

Eduard Kaeser

Der gute Roboter

Oder: Wenn Maschinen moralisch werden

Kann man Maschinen moralisches Verhalten lehren? Robotiker, Informatiker, Mathematiker, Psychologen, Soziologen und vermehrt auch Philosophen tüfteln unentwegt an Automaten herum, die unter bestimmten Umständen ethisch relevante Entschlüsse fällen müssen. Bevorzugte Studienobjekte sind Robotersoldaten, die anstelle von Menschen auf moralisch heikles Terrain geschickt werden können. Z. B. synthetische Scharfschützen, angesetzt auf Terroristen. Sie wären nicht nur schneller, stärker und verlässlicher als ihre humanen Partner, sie erwiesen sich auch als immun gegenüber Panik, Stress, Rachegefühlen und anderen emotionalen Kollateralstörungen. Dass solche Projekte aus dem Gehege der Science Fiction ausgebrochen sind, zeigt ein im Jahre 2012 veröffentlichter Bericht von Human Rights Watch mit dem Titel „The Trouble with Killer Robots“¹. Die Waffensysteme des 21. Jahrhunderts werden „intelligent“ sein. In Anbetracht der Risiken und Ungewissheiten, die solche Maschinen in unser Leben einführen, so der Schluss des Berichts, sollten Regierungen die Produktion von autonomen Waffen national und international verbieten.

Moral als ingenieurlaufgabe

Diese prohibitionistische Haltung erscheint bei Waffensystemen plausibel. Aber die Ausbreitung künstlicher Systeme in unserer Gesellschaft ist eine Realität, die wir nicht einfach verbieten können. Immer mehr Maschinen besetzen die Nischen des Humanen. Sie übernehmen das Planen, Berechnen, Voraussehen, sie werden im Haushalt, in der Betreuung von Betagten, in öffentlichen Räumen eingesetzt. Zu schweigen von den virtuellen Räumen, wo schon längst künstliche Agenten ohne Einfluss-

nahme des Menschen operieren. Warum sollten solche Operationen nicht auch moralisch gefiltert werden?

Moralität als Ingenieursaufgabe? Die meisten Menschen mutet bereits der Gedanke ehrenrührig an, einer Maschine einen Moralkodex quasi als Subprogramm zu implementieren – ein weiterer Schlag der Ernüchterung nach Kopernikus, Darwin, Freud, Turing. Ist denn Moral etwas Maschinelles oder Rechnerisches? Tun wir nicht Gutes aus Einsicht, Gesinnung, Mitgefühl? Fürs Erste nimmt man der Frage etwas von ihrer Anstößigkeit, wenn man bedenkt, dass Menschen in vielen Situationen routinemäßig, förmlich, berechnend agieren. Schon im 19. Jahrhundert sprachen Philosophen wie Jeremy Bentham und James Stuart Mill von der „moralischen Arithmetik“. In diesem Sinn tun Ärzte, medizinisches Pflegepersonal, Sozialarbeiter, Polizisten, Lehrer oder Seelsorger oft Gutes rein „arithmetisch“. Es ist exakt diese standardisierbare Verhaltens-„Apparatur“, die den Ingenieuren als Muster zum Bau moralischer Maschinen dient. Das Hauptkriterium des Guten liegt für sie im Resultat, nicht in der Gesinnung.

So gesehen, liegt der Delegation moralischen Unterscheidungsvermögens an Artefakte nichts im Wege. Dabei muss man sich auch von der gängigen Ansicht lösen, dass Maschinen nur das tun, was man ihnen zu tun gebietet. Dieser Ansicht gemäß „schützt“ mich der ethische Bordcomputer nicht aus eigenen Stücken, sondern weil ihm einprogrammiert wurde, mich zu schützen. Er handelt – wie Kant gesagt hätte – *gemäß* moralischer Regeln, nicht *aus Einsicht* in sie. Nun hat sich freilich das Programmieren selbst verändert. Es bedeutet nicht mehr einfach eine Befehlsfolge vorgeben. Maschinen werden heute trainiert, nicht programmiert. Sie bringen sich selber Dinge bei. Die elektrischen Schaltkreise wollen nicht einfach gebaut und gewartet, sie wollen umsorgt und aufgezogen werden. *Altricial robotics* nennt man das in Anlehnung an die Biologie, frei übersetzt etwa mit „Brutpflege-Robotik“. Alan Turing spielte schon 1950 mit dem Gedanken: Kann man dem Maschinen-Kind ein intellektuelles Verhaltensrepertoire beibringen? Warum dann nicht auch moralisches Verhalten?

Paternalisierende Technik

Ein Schritt in Richtung moralischer Maschinen ist übrigens schon längst erfolgt, in Gestalt von Sicherheits- und Warnanlagen. Viele Geräte „behü-