

Geert Keil

Was Roboter nicht können. Die Roboterantwort als knapp mißlungene Verteidigung der starken KI-These

in: *Der Mensch in der Perspektive der Kognitionswissenschaften*,
hrsg. von Andreas Engel und Peter Gold, Frankfurt am Main (Suhrkamp) 1998, 98-131.

Zusammenfassung

Theoretiker der *Künstlichen Intelligenz* und deren Wegbegleiter in der Philosophie des Geistes haben in den vergangenen Jahren auf verschiedene Weise auf Kritik am ursprünglichen Theorieziel der KI reagiert. Eine dieser Reaktionen ist die Zurücknahme dieses Theorieziels zugunsten der Verfolgung kleinerformatiger Projekte. Eine andere Reaktion ist die Propagierung konnektionistischer Systeme, die mit ihrer dezentralen Arbeitsweise die neuronalen Netze des menschlichen Gehirns besser simulieren sollen. Eine weitere ist die sogenannte *robot reply*, mit der sich dieser Aufsatz beschäftigt.

Die Roboterantwort besteht aus zwei Elementen. Sie enthält (a) das Zugeständnis, daß das Systemverhalten eines wie auch immer programmierten konventionellen Digitalrechners mit von Neumann-Architektur nicht schon menschenähnliche Intelligenz aufweist. Sie enthält (b) die Behauptung, daß es für bestimmte Arten von Maschinen *doch* zur Intelligenz reicht. In die Liga der intelligenten Wesen könnten Maschinen genau dann aufsteigen, wenn sie Roboter sind. Damit ist gemeint: wenn sie über Wahrnehmungskomponenten (Rezeptoren) und Handlungskomponenten (Effektoren) verfügen, mithilfe deren sie aktiv in kausale Interaktionen mit ihrer Umwelt eintreten können.

Ich werde die These vertreten, daß der Roboterantwort eine richtige Intuition zugrunde liegt, von der die Roboterfreunde sich aber zu einer kurzschlüssigen Folgerung verleiten lassen. Den Roboterfreunden ist darin zuzustimmen, daß sie mentale Zustände und Handlungskompetenz so eng aneinander binden. Wir sollten tatsächlich keinem Wesen, dem wir Handlungsfähigkeit zuerkennen, mentale Zustände absprechen. Doch Handelnkönnen und Geisthaben sind nicht hinreichend unabhängig voneinander, als daß man das eine als *Rechtfertigung* für die Zuerkennung des anderen verwenden könnte. Wir sollten Robotern beides absprechen.

Im Schlußteil des Aufsatzes werde ich zusätzlich behaupten, daß die Roboterantwort eine Abkehr vom ursprünglichen Theorieziel der KI darstellt. Die KI hat sich auf die funktionalistische These der multiplen Realisierbarkeit des Mentalen verpflichtet. Damit es so etwas wie Künstliche Intelligenz geben kann, muß Intelligenz aus ihrer Verstrickung mit kontingenten physischen Eigenschaften menschlicher Intelligenzträger herausgelöst werden können. Mit der Roboterantwort ist der Boden dieses Forschungsprogramms verlassen. Wenn sie nämlich zutreffen soll, wird völlig unklar, welche Merkmale des Menschen wir überhaupt noch als kontingent ansehen können. Sind die dem Roboter zugeschriebenen Fähigkeiten nicht nur an mentale, sondern auch an bestimmte physische Voraussetzungen gebunden, dann kehren auf diese Weise Restriktionen zurück, von denen die GOFAI (Good Old-Fashioned Artificial Intelligence) gerade abstrahieren zu dürfen glaubte. Durch den Übergang von mentalen zu körperlichen Operationen haben die Vertreter der Roboterantwort die These der multiplen Realisierbarkeit entscheidend geschwächt. Was die GOFAI anstrebte, war ein künstlicher *Geist*, nicht ein künstlicher Körper. Wenn wir nun menschenähnliche Roboter oder gar biologisch-technische Hybriden einführen müssen, um die Zuschreibung mentaler Zustände zu retten, läuft dies auf eine Abkehr vom ursprünglichen Theorieziel der KI hinaus. Dann sind wir längst auf dem Wege von der künstlichen Intelligenz zum künstlichen *Menschen*. Je mehr ein System können, je menschenähnlicher es sein muß, um als geistbegabt zu gelten, desto weiter bewegt sich die KI auf ihre Kritiker zu. Die Roboterantwort stellt also, aller Rhetorik zum Trotz, eine *Rückzugsposition* der KI dar. Alan Turing war auf der Suche nach „a fairly sharp line between the physical and the intellectual capacities of man“. Vieles spricht dafür, daß es diese Grenze nicht gibt.