

Das KI-Verfahren nach John Rawls¹

§ 3 Das KI-Verfahren in vier Schritten

1. Es ist wichtig einzusehen, daß das moralische Gesetz, der kategorische Imperativ und das KI-Verfahren drei verschiedene Dinge sind.

Das moralische Gesetz ist eine Idee der Vernunft. Sie bestimmt ein Prinzip, das für alle vernünftigen und rationalen Wesen (kurz: für alle vernünftigen Wesen) gilt, einerlei, ob sie – wie wir selbst – endliche Wesen mit gewissen Bedürfnissen sind oder nicht. Es gilt für Gott und die Engel sowie für vernünftige Wesen sonstwo im Universum (falls es welche gibt) ebenso wie für uns selbst.

Der kategorische Imperativ richtet sich – als Imperativ – nur an diejenigen vernünftigen Wesen, die, weil sie endliche Wesen mit gewissen Bedürfnissen sind, das moralische Gesetz als Einschränkung empfinden. Als Wesen dieser Art empfinden wir das moralische Gesetz so, und daher gibt der kategorische Imperativ an, wie dieses Gesetz für uns gelten soll (GMS, II, 12–15, S. 412–414).

Damit der kategorische Imperativ für unsere Situation gilt, muß er unseren Umständen im Rahmen der Naturordnung angepaßt werden. Diese Anpassung wird vom KI-Verfahren geleistet, das mit Hilfe der Naturgesetz-Formulierung die normalen Bedingungen des menschlichen Lebens in Rechnung stellt (GMS, II, 33, S. 421).

2. Nachdem diese einleitenden Bemerkungen gemacht sind, kann ich nun das KI-Verfahren in vier Schritten vorführen. Beim ersten Schritt haben wir die Maxime des Akteurs vor uns, die, wie wir annehmen wollen, von seinem Standpunkt rational ist; d. h. die Maxime ist rational, wenn man die Situation des Akteurs und die verfügbaren Alternativen ebenso berücksichtigt wie seine Wünsche, Fähigkeiten und Überzeugungen (von denen angenommen wird, daß sie unter den gegebenen Umständen rational sind). Kant bezeichnet eine Maxime als subjektives Prinzip. Sie ist ein Prinzip, nach dem das Subjekt handelt (GMS, II, 30, Anm., S. 421). Wenn die Maxime des Akteurs, wie hier unterstellt wird, von seinem eigenen Standpunkt rational ist, kann man sie subjektiv gültig nennen.

Ferner wird angenommen, daß die Maxime aufrichtig ist. D. h. sie reflektiert die wirklichen Gründe des Akteurs für seine beabsichtigte Handlung, wie er – dem klaren Denken unterstellt wird – sie aufrichtig beschreiben würde. Demnach gilt das KI-Verfahren für Maximen, zu denen klardenkende und rationale Akteure angesichts der von ihnen für relevant erachteten Merkmale ihrer Situation gelangt sind. Dem sollten wir hinzufügen, daß das Verfahren ebenso gut für Maximen gilt, zu denen rationale und aufrichtige Akteure unter Voraussetzung der normalen Umstände des menschlichen Lebens gelangen könnten (aber faktisch nicht gelangt sind).

Fazit: Die Maxime des Akteurs ist beim ersten Schritt sowohl aufrichtig als auch rational. Sie ist ein spezifischer hypothetischer Imperativ [...]; und da zu seiner Formulierung das Personal-

¹ Der folgende Text ist entnommen aus: John Rawls (2000): *Geschichte der Moralphilosophie. Hume – Leibniz – Kant – Hegel*, hrsg. von Barbara Herman, Frankfurt a. M. 2002, S. 230–35.

pronomen der ersten Person verwendet wird, wollen wir sagen, damit werde die persönliche Absicht des Akteurs, nach dieser Maxime zu handeln, ausgedrückt. Folgendes ist ihre Standardform:

- (1) Unter den Umständen U soll ich X tun, um Y herbeizuführen, es sei denn Z. (Hier ist X eine Handlung und Y ein Zweck, ein Sachverhalt.)

Zu beachten ist, daß die Maxime einen »um-zu«-Nebensatz enthält und sich damit auf einen Zweck bezieht. Nach Kant haben *alle* Handlungen Zwecke (MS, VI, S. 384f.). [...]

3. Der zweite Schritt verallgemeinert die Maxime des ersten Schritts. Das Resultat kann man als universelles, für jeden gültiges Gebot bezeichnen (was allerdings nicht Kants eigene Terminologie ist). Wenn dieses Gebot die Probe des KI-Verfahrens besteht, ist es ein praktisches Gesetz, ein objektives und für jedes rationale Wesen gültiges Prinzip (GMS, II, 30, Anm., S. 421). Demnach haben wir folgende Formulierung:

- (2) Unter den Umständen U soll jeder X tun, um Y herbeizuführen, es sei denn Z.

