Computerization, discretion, freedom

Sergey Bratus, Anna Shubina

January 5, 2013

The idea that widespread deployment of computer technology has the power to improve any given area of human endeavor appears to have taken deep roots in the public mind. Computer technology, once a concerted effort to embed it in an industry has been made, is assumed to be able to improve efficiency, lower costs, reduce waste, and so on.

Yet computerization changes every area of human activity it touches; the changes are both subtle and profound, as they modify human behaviors related to discretion, perceived competence, and, ultimately, moral choice. We discuss the reasons why computers must be watched very carefully by anyone concerned about freedom.

Computer Magic in the Public Mind

Proposals that rely on computing as a centerpiece of their promise do not tend to encounter much public scepticism – smart grid and the electronic health record infrastructure initiatives being the latest examples (compare these with proposals related to the good old "rocket science", which encounters much more sceptical reactions).

As it stands, the warnings of security professionals that our current commodity computer technology might not be ready for such massive endeavors are largely ignored and hardly ever reach the general public, much less receive its attention. Even cautionary stories of computer technology enabling dystopias ("Minority Report" and "Gattaca" among others) make the implicit assumption of enough "smarts" and reliability on the technology's part to sustain the prolonged or indefinite operation of the distopia.

It appears that a message to the contrary does not resonate with the public mind; trustworthiness of complex computer technology – if not now, then in the nearest future – is assumed to be a trivial matter, a forgone conclusion. Sceptics have truly become "vox clamantis in deserto".

The Magic, Old and New

> "One of the chief duties of the mathematician in acting as an adviser to scientists ... is to discourage them from expecting too much of

mathematics."

N. Wiener, "I Am a Mathematician, The Later Life of a Prodigy," MIT Press, 1964

Norbert Wiener wrote this at the time when mathematical modeling and its applications were seen as magic that, when applied to almost any area of human activity by competent enough practitioners, would provide instant insights, improvements, or both.

This perception was formed by the powerful and obvious successes of operations research and similar applied disciplines. In the popular mind, it was also both part and justification of the vision of the scientifically organized society to come, in which wrongs would be equitably righted and burdens equitably distributed through science – and what better means to do so than mathematics? So great was this trust in mathematics, that even obviously unrealistic claims of unprecedented and unstoppable superiority of planned economies that claimed their planning to be based on mathematical planning were given wide credence, including by famous economists such as Paul Samuelson [1].

Wiener, the father of Cybernetics, was among those responsible for its broadly perceived promise. We would do well to heed his warning and apply it to the apparent and widely acclaimed magic of *our* days: computerization.

1 "Computers Can Improve Anything"

Belief in great powers of a formal device or method must first take root among enough engineers and others having to do with technology; only then it can be taken up by the general public. Computers have garnered this belief early on, so that already in 1979 DeMillo, Lipton, and Perlis repeated Wiener's warning with regard to computing.

More importantly, these authors offered a convincing explanation of the technologists' growing belief in the power of computerized computation:

> Since "symbols" can be written and moved about with negligible expenditure of energy, it is tempting to leap to the conclusion that *anything* is possible in the symbolic realm. This is the lesson of computability theory (viz., solvable problems vs. unsolvable problems), and also the lesson of complexity theory (viz., solvable problems vs. feasibly solvable problems): physics does not suddenly break down at this level of human activity. It is no more feasible to construct *symbolic* structures without using energy than it is possible to construct *material* structures for free. But if symbols and material objects are to be identified in this way, then we should perhaps pay special attention to the way material artifacts are *engineered*, since we might expect that, in principle, the same limitations apply.
>
> — *Richard A. DeMillo, Richard J. Lipton, and Alan J. Perlis [2]*

It seems that the unrealistic expectations that Wiener warned about with regard to mathematical methods were revived by the promise of scaling up such symbolic computations with computers – as if the only limitations that kept the original magic from delivering were those of scale, not of natural laws.

In other words, using the magic of computers to get around the hard problems of the world is fraught with the same danger as calling on the magic of mathematical modeling had been in the 1960s: when limitations arising from natural laws get in the way, mathematics or computer science can, in time, discover, distill, and describe these limitations in their own formal ways, but not overcome them.

The hardness dictated by natural laws does not go away. It merely manifests itself in a different pattern of failures when people attempt to overcome it by piling on more effort. However, given the penetration of computers into all walks of life, this time the failures may have a bigger impact than just technical disappointments or disasters.

