

Interview mit Dietrich Dörner



Prof. Dr. sc. Dietrich Dörner studierte Psychologie an der Universität Kiel und wurde 1979 Professor für Allgemeine Psychologie an der Universität Bamberg. Von 1989 bis 1991 hatte er die Leitung der Arbeitsgruppe Kognitive Anthropologie am Berliner Max-Planck-Institut inne. Bis zu seiner Emeritierung im Sommer 2006 leitete er das Institut für Theoretische Psychologie an der Universität Bamberg. Zu seinen wichtigsten Arbeitsgebieten gehörten die kognitive Organisation des Problemlösens, menschliche Handlungssteuerung und die Modellierung von Emotionen.

KI: Herr Professor Dörner, was kann die Künstliche-Intelligenz-Forschung aus Ihrer Sicht von der Psychologie lernen?

Sie kann lernen, wie man integrierte Systeme baut, die nicht nur allein eine spezifische, sondern viele verschiedene Sachen machen können. Menschen können vielen Aufgaben zugleich nachgehen und diese miteinander integrieren. Lernende Systeme müssen so beschaffen sein, dass sie sich in vielen Umgebungen zurechtfinden, in verschiedenen Realitäten bewegen können. In dieser Beziehung sind natürliche Systeme, nicht nur Menschen, ganz toll.

Integrierte Systeme bauen

Nicht zuletzt kann man aus der Psychologie lernen, wie man Systeme bauen kann, die sich selber zum Gegenstand der Betrachtung machen, um sich auf dem Wege der Selbsterkenntnis zu modifizieren. Aus meiner Sicht wird dieses Thema in der heutigen psychologischen Forschung aber leider zu wenig behandelt.

KI: Ihr Forschungsgebiet ist die Theoretische Psychologie. Was ist Ihr Untersuchungsgegenstand?

Mich interessiert, wie die Seele konstruiert ist, wie sie zusammenhängt. Aristoteles, einer von den ersten Psychologen, hat bereits angefangen, eine funktionale Psychologie zu bauen. Die Psychologen haben das leider weitestgehend wieder vergessen, als sie um 1900 herum dachten, um eine richtige Wissenschaft zu sein, dürften sie nur Beobachtungen und Experimente machen, und um Gottes Willen nicht darüber nachdenken, wie die

Seele „innendrin“ aussieht. Ein bisschen hat sich das mit der Kybernetik in den fünfziger und sechziger Jahren, und mittlerweile in der Kognitionswissenschaft geändert.

KI: Ein bekanntes Buch aus Ihrer Feder heißt „Bauplan für eine Seele“. Was genau bauen Sie da?

Wir haben den Seelenbegriff des Aristoteles wiederbelebt, der vor 2300 Jahren schlicht und klar sagte: „Die Seele ist das Prinzip des Lebendigen“, oder: die Seele ist das System von Regeln, die funktionieren müssen, damit, wie er sagt, „ein Körper Leben hat“. Die Seele ist ein Regelsystem.

Wir gehen dabei mehr ingenieurtechnisch vor: Wir schauen, was Menschen können und überlegen: Wie muss eine Maschine aussehen, die das auch kann? Dann haben wir einen möglichen Kandidaten für eine Theorie von einem Aspekt des menschlichen Denkens, des Fühlens, des Wollens, oder was auch immer man betrachtet.

KI: Sie schlagen also vor, psychologische Forschung auf dem Weg der Simulation voranzutreiben?

Das halte ich für den einzigen gangbaren Weg. Die rein neurobiologische Hirnforschung bringt nicht viel Aufklärung über die Funktionen! Aus der Hirnforschung lernen wir etwas über die Funktionsweise einzelner Neuronen, wichtige Bahnen auf denen sie verschaltet sind und so weiter, aber wichtig ist doch: Was geschieht auf diesen Bahnen? Was geschieht, wenn wir einen roten Fleck auf dem Resonanztomogramm sehen?

Das Ziel unserer Arbeit ist es, das Gehirn im Computer nachzu-

bauen, aber im Mittelpunkt steht dabei nicht so sehr die neuroanatomische Struktur, sondern die Funktion.

Die Funktion des Gehirns im Mittelpunkt

Wir schauen beispielsweise: es muss einen Mechanismus geben, der zwischen verschiedenen Motiven auswählt. Wie kann der beschaffen sein? Wir überlegen uns, wie er strukturiert sein kann und dann bauen wir ihn. Wir gucken uns an, ob er dasselbe tut wie das Gegenstück bei Menschen, wenn sie Absichten und Motive wechseln. Wenn wir dann in die neurowissenschaftliche Forschung sehen, stellen wir fest: das könnte vielleicht der Thalamus sein, den wir da gebaut haben, denn der erfüllt diese Aufgabe und hat ganz ähnliche Merkmale, und auf diese Weise stellen wir die Beziehung zum Gehirn her.