Beim dritten Schritt sollen wir das universelle Gebot unter (2) in ein Naturgesetz verwandeln, um folgende Formulierung zu erhalten:

- (3) Unter den Umständen U tut jeder immer X, um Y hervorzubringen, als ob es einem Naturgesetz entspräche (als ob ein solches Gesetz durch Naturinstinkt in uns eingepflanzt wäre). (II, 37, S. 422 f.)

Der vierte Schritt ist der komplizierteste. [...] Der intuitive Grundgedanke ist folgender:

- (4) Wir sollen das Als-ob-Naturgesetz des 3. Schritts den schon gegebenen Naturgesetzen (gemäß unserem Verständnis dieser Gesetze) hinzufügen und uns dann möglichst gut überlegen, wie die Ordnung der Natur wohl beschaffen sein mag, sobald die Wirkungen des neu hinzugefügten Naturgesetzes genügend Zeit gehabt haben, um zum Tragen zu kommen.

Hier wird angenommen, daß sich aus der unter Schritt (3) vollzogenen Vermehrung der bisherigen Naturgesetze um ein weiteres Gesetz eine neue Naturordnung ergibt, die ihrerseits einen festen Gleichgewichtszustand hat, dessen relevante Merkmale wir herausbekommen können. Diese neue Naturordnung wollen wir eine »angeglichene soziale Welt« nennen. Ferner wollen wir annehmen, diese soziale Welt stehe mit der unter Schritt (1) genannten Maxime in Verbindung. Außerdem unterstellen wir dem Akteur eine gesetzgeberische Absicht, gleichsam als wollte er einer solchen Welt das Gesetz vorschreiben. Hier läuft der Grundgedanke darauf hinaus, daß ein vernünftiger Idealakteur, der sich überlegt, ob er nach der Maxime unter Schritt (1) handeln soll, implizit jene Forderungen der reinen praktischen Vernunft akzeptiert, die von den Schritten bis einschließlich (4) dargestellt werden.

4. Kants kategorischer Imperativ läßt sich nun wie folgt formulieren: Man darf nur dann nach der unter Schritt (1) genannten rationalen und aufrichtigen Maxime handeln, wenn die beiden folgenden Bedingungen erfüllt sind:

Erstens, man muß als aufrichtiger, vernünftiger und rationaler Akteur dazu imstande sein, nach dieser Maxime zu handeln, wenn man sich als Angehörigen der damit verbundenen angeglichenen Welt und damit als jemanden sieht, der innerhalb dieser Welt und gemäß ihren Bedingungen handelt.

Zweitens, wir müssen diese angegliche soziale Welt als solche wollen können und sie bejahen, falls wir ihr angehören.

Also gilt: Wenn wir außerstande sind, diese angegliche soziale Welt zu wollen und zugleich als Angehörige dieser Welt nach jener Maxime zu handeln, können wir jetzt nicht nach dieser Maxime handeln, obwohl sie laut Voraussetzung unter den gegebenen Umständen völlig rational ist. Der kategorische Imperativ in der vom KI-Verfahren dargestellten Form gilt für uns unabhängig davon, welche Konsequenzen sich aus unserer Befolgung des Imperativs im Hinblick auf unsere natürlichen Wünsche und Bedürfnisse ergeben mögen. Das reflektiert den Vorrang der reinen praktischen Vernunft gegenüber der empirischen praktischen Vernunft.

Schließlich sei noch angemerkt: Man sollte nicht vergessen, daß sich diese Wiedergabe des KI-Verfahrens auf die Naturgesetz-Formulierung stützt. Diese Formulierung lautet:

Handle so, als ob die Maxime deiner Handlung durch deinen Willen zum allgemeinen Naturgesetze werden sollte. (II, 33, S. 421)

Diese Formulierung haben wir als unterstellte Gesetzgebungsabsicht interpretiert: Es ist so, als hätten wir die Macht der gesetzgeberischen Vernunft und müßten sie – als Bedingung des Handelns nach unserer Maxime – zum Einsatz bringen. Geprüft werden muß, ob wir das, was wir jetzt in der angeglichenen sozialen Welt beabsichtigen, tun können, und ob wir diese Welt auch wollen können.

§ 4 Kants zweites Beispiel: das unredliche Versprechen

1. Kant nennt vier Beispiele, deren zweites und viertes zumindest auf den ersten Blick einleuchtender wirken als die übrigen beiden. Nun möchte ich knapp veranschaulichen, wie sich die vier Schritte des KI-Verfahrens in diesen Fällen anwenden lassen. Ich beginne mit dem zweiten Beispiel, in dem es um das unredliche Versprechen geht (GMS, II, 36, S. 422).