2 The Failures and Fallacies of Computerization

> "However, the assumption that the aforementioned benefits [safer care, reductions in administrative costs, improved clinical performance, better communication between patients and caregivers] are highly correlated with health IT has not been adequately tested and there are some indications that the features needed to acquire one benefit may actually frustrate efforts to achieve another. In particular, there is a growing concern that health IT designs that maximize the potential for administrative and economic benefit may be creating new paths to failure."
>
> *The New Institute of Medicine Report on Health IT and Patient Safety*citespectrum,

It seems to me that the conclusions reached by this committee regarding computerization of health IT should have been a common sense *starting point* when an application of computer technology is being considered, for the following reason:

"Computerizing" a complex human activity means replacing human effort in some tasks by programs (in software or hardware), and altering other tasks so that they can feed into these programs.[1] The stated goal is usually better efficiency or precision. However, if the match between the replacement's logic, the other task changes necessary to accommodate it, and the actual work flow is not perfect, human time and effort will be spent to account for the differences – likely interfering with overall results.

Such replacement is particularly fraught with danger whenever desired performance of the task involves human discretion. There are many reasons for

[1] This applies to all human activities, from accounting to assembly by robotic arms: activities surrounding the automated task must also change.

this. For example, the newly "computerized" task must work not just in its most common circumstances, but also in exceptional circumstances that require the performer to do something differently. Discretion can be "shifted" to neighboring tasks as a part of their necessary adjustment to automation, but it cannot be eliminated if the activity to be automated is sufficiently complex.

Obviously, the nature and variation of the task must be well-understood and well-described to programmers before replacement is attempted. However, the amount of information required and the effort of gathering and organizing this information may be non-trivial and costly – potentially more costly than the actual improvement from computerization.

Replacement can have unexpected effects on other related tasks, and especially on the human involved in the computerized task when human involvement is still required. Should the program logic fail to exactly match what needs to be done, the human will find himself working around or "fighting" the system, spending effort and, possibly, compromising his performance on other tasks.

In particular, the human will have to mentally compute the difference between what he would normally do and what the program will do for him and to find a way to get that difference done in the new environment, which does not account for the necessity to get it done.

The inability of human-replacing computers (or "robots") to make proper exceptions has long been a theme of science fiction (and likely helped create the phenomenon of science fiction as such). The inspiration behind these dystopian visions, however, is not merely fear of the unknown or of the technological future sweeping in too fast. It was a natural extension of the mundane experience of a rule-driven computation in which human discretion is limited by circumstances or by design: the bureaucratic process.

3 The Perfect Bureaucrat

The computer is nothing but a fast, implacable, zero-discretion bureaucrat. In fact, this is precisely how Richard Feynman chose to describe it in his "Lectures on Computation". The only discretion and context awareness in the system can only come from human overrides.

As computer security professionals – that is, specialists in computer-related *failures* – it is our duty to explain the drawbacks of computerized decision making.

- Discussion focuses on "computers" (users imagine PCs, PDAs, and other cool and impressive devices), whereas the real issue is "computation" (even in security, the focus is finally shifting to "malicious computation" without regard to the amount of foreign code introduced, a revolutionary development of early hacker research).

- Once we focus on computation rather than "computer", we realize that it predates the computer as such – as Ross Andersson argues, massive computations in the sense of rule-guided, multi-step procedures, with almost

no discretion on the part of low-ranking participants were actually being performed by roomfuls of clerks, bureaucratic hierarchies (in China, for thousands of years), and even armies.

- A significant part of the public is familiar with bureaucratic procedure and understands the drawbacks created by strictly following rules regardless of context and zero-discretion policies.

4 Discretion

I am indebted to Eleanor Saitta for the following observations: *Many organizations tend to view discretion as a bug, rather than an inherent part of their production process.*

Why do organizations underestimate the amount of discretion involved in their activities ("computations")? This is an organizational counterpart to overestimating one's competence – organizations underestimate the complexity of what they do and overestimate their rules and procedures.

In order to analyze shifts of discretion, organizations must first understand how much they depend on discretion and what role it plays in the overall computation. When they lack such understanding, they tend to overestimate the returns of computerization.

5 Freedom

Rules you cannot bend, when you clearly see you should, create helplessness. There is nothing more harmful to freedom than helplessness. But computers are great (and guiltless) enforcers of rules, and therefore great tools to teach helplessness. Therefore they must be watched very carefully by anyone concerned about freedom.

