et al., *Automatic External Defibrillators: Clinical Issues for Cardiology*, 73 CIRCULATION 381, 384 (1986).

94. *See* Richard O. Cummins et al., *Encouraging Early Defibrillation: The American Heart Association and Automated External Defibrillators*, 19 ANNALS EMERGENCY MED. 1245 (1990). The author admonished:

> A person in cardiac arrest from ventricular fibrillation must be defibrillated as quickly as possible to have any chance at successful resuscitation. This treatment is standard of care in the legal sense that all reasonable and prudent medical practitioners agree that early defibrillation is the definitive intervention for ventricular fibrillation. Inability to provide rapid defibrillation for a patient in ventricular fibrillation in medical settings in which defibrillators are available would be a breach of standard of care. Such a breach of commonly accepted care when a duty to respond exists is the first requirement for malpractice litigation. Does this mean that an EMS system is vulnerable to malpractice allegations if it does not provide a prehospital, nonparamedic-based defibrillation program? The answer to this question is unclear.

Id. at 1247.

situated mechanized medical provider.[95] We demand a more insightful response from human physicians and emergency medical technicians than from computerized systems. We count on the human medical providers to bring broad experience to bear in exercising their judgment and to consider an extensive number of variables before determining what response to an emergency is medically appropriate. Our expectations of a "mechanical medic" are of necessity more circumscribed. We do not expect a computerized medical device to consider more than the limited data it has been designed to gather.[96] While automatic computerized defibrillators are competent to resuscitate victims of cardiac arrest who would otherwise perish, the impersonal acumen of these mechanical medics is demonstrated in a stereotypical way. The heroic action by which they accomplish resuscitation of victims is achieved in a routine and perfunctory manner as they apply and act on algebraic formulas to evaluate aspects of the patient's symptoms or other indications.[97]

---

95. In evaluating the standard of behavior that the law demands of the human physician or emergency medical technician, we look to the model of what an ordinary prudent practitioner would have done under similar circumstances. When we deal instead with a machine we must ask what an ordinary prudent machine would have done in these circumstances. But what is an ordinary prudent machine, and what should it do? Reasonable men at the close of the twentieth century must seek out and make use of the best available technology. Today, the state of technology is but one of the circumstances in the life of the reasonable man that will be taken into consideration as the law assesses liability. *See* The T. J. Hooper, 60 F.2d 737 (1932) (ship held to be unseaworthy for failure to use the best available technology, despite the fact that the ship was in compliance with industry custom and practice).

96. *See* W. Douglas Weaver et al., *Use of the Automatic External Defibrillator in the Management of Out-of-Hospital Cardiac Arrest*, 319 NEW. ENG. J. MED. 661, 665 (1988). It has been noted that early prototype automated defibrillators often failed to recognize many cases of ventricular fibrillation. Indeed, it is pointed out by Cummins that although the automated defibrillators utilized by trained Emergency Medical Technicians were as effective as conventional manual systems, nevertheless "an A[utomatic] E[xternal] D[efibrillator] will not, with present rhythm analysis systems, identify [certain types of erratic heart activity]. In comparison, E[mergency] M[edical] T[echnicians], when so instructed can do so." Richard O. Cummins et al., *Automatic External Defibrillators Used by Emergency Medical Technicians: A Controlled Clinical Trial*, 257 JAMA 1605, 1609 (1987).

97. Brian Smith of Xerox PARC explained:

> All expert systems are based on models . . . .
>
> . . . .
>
> For example, a medical expert system designed to administer drugs might model drug absorption in terms of a scalar quantity proportional to the square of a patient's height, or proportional to the weight (neither model, of course, would be expected to be entirely accurate).
>
> When expert systems are actually deployed, however, they interact with the world itself, not with models. For example, when drugs are actually administered, or when offices are actually equipped with expert systems intended to work alongside people, we have full, thick situations to deal with,

Ironically, medical researchers have suggested that it is precisely because mechanical medics perform tasks in this perfunctory and stereotypical way that the medical community may favor medical intervention by automated intermediaries over equivalent manual procedures undertaken by nonprofessionals.[98] What makes the evolution of automatic defibrillators over the past few years so telling is not only that this technology is a prototype for more refined computerized stand-ins, but also that the technology epitomizes an actualization of the autonomous intelligent artifacts we have been considering. If John Snapper's conjectural "mechanical medic" is becoming an everyday appliance, the question of legal responsibility is no longer merely an esoteric issue of academic interest.

