

20 November 2008

By: Dan Talpalariu, Science Editor



Saving the lives of some humans might mean endangering the lives of others
uni-karlsruhe

[Robots Should Pose No Threat to Humans](#)

A few pieces of advice on how to design such machines

The exponential progress of technology helps specialists design and build ever more advanced machines expected to fully replace humans soon. They already look quite [similar to us](#), or are able to [chat](#), although they are much [stronger](#), [play musical instruments](#) better, beat humans at chess, [ride monocycles](#), [drive cars](#), can [adapt to injuries](#) or [explore the oceans](#) for us. Bearing these in mind, it shouldn't be long until they could make their own decisions. But is this a good thing?

Along with the technology, the number of persons worried by the development of the robotics field also increases. Their main concern is that, at some point, more advanced robots will be subjected to difficult moral and ethical problems which could imply harming humans "for the greater good," or simply because they're ordered to, such as the army has a plan to limit human soldier casualties by sending robots to combat. Wendell Wallach, an ethicist from Yale University, and Colin Allen, a cognitive science historian and philosopher from Indiana University, have tried to offer a way of avoiding that robots would ever pose a threat to humans by providing a series of simple rules comprised in their recent book entitled "[Moral Machines: Teaching Robots Right from Wrong](#)". They strongly advise humans not to put machines in high-risk situations anymore so that the consequences of their actions are easier to predict and monitor. Somewhat in the same respect, they should not be provided with weapons, contrary to the military plans and actions worldwide. The authors also suggest embedding rules (like those imagined by the famous SF writer Isaac Asimov: never to hurt a human, obeying them and treating self-preservation as the least important), principles (like "greatest good for the greatest number", or to "treat others as you would wish to be treated"). But all these are morally and ethically (and so far also technologically) limited, since they would allow for very complicated contradictions that the robot would not now how to address. As a solution, the authors suggest designing robots with the ability to learn along the way like humans do, to be able to express emotions and decipher body language or situational contexts based on previous experiences. The combination of all these ground rules would perhaps bring us closer to a more friendly and helpful machine world in the future.