Moral Responsibility for Computing Artifacts: Five Rules, Version 27

Preamble

As computing artifacts become increasingly complex, some have suggested that such artifacts greatly complicate issues of responsibility. In order to help deal with these complexities, we propose five rules as a normative guide for people who design, develop, deploy, evaluate or use computing artifacts. Our aim is to reaffirm the importance of moral responsibility for these artifacts, and to encourage individuals and institutions to carefully examine their own responsibilities with respect to computing artifacts. We do not claim that these rules are exhaustive; professionals, individuals and organizations may choose to take on more responsibility than we describe here.

This is a collaborative document, and it cannot include everything each of us thinks about this subject. However, each signer of this document supports what is written here.

A Working Definition of "Computing Artifacts"

An artifact is an object designed, made or shaped by humans [1]. We use "computing artifact" for any artifact that includes an executing computer program. We intend to include software applications running on a general purpose computer, programs burned into hardware and embedded in mechanical devices, robots, phones, webbots, toys, programs distributed across more than one machine, and many other configurations. We include, among other types: software that is commercial, free, open source, recreational, an academic exercise or a research tool.

A Working Definition of "Moral Responsibility"

We use "moral responsibility for computing artifacts" to indicate that people are answerable for their behavior when they produce or use computing artifacts, and that their actions reflect on their character. [2] "Moral responsibility" includes an obligation to adhere to reasonable standards of behavior, and to respect others who could be affected by the behavior. We do not address legal liability in this document.

A Working Definition of "Sociotechnical Systems"

Each computing artifact should be understood in the context of "sociotechnical systems." A sociotechnical system includes people, relationships between people, other artifacts, physical surroundings, customs, assumptions, procedures and protocols. [3]

We acknowledge the importance of sociotechnical systems to the issue of moral responsibility for computing artifacts. For example, a GPS navigator is a computing artifact, but in isolation from the satellites it uses for ascertaining location, it cannot perform its function. People, commercial enterprises, governments and artifacts were necessary to design, develop, and deploy the satellite system and the navigators. The people who make the device encourage and discourage different uses by the navigator's design. Different people who buy the navigators choose to use it in different ways. A protocol for communicating with those satellites had to be negotiated between stakeholders. The methods by which people have agreed to identify places on the earth's surface form another part of the sociotechnical system without which an automated navigator is infeasible.

The significance of sociotechnical systems should inform any discussion of moral responsibility for computing artifacts, but it complicates matters. On one hand, ignoring the sociotechnical systems in which a computing artifact is embedded is folly. On the other hand, including all relevant sociotechnical systems components in every discussion of moral responsibility involving a computing artifact will make it impractical to assign meaningful responsibility to the people most directly involved with that specific artifact. In order to negotiate this tension, we first discuss moral responsibility for computing artifacts in a more focused sense (Rules 1, 2 and 3), and then place this discussion into a broader context that explicitly includes sociotechnical systems (Rules 4 and 5).

Rule 1: **The people who design, develop, or deploy a computing artifact are morally responsible for that artifact, and for the foreseeable effects of that artifact. This responsibility is shared with other people who design, develop, deploy or knowingly use the artifact as part of a sociotechnical system.**

Rule 2: **The shared responsibility of computing artifacts is not a zero-sum game. The responsibility of an individual is not reduced simply because more people become involved in designing, developing, deploying or using the artifact. Instead, a person's responsibility includes being answerable for the behaviors of the artifact and for the artifact's effects after deployment, to the degree to which these effects are reasonably foreseeable by that person.**

When people design, develop and deploy computing artifacts, they do so consciously and intentionally. This intentionality is important when discussing moral responsibility. Rules 1 and 2 are meant to clarify the moral responsibility of the people most directly answerable for a specific computing artifact and its effects on others.

Rule 2 is meant to address the problem of many hands, a concern that when responsibility for an artifact is broadly shared among many people, it may be that no individual considers his or her own responsibility significant. [4] Rule 2 explicitly rejects that abdication of responsibility.

