



Robert J. Sawyer is a Hugo and Nebula Award-winning science-fiction writer. E-mail: sawyer@sawyer.com

## Robot Ethics

C-3PO AND R2-D2 MAY BE TWO OF THE WORLD'S MOST FAMOUS FICTIONAL ROBOTS, but a quasi-robot named MQ-5B/C is perhaps more interesting just now. On 1 September 2007, operators used this unmanned airborne drone to locate and drop a bomb on two individuals who appeared to be planting explosives near Qayyarah, Iraq. As we make robots more intelligent and autonomous, and eventually endow them with the independent capability to kill people, surely we need to consider how to govern their behavior and how much freedom to accord them—so-called roboethics. Science fiction dealt with this prospect decades ago; governments are wrestling with it today. Why now? It's not only because robots are killing people. It's also because they have become household consumer-electronics items and because some now look and act like humans (Honda's Asimo can even dance). We have an instinctive reaction that a threshold has been crossed.

The notion of killer robots is a mainstay of science fiction; but then again, so is the idea of robots with built-in safeguards against that. In his 1942 story "Runaround," Isaac Asimov offered his now-famous Three Laws of Robotics: A robot may not injure a human being or, through inaction, allow a human being to come to harm; a robot must obey orders given to it by human beings except where such orders would conflict with the First Law; and a robot must protect its own existence as long as such protection does not conflict with the First or Second Law. Most of Asimov's stories deal with things going awry because these laws don't equip robots to tackle real-world situations. In his 1947 story "With Folded Hands," Jack Williamson had robots adhere to an even simpler directive: To serve and obey, and guard men from harm. That, too, had an unwelcome result: a totalitarian society in which robots prohibit humans from participating in almost all activities, lest one of us be injured.

Indeed, all attempts to govern complex robotic behavior with coded strictures may be misguided. Although the machines will execute whatever logic we program them with, the real-world results may not always be what we want. And yet, we seem unable to resist trying, and so governments are now drafting their versions of Asimov's and Williamson's laws. This year, South Korea's Ministry of Commerce, Industry, and Energy established a Robot Ethics Charter, which sets ethical guidelines concerning robot functions. The move anticipates a time when intelligent service robots are part of daily life. EURON (the European Robotics Research Network) also announced plans to develop guidelines for robots in five areas: safety, security, privacy, traceability, and identifiability. Japan's Ministry of Economy, Trade, and Industry has joined in too. With an aging population and robot caregivers being developed there (and elsewhere in the world), the Japanese foresee robots in many homes and have issued policies for how they should behave and be treated.

The United States has yet to jump on the roboethics bandwagon. That many U.S. robots are created for the military and designed to harm humans may be the reason. Still, it is likely that the most interesting litigation defining robot responsibilities and rights will emerge in the United States. For starters, a Michigan jury awarded the family of the first human ever killed by a robot (accidentally, in 1979) \$10 million, which was, at that time, the largest personal-injury award in the state's history.

Again, science fiction may be our guide as we sort out what laws, if any, to impose on robots and as we explore whether biological and artificial beings can share this world as equals. Isaac Asimov's 1954 novel *The Caves of Steel* describes a fully equal robotic partner of a police officer. Lester del Rey's 1938 story "Helen O'Loy" portrays what might be one viable future: a man marrying a robot woman, and living, as one day all humans and robots might, happily ever after. I, for one, look forward to that time.

— Robert J. Sawyer

10.1126/science.1151606

