

Autonomy Recommendation for ACM Code of Ethics and Professional Conduct

The [Association for Computing Machinery \(ACM\)](#) is dedicated to Advancing Computing as a Science & Profession. "We see a world where computing helps solve tomorrow's problems – where we use our knowledge and skills to advance the profession and make a positive impact."

The ACM Committee on Professional Ethics (COPE) is responsible for promoting ethical conduct among computing professionals, and the [ACM Code of Ethics and Professional Conduct](#) is undergoing a [comprehensive review](#). We have offered the the following [recommendation](#).

We have reviewed "ACM Code of Ethics and Professional Conduct" (Draft 3) and applaud this important guidance for computing professionals [1].

A significant area of ongoing activity is not yet addressed: remote and autonomous systems. While several aspects of the draft Code pertain, the design and deployment of such systems pose special challenges that computing professionals must specifically consider. As noted in the Turing Award Lecture of fifty years ago: ethics, professional behavior, and social responsibility cannot be separated from the diverse fields in which computer science is applied [2].

Many robotic systems are being deployed with high degrees of autonomy, long operational endurance, and independence from direct human supervision. Examples include self-driving cars, unmanned air vehicles (drones), military sentry vehicles, and many others. Whether manned, unmanned, civil or military, such systems have significant potential for applying indiscriminate lethal force at a distance. Complex situations, unforeseen interactions, and emergent behaviors often occur that are beyond the original scope or intent of designers and engineers.

Special considerations are necessary for such machines, since preprogrammed machine responses remain inadequate in isolation. Protections for human life must be considered and engineered into systems capable of prolonged operations beyond the range of direct remote control. A critical enabler is available to help: the combination of human judgement and artificial intelligence can yield more effective systems than is possible by either alone [3]. Thus sufficient human supervisory guidance, and permission checks for recognizably dangerous situations, must be available for systems that are allowed to operate autonomously. Simply put: if a human is not in charge, then no one is in charge.

Constraints on action (such as limits of authority, and conditions requiring explicit human approval) can be achieved for remote systems presenting potential hazard to life. For example, recent work has shown that human ethical considerations can be expressed using validatable syntax and logical semantics when defining executable robot missions [4]. Indeed, if ethical approaches combining machine and human capabilities can better ensure human safety, it is unethical to not consider them. Understanding such issues when engineering systems with autonomy requires the technical expertise and moral judgement of computer-science professionals.

We recommend adding a section to the ACM Code that articulates these vital concerns. Suggested draft Professional Responsibilities paragraph 2.10 follow.

=====
"Recognize potential risks associated with autonomy. Systems operating remotely or with minimal human supervision (for example, drones or driverless vehicles) may have the capacity for inflicting unintended lethal force. Safeguards, legal requirements, moral imperatives, and means for asserting direct human control must be considered, in order to avoid the potential for unintended injury or loss of life due to emergent behavior by robotic systems."
=====

Our chosen wording of "must" vice "should" is intentional, since recognizing such risks meets thresholds described in [1] and failure to consider such issues is negligent.

Ethical constraints on robot mission execution are possible today. There is no need to wait for future developments in Artificial Intelligence (AI). It is a moral imperative that ethical constraints in some form be introduced immediately into the software of all robots that are capable of inflicting unintended or deliberate harm to humans or property.

Very respectfully submitted.

Don Brutzman, Bob McGhee, Curt Blais and Duane Davis
Naval Postgraduate School (NPS), Monterey California USA

[1] Don Gotterbarn, Amy Bruckman, Catherine Flick, Keith Miller, and Marty J. Wolf, "ACM Code of Ethics: A Guide for Positive Action," *Communications of the ACM (CACM)*, vol. 61 no. 1, pp. 121-128. http://mags.acm.org/communications/january_2018/?CFID=845243269&CFTOKEN=68679870&pg=123#pg123

[2] Richard W. Hamming, "One Man's View of Computer Science," ACM Turing Award Lecture, *Journal of the ACM (JACM)*, vol. 16 no. 1, January 1969. <https://dl.acm.org/citation.cfm?id=1283923>

[3] Richard W. Hamming, *Learning to Learn: The Art of Doing Science and Engineering*, CRC Press, 1997.

[4] Don Brutzman, Curtis Blais, Duane Davis, and Robert B. McGhee, "Ethical Mission Definition and Execution for Maritime Robots under Human Supervision," *IEEE Journal of Oceanic Engineering (JOE)*, January 2018, <http://ieeexplore.ieee.org/document/8265218>