B. Risk-Utility Analysis

An issue that must be confronted in the assimilation of automated facilities into society is the extent to which human traits, such as judgment and common sense, must be incorporated into the design of an

---

> of at least potentially arbitrary complexity. Furthermore, the success of expert systems ultimately depends on their ability to deal with these rich, embedded situations. Their success, in other words, isn't exhausted by their ability to deal appropriately with the model used in their construction, or ended in their knowledge bases.
>
> . . . .
>
> . . . We have virtually no techniques, on the other hand, with which to study the latter relationship, between model and world. We are largely unable, therefore, to assess the appropriateness of models, or to predict when models will fail. All that we do when we prove a program "correct" is to prove it will behave as specified with respect to a model. It would be something quite else—something we don't know how to do—to prove that a system will in fact do the "correct" thing once embedded into a real situation.

Brian Smith, *Models in Expert Systems*, PROC. NINTH INT'L JOINT CONF. ON ARTIFICIAL INTELLIGENCE 1308 (1985).

98. Richard O. Cummins et al., *Automatic External Defibrillators Used by Emergency Medical Technicians: A Controlled Clinical Trial*, 257 JAMA 1605 (1987). Cummins contends:

> The medical decisions of rhythm diagnosis and countershock delivery are removed from the prehospital rescuers and placed with a device whose performance at rhythm identification is already known. This placement of medical decision making with an automatic device may make the concept of defibrillation by E[mergency] M[edical] T[echnicians] more acceptable to both physicians and prehospital personnel.

Id. at 1609.

automated process. Legal constraints on the process of mechanization exist to protect the public from dangers generated by automated systems. We have seen that inappropriate transmutations of human functions to machines have sometimes been expressly prohibited by legislators or circuitously interdicted by courts through imposition of an onerous measure of tort liability. Oversight by legislators, administrators, and courts regulates interchange between humans and mechanical surrogates, restricting this delegation when machines are not sufficiently competent to perform a given task.

In some instances, however, the decisionmaking processes programmed on a computer might well outperform some competent human practitioners, though failing to reach the standard of the best. Cases where injuries that result from mechanized treatment would not have occurred had a superior human practitioner been present will likely be decided within a risk-utility framework. Computerized automated defibrillators were developed to stand in for physicians because time is so critical to the success of defibrillation. The computerized equipment makes immediate defibrillation practicable in emergency situations in which even lay persons trained to respond to cardiac arrest are not instantly available. The fact that early prototype automated defibrillators often failed to recognize many cases of ventricular fibrillation[99] was considered and weighed in light of the exigencies of providing the instantaneous medical countermeasures indispensable to restoring a victim's vitality.

Application of risk-utility analysis to cases involving the automation of medical treatment does not occasion any departure from commonplace fairness-based conceptions. The analysis is logically identical to that performed when deciding where to locate a potentially dangerous facility. Once the decisionmaker has determined that the dangerous facility must constructed, the best location is the one where the least number of people are endangered. While this decsion is made mindful of the fact that some people will be exposed to harm, it is presumably a responsible decision.

Human intellect and standard computers are potent intellective resources, each adapted to different assignments. The mechanism of common computers significantly outperforms mentality in the sheer speed with which mathematical and logical calculations are performed. Computerized equipment is, more often than not, bungling and ineffectual

---

99. *See* W.D. Weaver et al., *Use of the Automatic External Defibrillator in the Management of Out-of-Hospital Cardiac Arrest*, 319 NEW ENG. J. MED. 661, 665 (1988).

at carrying out routine common sense tasks. The limitations of traditionally designed computer systems have bedeviled attempts to devise universal robots, and account in part for the sluggishness with which advances in artificial intelligence are being achieved. Current intelligent artifacts and computerized systems are able to function effectively only in particular ambiences and are suitable only for unique transactions for which appropriate behaviors may be pre-programmed. Automated defibrillator technology confronts the law with the challenge of resolving a number of perplexities spawned by automated facilities, such as whether a human actor ultimately must always be in charge[100] and whether computerized systems should be permitted to make decisions that humans cannot override.[101] Such unorthodox questions remain unanswered, and to this point essentially unasked, principally because courts have not yet had many occasions to examine them. In the brave new world circumscribed by automated facilities and processes, judges as well as engineers will participate in making pivotal decisions about the design specifications of emerging technology.