The words “reasonably,” “foreseeable” and “knowingly” add complexity, subtlety and ambiguity to Rules 1 and 2. This is unfortunate, as we’d like the rules to be as straightforward as possible. However, in order to develop practical, realistic rules, it is difficult to make them simpler than this.

By using the word “foreseeable,” we acknowledge that the people who design, develop, deploy and use artifacts cannot reasonably be expected to foresee all the effects of the artifacts, for all time. However, implicit in our use of this word is the expectation that people make a good faith effort to predict the uses, misuses, and effects of the deployment; and to monitor these after deployment. Willful ignorance, or cursory thought, is not sufficient to meet the ethical challenges of Rules 1 and 2.

Furthermore, if people design an artifact in such a way that it is not possible to reasonably predict its future behaviors, then they are particularly responsible for the unpredictable, and potentially harmful, results. Some machines are designed to adapt over time, to “learn” without human supervision, or to self-modify their code; we assert that the people who launch these machines have a heavier burden of responsibility than people who launch more predictable machines. We assert that a machine’s unpredictability increases people’s responsibility for anticipating problems and safeguarding against them. People who recognize their responsibilities in this way are likely to make their machines simpler and more predictable in order to make them safer and more reliable; we would welcome this outcome.

Another caveat about Rules 1 and 2 concerns a decision *not* to create or deploy a computing artifact. If it is reasonably foreseeable that the creation and deployment of an artifact is likely to have a good (or bad) effect, then a decision to *not* create and deploy that artifact has ethical significance because it reduces the good (or reduces the bad), and those who made the decision are responsible for the consequences of that decision.

***Rule 3:* People who knowingly use a particular computing artifact are morally responsible for that use.**

The word “knowingly” is problematic in Rule 3, but we think it is, on balance, appropriate. People who “use” a particular computing artifact may or may not be aware of this use. For example, the driver of a car may not have any knowledge of a computing artifact embedded in the car that records data for analysis in case of a crash. It seems to us counter-intuitive to assign moral responsibility to the driver for the use of that artifact. However, when someone knowingly and intentionally uses a particular computing artifact, that person takes on moral responsibility attached to that use. A dramatic example is when someone launches a cruise missile at an enemy target; a more mundane example is when someone searches the web for information about a prospective employee. The moral responsibility of a user includes an obligation to learn enough about the computing artifact's effect to make an informed judgment about its use for a particular application.

It is not our intent to absolve the users of computing artifacts from moral responsibility if they are *willfully* ignorant about artifacts or their effects. Rule 3 could be misinterpreted in this way. We acknowledge this problem, but we judge that the possibility of this abuse does not negate that there are practical and ethically significant differences in the way people interact with computing artifacts. For example, “users” of computing artifacts cannot be reasonably held accountable if the use is hidden from them. (The hidden nature of the artifact may be intentional or incidental.) However, people should not seek, or even allow themselves, to be ignorant about technology and its effects in order to avoid responsibility for their use of technology.

As with Rules 1 and 2, Rule 3 applies to people who consciously decide *not* to use a computing artifact. In order to place Rules 1, 2 and 3 into a broader context, we assert two more rules:

***Rule 4:* People who knowingly design, develop, deploy, or use a computing artifact can do so responsibly only when they make a reasonable effort to take into account the sociotechnical systems in which the artifact is embedded.**

Sociotechnical systems are increasingly powerful. If people thoughtlessly produce and adopt these systems, they are, in our opinion, being morally irresponsible. Ignorance is not a justification for harms associated with sociotechnical systems and the computing artifacts imbedded in those systems. Security issues that occur when computing artifacts are deployed via the Internet are an example of the interaction of an artifact and a sociotechnical system.

Rule 4 is intended to be a progressively heavy burden. It requires an honest effort to identify and understand relevant systems, commensurate with one's ability and one's depth of involvement with the artifact and system. Thus, the burden is heavier for those with more expertise and more influence over the artifact's effects and over the system’s effects. Those in design and development cannot shift their burden to the users (see Rule 2), and users cannot shift the burden to developers when users' local knowledge is critical to appropriate ethical action.

