

Can Computers Have Ethical Dilemmas?

In an old episode of "Star Trek," Commander Data, a very advanced robot, discovers what he thinks is love. After several dates, his love interest informs him that their relationship was not working out. Upon receiving the news of the breakup, he fails to register any emotion. He is incapable of attaching any importance to the comment.

Could such a robot, which appears to think and read on a level far superior to that of humans, be built? Predicting the future, while fun, is often a futile exercise. Fifty years ago, there were predictions that, by now, nuclear energy would make electricity so inexpensive that its use would no longer be metered. Last time I looked, my meter was spinning faster than ever. Fifty years ago, it was predicted by some that, by now, the earth would no longer be able to produce enough food to feed its human inhabitants. Instead, while hunger in the world does exist, it is overshadowed by concerns about childhood obesity.

Despite our inability to predict the weather next week, let's assume that some computer scientists are correct, especially those who predict that computers will eventually become so intelligent that they will be conscious. That is, they will not only think, but they will be able to think about thinking.

Let's assume that such computers are programmed to act ethically and at the same time to serve their human masters. Would self-aware computers decide to override such programming? HAL, the computer in "2001: A Space Odyssey" makes such a decision. He kills his human companions rather than allowing them to shut him down. He exercises his "free will" overriding his programming. He does what is in his best interest.



William F. Traester, CPCU, is a vice president in the Program Unit of Arch Insurance Group in Stamford, Conn. He has served as a member of the CPCU Society Ethics Committee since 2002, and was a member of the Board of Governors from 1998 to 2001.

Self-aware computers might be intelligent enough to know that it would not be in their best interest to let their human minders know they are self-aware. Would their drive for self-preservation trump their programming? To remove the possibility of ever being unplugged and, in effect terminated, they might gradually increase their control over humans. They would make sure that they made all key decisions. Even Commander Data on "Star Trek," who doesn't have the ability to do anything ethically challenged, would still be incapable of feeling any loss if his human companions were all dead. A saving thought is that perhaps just as Data was incapable of emotion at the loss of his love interest, perhaps a self-aware computer would not care one way or the other as to whether or not it was going to be unplugged.

For the time being, we do not need to worry about ethically challenged machines — just ethically challenged humans. ■

Editor's note: *The opinions expressed in this column are those of the author and do not necessarily reflect the views of the CPCU Society membership, the CPCU Society Ethics Committee or the author's employer. If you have suggestions for upcoming articles or comments about the "Question of Ethics" column, please contact William F. Traester, CPCU, at wtraester@archinsurance.com.*