---

100. Where an erroneous computer printout indicated that an automobile loan was in default, and the creditor preferred relying on the computer-generated record rather than exercising independent judgment and considering other evidence showing that the debt had been paid, courts have been outraged to the point of awarding punitive damages. In Ford Motor Credit Co. v. Swarens, 447 S.W.2d 53, 57 (Ky. 1969), the court asserted a standard under which "the law must require that men in the use of computerized data regard those with whom they are dealing as more important than a perforation on a card." *See also* Ford Motor Credit Co. v. Hitchcock, 158 S.E.2d 468 (Ga. Ct. App. 1967); Price v. Ford Motor Credit Co., 530 S.W.2d 249 (Mo. Ct. App. 1975).

101. James Moor convincingly demonstrated that the sentiment that "computers should never make any decisions which humans cannot override," which he characterized as a "dubious maxim," is flawed. *See* Moor, *supra* note 39, at 226. There are circumstances, he argues, in which it would be morally preferable that humans not be authorized to override computer decisions. If we were to find that operating automobiles automatically by computer would bring about a substantial reduction in automobile accidents and casualties, and we were to find as well that allowing humans to override computer driving decisions led to an escalation in the accident rate, those facts would give rise to persuasive moral and prudential arguments in favor of leaving driving to computers and not permitting humans to override this automatic process. Moor regards computer decisionmaking instrumentally and sets out his position as follows:

> [F]or particular situations we must determine whether using computer decision makers will better promote our values and accomplish our goals . . . . Within the context of our basic goals and values (and the priorities among them) we must empirically determine not only the competence of the compu-ter decision maker but the consequences of computer decision making as well.

Id. at 227.

AFTERWORD: ADAPTING THE LAW FOR ARTIFICIAL INTELLIGENCE

Somewhere along the continuum between man and intelligent artifacts are animals such as dogs, which exhibit degrees of consciousness and are accordingly deemed "responsible" for their actions to a certain extent.[102] Is it not possible that in the future a machine will be conscious on a level similar to animals?[103] Artificial intelligence researchers have persuasively argued that heterogeneous species of consciousness exist,

---

102. Wilks, *supra* note 4, at 1279, calls to our attention the fact that:

> In English common law, at least, there is already a well established and still operative precedent for a category of entities which are neither human, nor totally without responsibility. They are animals like dogs, which certainly pass the test of having the appropriate attributions made to them, at least by a large part of the population. They are quite distinct from ferae nature like tigers: if you keep a tiger and it does any wrong, you are responsible, for they are taken to be simple machines in your keeping. With dogs the situation is more complex and normally, though inaccurately, summed up in the cliche "every dog is allowed one bite," the point being that a dog is not deemed savage simply because it bites someone once. It may, like us be acting out of character. Whereas to be a savage dog is to be a habitual biter and in particular to have a savage character known to its owner. Tigers are not to be thought of as having characters to act out of: they are just machines that bite. This notion of having a character one could act out of is tightly bound up with the notions of moral and legal responsibility and blame. Dogs are blamed and punished in analogous ways to people—in some countries both can be executed—and that is only because they share very similar (though importantly different) physiological structures. The problem with machines and their programs, even if we were to squeeze them into the same category as dogs, would be how to blame and punish them.).

103. Samuel Butler forewarned:

> [T]here is no security . . . against the ultimate development of mechanical consciousness in the fact of machines possessing little consciousness now. A mollusc has not much consciousness. Reflect on the extraordinary advance which machines have made during the last few hundred years, and note how slowly the animal and vegetable kingdoms are advancing. The more highly organized machines are creatures not so much of yesterday, as of the last five minutes, so to speak, in comparison with past time. Assume for the sake of argument that conscious beings have existed for some twenty million years: see what strides machines have made in the last thousand! May not the world last twenty million years longer? If so, what will they not in the end become? Is it not safer to nip the mischief in the bud and to forbid them further progress? But who can say that a vapor engine has not a kind of consciousness? Where does consciousness begin and where end? Who can draw the line? Who can draw any line?