Behind a great many ideological rifts is the difference of belief in the degree of other humans' competence, and, therefore, in the value of discretion. The same is true for managerial styles.

One recurrent attitude is that [other or all] humans are better off with a stricter set of rules to define their daily activities. The damage that unbendable, zero-discretion rules clearly do in some cases is on average believed to be compensated by the good they create across the society, which otherwise would suffer more damage from poor application of discretion and lack of control. In short, the underlying conviction is that humans, on average, cannot be trusted with discretion, for their own good and efficiency, and the more of their material activities could be described as strict procedures to follow, the better.

From this position, introduction of additional (expert-designed) rules and controls to see that these are obeyed is by default seen as beneficial. It is seen as safer to err on the side of adopting technologies that introduce new possibilities for rule-making and enforcement: book-keeping, auditing, bureaucratic

procedure, registration, certification. Computerization is merely a recent addition to this list.

However, to fully take advantage of these technological improvements, humans must also learn [or be trained, or obligated] to delegate their decisions to certified procedures and professionals trained to apply the rules. Depending on the rule system and the activity, exclusive delegation may even be *required* for the stability of the rule-based mechanism, or at least the perceived fairness of its outcomes.

The best kind of delegation, is, of course, voluntary, as it does not require dealing with the negative consequences of perceived coercion. For this, the feeling of helplessness is very handy for those whose modus operandi for improving the world is to reduce human discretion.

6 Challenges the future holds

The accumulation of rules in rule-based artificial intelligence systems is known to practitioners to make adjusting or retraining them barely tractable. Often, new rules are introduced to compensate for effects of previously introduced ones or their combinations; the overall effect is more complex interactions and reduced tractability. Similar effects plague organizational policies - rules and checks that may have served to compensate for a past mistake or simply to make someone's job easier pile up over time – and may outlive the circumstances that called them into being. Accordingly, organizations accumulate informal lore on rules that may be bent and forms that may be skipped.

However, computerization creates environments where rules can no longer be informally bent or bypassed at the worker's discretion; what's written in code may as well be written in stone. Worse, the apparent ease of manipulating computerized rules may encourage more eager rule-making – but in complex systems such ease is a dangerous illusion, due to the above mentioned tractability problems.

Indeed, computerized environments are seeing a different culture than the lore of allowed discretion – that of "fighting the system", or "tricking the system" in old bureaucracies, notorious for their Kafkaesque intractability.

This presents a new and yet not well-understood challenge to designers of systems intended to "computerize" human activities to previously impossible degrees and levels of detail: they must understand the role of discretion in these activities and make sure to accommodate it. Such accommodation will certainly require new methodology for understanding both processes and the effects of their uniformization, intentional or incidental to computerization.

It is debatable to which extent a system's designers are liable for its misuse due to misunderstanding of its inherent limitations and overenthusiastic misapplication. Yet it is clear that such misuse will eventually reflect on the designers and technologies, and therefore designers must think, e.g., on how to avoid empowering compulsive and excessive rule-making – just as they currently think of preventing other kind of user error. This, in turn, will require studies of sys-

tematic organizational errors due to the advanced technology-induced illusion of control and overestimation of the rule-makers' competence.

The hardest of all will be the challenge of predicting and offsetting broader societal impacts of ubiquitous limitations of discretion. We hardly have a clear idea of the limits beyond which shrinking of individuals' discretion becomes a threat to personal autonomy, turning individuals to submissive and uncritical rule-followers; however, we know that technically enforced rule-following in well-organized societies can be the slipperiest of slopes. We must therefore view those aspects of computer technologies that limit human discretion with utmost caution – and, amid enthusiasm for their obvious benefits, it is our primary professional duty to also warn the society of their less obvious dangers to personal autonomy and freedom.

References

[1] `http://freedomkeys.com/samuelson.htm`

[2] *Social Processes and Proofs of Theorems and Programs*, Richard A. DeMillo, Richard J. Lipton, and Alan J. Perlis, Yale University Technical Report TR-82, p.7, `http://www.cs.yale.edu/publications/techreports/tr82.pdf`

[3] *The New Institute of Medicine Report on Health IT and Patient Safety*, `http://spectrum.ieee.org/riskfactor/computing/it/new-institute-of-medicine-report-on-health-it-and-patient-safety`