

Living Safely with Robots, Beyond Asimov's Laws

June 22 2009, By Lisa Zyga



TOPIO 2.0 - TOSY Ping Pong Playing Robot version 2 at Nuremberg International Toy Fair 2009. Image: Wikimedia Commons

(PhysOrg.com) -- "In 1981, a 37-year-old factory worker named Kenji Urada entered a restricted safety zone at a Kawasaki manufacturing plant to perform some maintenance on a robot. In his haste, he failed to completely turn it off. The robot's powerful hydraulic arm pushed the engineer into some adjacent machinery, thus making Urada the first recorded victim to die at the hands of a robot."

In situations like this one, as described in a recent study published in the *International Journal of Social Robotics*, most people would not consider

the accident to be the fault of the robot. But as robots are beginning to spread from industrial environments to the real world, human safety in the presence of robots has become an important social and technological issue. Currently, countries like Japan and South Korea are preparing for the “human-robot coexistence society,” which is predicted to emerge before 2030; South Korea predicts that every home in its country will include a robot by 2020. Unlike industrial robots that toil in structured settings performing repetitive tasks, these “Next Generation Robots” will have relative autonomy, working in ambiguous human-centered environments, such as nursing homes and offices. Before hordes of these robots hit the ground running, regulators are trying to figure out how to address the safety and legal issues that are expected to occur when an entity that is definitely not human but more than machine begins to infiltrate our everyday lives.

In their study, authors Yueh-Hsuan Weng, a former staff of Taiwan’s Conscription Agency, Ministry of the Interior, and currently visiting at Yoshida, Kyoto, Japan, along with Chien-Hsun Chen and Chuen-Tsai Sun, both of the National Chiao Tung University in Hsinchu, Taiwan, have proposed a framework for a legal system focused on Next Generation Robot safety issues. Their goal is to help ensure safer robot design through “safety intelligence” and provide a method for dealing with accidents when they do inevitably occur. The authors have also analyzed Isaac Asimov’s Three Laws of Robotics, but (like most robotics specialists today) they doubt that the laws could provide an adequate foundation for ensuring that robots perform their work safely.

One guiding principle of the proposed framework is categorizing robots as “third existence” entities, since Next Generation Robots are considered to be neither living/biological (first existence) or non-living/non-biological (second existence). A third existence entity will resemble living things in appearance and behavior, but will not be self-aware. While robots are currently legally classified as second existence

(human property), the authors believe that a third existence classification would simplify dealing with accidents in terms of responsibility distribution.

One important challenge involved in integrating robots into human society deals with “open texture risk” - risk occurring from unpredictable interactions in unstructured environments. An example of open texture risk is getting robots to understand the nuances of natural (human) language. While every word in natural language has a core definition, the open texture character of language allows for interpretations that vary due to outside factors. As part of their safety intelligence concept, the authors have proposed a “legal machine language,” in which ethics are embedded into robots through code, which is designed to resolve issues associated with open texture risk - something which Asimov’s Three Laws cannot specifically address.

“During the past 2,000 years of legal history, we humans have used human legal language to communicate in legal affairs,” Weng told *PhysOrg.com*. “The rules and codes are made by natural language (for example, English, Chinese, Japanese, French, etc.). When Asimov invented the notion of the Three Laws of Robotics, it was easy for him to apply the human legal language into his sci-fi plots directly.”

As Chen added, Asimov’s Three Laws were originally made for literary purposes, but the ambiguity in the laws makes the responsibilities of robots’ developers, robots’ owners, and governments unclear.

“The legal machine language framework stands on legal and engineering perspectives of safety issues, which we face in the near future, by combining two basic ideas: ‘Code is Law’ and ‘Embedded Ethics,’” Chen said. “In this framework, the safety issues are not only based on the autonomous intelligence of robots as it is in Asimov’s Three Laws. Rather, the safety issues are divided into different levels with individual

properties and approaches, such as the embedded safety intelligence of robots, the manners of operation between robots and humans, and the legal regulations to control the usage and the code of robots. Therefore, the safety issues of robots could be solved step by step in this framework in the future.”

Weng also noted that, by preventing robots from understanding human language, legal machine language could help maintain a distance between humans and robots in general.

“If robots could interpret human legal language exactly someday, should we consider giving them a legal status and rights?” he said. “Should the human legal system change into a human-robot legal system? There might be a robot lawyer, robot judge working with a human lawyer, or a human judge to deal with the lawsuits happening inter-human-robot. Robots might learn the kindness of humans, but they also might learn deceit, hypocrisy, and greed from humans. There are too many problems waiting for us; therefore we must consider if it is a better to let the robots keep a distance from the human legal system and not be too close to humans.”

In addition to using machine language to keep a distance between humans and robots, the researchers also consider limiting the abilities of robots in general. Another part of the authors’ proposal concerns “human-based intelligence robots,” which are robots with higher cognitive abilities that allow for abstract thought and for new ways of looking at one’s environment. However, since a universally accepted definition of human intelligence does not yet exist, there is little agreement on a definition for human-based intelligence. Nevertheless, most robotics researchers predict that human-based intelligence will inevitably become a reality following breakthroughs in computational artificial intelligence (in which robots learn and adapt to their environments in the absence of explicitly programmed rules). However, a growing number of

researchers - as well as the authors of the current study - are leaning toward prohibiting human-based intelligence due to the potential problems and lack of need; after all, the original goal of robotics was to invent useful tools for human use, not to design pseudo-humans.

In their study, the authors also highlight previous attempts to prepare for a human-robot coexistence society. For example, the European Robotics Research Network (EURON) is a private organization whose activities include investigating robot ethics, such as with its Roboethics Roadmap. The South Korean government has developed a Robot Ethics Charter, which serves as the world's first official set of ethical guidelines for robots, including protecting them from human abuse. Similarly, the Japanese government investigates safety issues with its Robot Policy Committee. In 2003, Japan also established the Robot Development Empiricism Area, a "robot city" designed to allow researchers to test how robots act in realistic environments.

Despite these investigations into robot safety, regulators still face many challenges, both technical and social. For instance, on the technical side, should robots be programmed with safety rules, or should they be created with the ability for safety-oriented reasoning? Should robot ethics be based on human-centered value systems, or a combination of human-centered value systems with the robot's own value system? Or, legally, when a robot accident does occur, how should the responsibility be divided (for example, among the designer, manufacturer, user, or even the robot itself)?

Weng also indicated that, as robots become more integrated into human society, the importance of a legal framework for social robotics will become more obvious. He predicted that determining how to maintain a balance between human-robot interaction ([robot](#) technology development) and social system design (a legal regulation framework) will present the biggest challenges in safety when the human-robot

coexistence society emerges.

More information:

www.yhweng.tw

“Toward the Human-Robot Co-Existence Society: On Safety Intelligence for Next Generation Robots.” Yueh-Hsuan Weng, Chien-Hsun Chen, and Chuen-Tsai Sun. International Journal of Social Robotics. DOI 10.1007/s12369-009-0019-1.

Copyright 2009 PhysOrg.com.

All rights reserved. This material may not be published, broadcast, rewritten or redistributed in whole or part without the express written permission of PhysOrg.com.

Citation: Living Safely with Robots, Beyond Asimov's Laws (2009, June 22) retrieved 22 March 2023 from <https://phys.org/news/2009-06-safely-robots-asimov-laws.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.