SAMUEL BUTLER, EREWHON 233-34 (2d ed. 1917).

ranging from the profound consciousness of a human being to the trivial, indeed inconsequential, consciousness of unattended machines.[104] Artificial intelligence seeks techniques that permit computers to accomplish tasks that, when performed by humans, are thought to require intelligence.[105] Paul Armor proclaims that the goal of artificial intelligence is "to push machine behavior out into this continuum."[106]

No one knows how behavioral adaptability in a machine will be achieved, but artificial intelligence researchers believe that the most promising prospect is the simulation of biological learning processes. Human behavior, in contradistinction to the actions of computerized equipment, is learned rather than pre-programmed. Humans remain in a lengthy state of immaturity, progressing from witless babies to mischievous toddlers. The illusive and seemingly aimless play of juveniles actually develops cognitive skills that enable them to perceive and interpret their environment. This enlightenment by operant conditioning produces flexible individuals with prodigious behavioral adaptability and resiliency.

To fashion a computerized system or robot capable of operating in a universal or generic setting will likely require the duplication or simulation in machines of the methods by which humans learn. Programs already exist that simulate nondeterministic automatons capable of evolution,[107] and the next generation of supercomputers will arguably enable

---

104. *See, e.g.*, J. CULBERTSON, THE MINDS OF ROBOTS 78 (1963); Paul Armor, *Attitudes Toward Intelligent Machines*, *in* COMPUTERS & THOUGHT 391-92 (Edward A. Feigenbaum & Julain Feldman eds. 1963); *see also* Terell Ward Bynum, *Artificial Intelligence, Biology and Intentional States*, 16 METAPHILOSOPHY 355 (1985). Bynum states:

> Even a little thought about consciousness, however, quickly reveals that consciousness is a matter of degree. There is a whole range of states, for example, between someone in a deep coma, who is not conscious at all, and someone fully conscious—there are various levels of "semi-consciousness," "faintness," "fogginess," and so on. It is true that relatively simple robots—like Chesster—which we are able to build today do not appear to be "fully conscious" like a person. But are we sure that machines cannot have a low level, or some kind of consciousness?

Id. at 365.

105. This definition was adapted from Marvin Minsky, who described artificial intelligence as "the science of making machines do things that would require intelligence if done by man." Marvin L. Minsky, *Introduction* to SEMANTIC INFORMATION PROCESSING at v. (Marvin L. Minsky ed., 1968).

106. Armor, *supra* note 104, at 391-92.

107. Professor Thomas Ray, a plant biologist at the University of Delaware, recently created a computer program, "Tierra," with simulated organisms that transmit "genetic

artificial intelligence scientists to create computer models that possess the intelligence of a rodent within a decade.[108] "The dream, of course, is to evolve programs that do the things we want. We could reward or punish them to solve the problems we set."[109] Advances in technology might ultimately produce sociable machines fitted not merely to processing and reacting, but manifesting anthropomorphic properties and suited to comporting themselves with common sense.

Professor Ray's work suggests that it may eventually be possible to fabricate an "intentional" apparatus capable of spontaneous learning without a program that constrains it to adjust its behavior in stereotypical ways.[110] Once the creator endows it with an architecture and a learning algorithm, the intentional automaton will learn from experience by adapting its behavior patterns in response to newly encountered circumstances. Because these systems are nondeterministic, the artifact's behavior would be contingent on the chance events to which it was exposed. When such automatons develop to the point at which they can function in a complex environment, they will have become perceptive and introspective agents capable of decisionmaking in situations the programmer did not anticipate.

An artifact that becomes autonomous of its programmer is jurisprudentially exciting. An apparatus that learns to make appropriate decisions in the process of adapting to its surroundings may, as Professor Bechtel argues, properly be said to have selected among the alternative choices on the basis of its own deep-seated and indigenous beliefs and desires.[111] The decisions of such artifacts could be characterized as inborn, intrinsic, and voluntary in the sense of being free of extrinsic coercion. Thus, when the apparatus evolves to a point where it is integrated into its environment, it may be considered morally accountable. Anticipated intelligent artifacts are morally responsible in the sense that had they evolved otherwise, they would presumably have behaved

---

mutations" to descendant organisms in a manner mimicking biological evolution. Thomas S. Ray, *Evolution and Optimization of Digital Organisms*, in SCIENTIFIC EXCELLENCE IN SUPERCOMPUTING: THE IBM 1990 PRIZE PAPERS (Keith R. Billingsley, Ed Derohanes, Hilton Brown, III eds., 1991). For a popular description of Professor Ray's work, see Malcome W. Browne, *Lively Computer Creation Blurs Definition of Life; Software Forms, Obeying Darwin's Rules, Vie to Avoid the "Reaper,"* N Y. TIMES, Aug. 27, 1991, at B5.

