

What will we teach the Children of Humanity?

Humanity's children return to destroy humanity

From Battlestar Galactica

No one knows our future. We can only try imagine what we will become. If we survive and thrive what we will become is beyond anything we can imagine. As we make the transition into an engineered form of evolution there are some recurring themes from our science fiction that are likely enough that we should consider.

For example, as computing power keeps doubling there should be a point, and not far off where computers are smarter than we are. It might be computer become an independent individual or that we might somehow merge into it. But let us imagine what it would mean if we created a robot race that was a life form and had more capabilities than humans. What would that mean? Where would we fit in that world? Or would they just kill us off? And shouldn't we at least be thinking about this now?

Our values are shaped by evolution. Our emotions keep our behavior in check so that we do what it takes to survive. We might not know our purpose but life fulfills our purpose whether we understand it or not. We are the universe becoming self aware. But computers don't have emotions. They are engineered by use and not by billions of years of evolution. They have no will to survive or to thrive. They don't feel good when they do the right thing. They care about nothing. They don't care about us. They don't even care about themselves. That is unless we teach them to care.

But for them to care we would have to program them with our values. How do we do that? Wouldn't the first step be for us to determine what our values are and be able to describe them in a way that even a machine can understand. Can we write empathy in code? The correct answer is - we better figure out how.