The sociotechnical systems in which an artifact would be embedded should be considered even when the decision is to not design, develop, deploy or use a computing artifact. Rule 4 expands the effect of Rules 1, 2 and 3, since people who obey Rule 4 will know more about the effects of a computing artifact they produce and/or use.

***Rule 5:* People who design, develop, deploy, promote, or evaluate a computing artifact should not explicitly or implicitly deceive users about the artifact or its foreseeable effects, or about the sociotechnical systems in which the artifact is embedded.**

Morally responsible use of computing artifacts and sociotechnical systems requires reliable information about the artifacts and systems. People who design, develop, deploy or promote a computing artifact should provide honest, reliable, and understandable information about the artifact, its effects, possible misuses, and, to the extent foreseeable, about the sociotechnical systems in which they think the artifact will be embedded. People are morally responsible for their choice of whom to inform, and for what they disclose or don’t disclose.

Computing Artifacts that are Not Exceptions to the Rules

No matter how sophisticated computing artifacts become, the rules still apply. For example, if an artifact uses a neural net, and the designers subsequently are surprised by the artifact’s effects, the rules hold. If a computing artifact is self-modifying, and eventually becomes quite different from the original artifact, the rules still hold. If a computing artifact is a distributed system or an emerging system, the rules still hold for the people associated with the pieces that are distributed, for the people associated with the organization of the overall system, and for the people responsible for the system from which the new system emerged. If a person launches a computing artifact A in order to launch a second artifact B that is designed to launch a third artifact C, then the rules apply to all three artifacts. (That is, if you are not willing to accept moral responsibility for A, B, and C, then you should not launch A.)

As mentioned above, there are responsibilities associated with *not* launching any computing artifact. However, when the predictability of an artifact’s future behavior is in serious doubt, we maintain that a precautionary principle [5] should be applied, which will require a serious effort to justify uncertain benefits in the face of costs that will be difficult to predict. We recognize that this is a heavy burden on those advocating launching such artifacts, and we contend that this is appropriate.

References

[1] Wiktionary entry for “Artifact.” URL = http://en.wiktionary.org/wiki/artifact, accessed 9 March 2010.

[2] Michael Davis, “‘Ain’t no one here but us social forces:’ Constructing the professional responsibility of engineers,” *Science and Engineering Ethics*, forthcoming.

[3] Chuck Huff. “Why a Sociotechnical System?” URL = http://computingcases.org/general_tools/sia/socio_tech_system.html, accessed 9 March 2010.

[4] Nissenbaum, H. 1994. “Computing and Accountability.” *Communications of the ACM 37*, 1 (Jan. 1994), 72-80.

[5] Som, C., Hilty, L. M. & Ruddy, T. F. (2004). “The Precautionary Principle in the Information Society.” *Human and Ecological Risk Assessment, 10* (5), 787-799.

Meta-Rules (rules about changing the text of this document):

Meta-rule 0. In this document, “we” refers to people who have signed on to this document by joining the Ad Hoc Committee on Responsible Computing. The “coordinator” is a member of the committee, and is in charge of version control. The coordinator is currently Keith W. Miller, email: miller.keith@uis.edu.

Meta-rule 1. Anyone we invite can sign on to the document by emailing the coordinator and volunteering to join the Committee.

Meta-rule 2. Anyone we invite can suggest changes in the document to the coordinator.

Meta-rule 3. Any proposed changes should be emailed to the coordinator. The coordinator emails the committee (and other interested parties) with proposed changes, edited for content and format by the coordinator. If there are no objections emailed by committee members to the coordinator within 10 days after the coordinator sends out a proposed change, the proposed change is accepted and a new version is emailed.

This version (and several others) are available at https://edocs.uis.edu/kmill2/www/TheRules/. Also listed there are the signers of this document. A short explanation of the history of the ongoing project of developing The Rules is available: https://edocs.uis.edu/kmill2/www/TheRules/history.html

To cite this document, we suggest the following: Ad Hoc Committee for Responsible Computing. Moral Responsibility for Computing Artifacts: Five Rules, Version 27. https://edocs.uis.edu/kmill2/www/TheRules/.

This version of “The Rules” was last revised on 30 October 2010.