108. Browne, *supra* note 107, at B5.

109. Mike Edelhart, *The Cradle of Artificial Life*, PC COMPUTING, Feb. 1991, at 152, 154 (quoting Danny Hillis, founder of Thinking Machines Corp., Cambridge, MA).

110. Ray, *supra* note 107.

111. Bechtel, *supra* note 17, at 304-05.

differently.[112]

Time will reveal whether life-like computerized systems will obtain recognition as consummate legal persons enjoying "legal rights."[113] Ultimately, this unfolding technology may yield self-regulating automatons with an anthropological temperament and perhaps even a "synthetic" consciousness. How might humankind relate to artifacts that exhibit simulated emotions such as sympathy, disappointment, or even pain? What pain, for example, "is" must be determined on the basis of our understanding of the function pain serves, regardless of the underlying physical process. If computers eventually comport themselves with true understanding and common sense and manifest some form of consciousness, the discontinuity between man and machine[114] will have become so tenuous that society may have no choice but to accept machines as "legal persons" with rights as well as duties.[115]

The development of conscious machines is unlikely to be achieved in

---

112. Snapper's position, by way of comparison, is that "[i]gnoring science fiction examples, a computer cannot be anguished, and we should not feel pity for a computer or pardon a computer. Thus a central aspect of our discussion of human action and human responsibility does not apply to computer decisions." Snapper, *supra* note 77, at 294. "One might question what it would mean to forgive a computer," Bechtel responded. "I would suggest one thing it might mean is that we would not localize the explanation for the malfunction in the way the computer had adapted to its environment, but, perhaps, in the unusual character of the circumstances or in deficiencies in the environment in which it learned its typical response patterns." Bechtel, *supra* note 17, at 305.

113. Willick, *supra* note 25, at 1271.

114. Whether an entity is man or machine is to some a matter of attitude. H. R. Halderman gave the following account of his interface with President Nixon:

> Nixon viewed Klein and Finch in human terms, as people, which meant he would have had trouble dealing with them on an official basis. He didn't see me as a person or even, I believe, as a human being. I was a machine. A robot. Shortly after it came out I saw the movie Star Wars: there is a robot, a metal machine clanking along doing what it's told by a computer-like mind. From Nixon's viewpoint, that's what I was. And I was a good machine. I was efficient, I didn't require a lot of 'oiling'—and he wasn't good at 'oiling,' or what LBJ called 'schmoozing.'

HARRY R. HALDERMAN & JOSEPH DIMONA, THE ENDS OF POWER 74 (1978). When the Fujitsu Fanuc robot-making factory in Japan began utilizing robots to make other robots, the union's leader demanded, "We want to see robots join the union." Fujitsu, sympathetic to the union's plight, offered to pay union dues on behalf of its robots, but the Japanese Labor Ministry ruled out this approach and proclaimed, "Robots cannot join a union like human workers." Leonard Silk, *Economic Scene: Strange New Robotic World*, N.Y. TIMES, May 4, 1992, at D2.

115. This view is further developed in Marshal S. Willick, *Artificial Intelligence: Some Legal Approaches and Implications*, 4 AI MAG. 5 (1983). *See also* Dan Lloyd, *Frankenstein's Children: Artificial Intelligence and Human Value*, 16 METAPHILOSOPHY 307, 308 (1985).

the near future or even in our lifetime. However, the technological impetus that has already brought "mechanical medics" into being will undoubtedly give birth to other autonomous machinery discharging essential social obligations. Fortunately, the slow pace of artificial intelligence leaves adequate time for us to anticipate and prepare for the tremendous social transformations it may generate as the attributes separating humans from their automated offspring become ever less distinct.