Schritt (1): Unter den Umständen U (d. h. wenn ich finanziell in Verlegenheit bin und Geld brauche, aber weiß, daß ich die Schulden nicht zurückzahlen kann, und keine Rückzahlung beabsichtige) soll ich ein unredliches Versprechen geben, um mir einen persönlichen Vorteil zu verschaffen.

Schritt (2): Unter den Umständen U soll jeder ein unredliches Versprechen geben usw., wie oben.

Schritt (3): Unter den Umständen U gibt jeder ein unredliches Versprechen (oder versucht ein solches Versprechen zu geben) usw. (als ob es einem Naturgesetz entspräche).

Schritt (4): Man füge das Naturgesetz des 3. Schritts den übrigen (uns bekannten) Naturgesetzen hinzu und ermittle den resultierenden Gleichgewichtszustand. Diese angegliche so-

ziale Welt ist so beschaffen, daß niemand in ihr unter den Umständen U ein unredliches Versprechen geben kann, so gern er es auch täte.

Das Naturgesetz unter (3) ist ein psychologisches Gesetz: Wir sind bestrebt, unredliche Versprechen zu geben, so als ob es einem Naturgesetz unterläge. Aber da unter bestimmten Umständen andere Gesetze zum Zuge kommen und dieses Gesetz in seiner Wirkung behindern können, sagen wir nicht rundheraus, jeder gebe unredliche Versprechen. Es mag zwar als psychologisches Gesetz gelten, daß jeder unter diesen Umständen dazu bereit ist, es zu versuchen; doch andere Gesetze können es nach sich ziehen, daß ein unredliches Versprechen im gegebenen Fall unmöglich ist.

Die Maxime des unredlichen Versprechens scheitert an der Probe des begrifflichen oder gedanklichen Widerspruchs, denn ein rationaler Akteur kann in der sozialen Welt der gesetzgeberischen Absicht nicht nach dieser Maxime zu handeln beabsichtigen. Das folgt daraus, daß rationale Akteure, wenn sie etwas zu tun beabsichtigen, die begründete Überzeugung haben müssen, daß sie es tun können und daß es unter den gegebenen Umständen in ihrer Macht liegt. Eine Absicht ist so etwas wie ein Plan; und es ist nicht rational, eine Handlung zu planen, von der man weiß, daß man sie nicht ausführen kann.

2. Zwei Absichten, die sich als unverträglich erweisen können, werden hier deshalb eingeführt, weil klar werden soll, wieso Kant von der Selbstwidersprüchlichkeit der Maxime des Akteurs sprechen kann. Aber warum glaubt Kant eigentlich, in der angeglichenen sozialen Welt könne niemand ein falsches Versprechen geben? Er äußert sich knapp und sagt, die Allgemeinheit des Gesetzes »würde das Versprechen und den Zweck, den man damit haben mag, selbst unmöglich machen, indem niemand glauben würde, daß ihm was versprochen sei, sondern über alle solche Äußerung als eitles Vorgeben lachen würde« (GMS, II, 36, S. 422).

Nun setzt Kant es offensichtlich als Naturgesetz voraus, daß die Menschen aus Erfahrung lernen und sich an Vergangenes erinnern. Demnach gilt: Sobald es gleichsam ein Naturgesetz wird, daß jeder (unter bestimmten Umständen) falsche Versprechungen macht, wird die Existenz dieses Gesetzes öffentlich bekannt. jeder weiß darüber Bescheid, und weiß, daß es auch anderen bekannt ist usw. Dabei brauchen wir nicht anzunehmen, daß alle Naturgesetze öffentlich bekannt sind, denn offensichtlich sind sie es nicht. Aber um die Forderungen des KI-Verfahrens in der Terminologie der Naturgesetz-Formulierung zu interpretieren, ist es nicht unangemessen, die öffentliche Kenntnisnahme der von den nach bestimmten Maximen handelnden Personen erzeugten Als-ob-Naturgesetze zu unterstellen.

Um das explizit zu machen, sagen wir: Im Gleichgewichtszustand der angeglichenen sozialen Welt sind die Als-ob-Naturgesetze des 3. Schritts öffentlich als Naturgesetze anerkannt, und das KI-Verfahren muß dementsprechend angewandt werden. Diese öffentliche Anerkennung der aus den unter Schritt (1) genannten Maximen hervorgehenden Als-ob-Naturgesetze wollen wir als Öffentlichkeitsbedingung für allgemeine moralische Gebote bezeichnen. Nach Kants Auffassung gehören akzeptable Gebote dieser Art sozusagen zur öffentlichen Moralgesetzgebung einer sittlichen Gemeinschaft.