KI: Offenbar hat Ihre Zielrichtung Ähnlichkeit zu den ursprünglichen Zielen der KI-Forschung...

Ich bin durch die Kybernetik zur Psychologie gekommen. Als ich 1960 zu studieren begann, war die Rede von Elektronengehirnen, und es wurde damals bereits diskutiert, ob man damit das Gehirn nachbauen könnte. Karl Steinbruch und die Kybernetiker um Norbert Wiener sagten: Regelungstechnik, das ist Leben. Dabei spielte das Verständnis der Motivationen eine große Rolle. Ich bin in die Psychologie hineingegangen mit der Idee, menschliches, geistiges, psychisches Geschehen zu modellieren und nachzubauen. Zu diesem Zeitpunkt arbeitete auch Friedhart Klix in Ostberlin an diesem Thema. Es gab nur wenige Leute auf der Welt,

die damals ebenfalls daran arbeiten, wie Earl Hunt, Marvin Minski, Edward Feigenbaum und einige andere. Sie hatten allerdings weniger die Kognition im Blick, sondern vor allem die Motivation.

Übrigens bin ich in hohem Maße auch von Stanislaw Lem beeinflusst worden. Lem bevölkerte seine Science-Fiction-Romane mit künstlichen Systemen, die ganz klar eine Seele hatten, mit fühlenden Robotern. Er hatte dabei vielleicht nicht immer eine ganz klare Vorstellung von den Einzelheiten, aber seine Vorstellungen gingen recht weit! Das hat mich sehr beeindruckt. Dabei setzte er jedoch viele Dinge voraus, bei denen ich dachte: das muss man erst einmal beweisen! Wir haben ein Jahr in Berlin gemeinsam am Wissenschaftskolleg verbracht, und wir haben uns sehr gut verstanden.

KI: Gegenwärtig ist die KI-Forschung anscheinend mit ganz anderen Themen beschäftigt: mit Lösungen für technische Systeme, Wissensmanagement, Logik und auch Robotik...

Ja, leider!

KI: ...und solche Architekturen, wie Sie sie vorschlagen, sind heute vor allem in der Kognitionswissenschaft zu finden. Ist die Kognitionswissenschaft die neue KI?

Auch in der Kognitionswissenschaft lässt die Integration sehr auf sich warten. Allein der Begriff „Kognition“: Erkenntnistätigkeit, das bezieht sich nur auf einen Teil der Gehirntätigkeit, da fehlen die Motive und die Emotionen und vieles mehr.

In der Kognitionswissenschaft lässt die Intergration auf sich warten

Insbesondere die Kognitionspsychologie ist leider keine Fortsetzung der Kybernetik. Die Kybernetik ist Ende der sechziger Jahre untergegangen.

KI: Welche Gebiete halten Sie für relevant, um ein integriertes System zu bauen?

Im Grunde sind alle relevant! Man darf nicht glauben, man könnte, wenn man menschliches Verhalten erklären möchte, so etwas wie Emotionen und Motive einfach weglassen. Wie das zum Beispiel bei ACT-R, der kognitiven Architektur von John Anderson, geschieht. Seine Arbeit ist bahnbrechend, und hat viele Forscher an das Thema herangeführt. Er hat aber auch viel Schaden angerichtet, denn er hat so getan, als könnte man mit der primitiven Struktur, die ACT-R hat, kompliziertes menschliches Verhalten erklären. Das kann man nicht.

Die Kognitionswissenschaft ist immer mehr in Details ausgewichen, in Experimente, in denen eine Korrelation zwischen menschlichem Verhalten und Modell gezeigt werden kann – aber mit menschlichem Verhalten nur unter ganz, ganz spezifischen Bedingungen! Die Kognitionspsychologie ist gewissermaßen empiristisch entartet: Als Modell für menschliche Komplexität fallen diese Systeme sehr kümmerlich aus.

KI: Wie unterscheiden sich Ihre Computermodelle des menschlichen Verhaltens von solchen Ansätzen?

Das Wesentliche ist, dass wir immer Motivation und Emotion integriert haben.

KI: Wie funktioniert denn Motivation? Können Sie das in wenigen Sätzen skizzieren?

Das kann man ganz einfach erklären! Das Grundprinzip ist der Regelkreis. Motivation besteht im Wesentlichen aus einer Reihe von Systemen, die versuchen, irgendetwas konstant zu halten oder zu optimieren. Das, was da konstant gehalten und optimiert wird, sind einerseits physiologische Dinge, andererseits – und das ist sehr wichtig – ist es das, was die russischen Psychologen in etwas irreführender Weise „kognitive Emotionen“ nennen. Das ist einerseits ein Bedürfnis nach der Fähigkeit möglichst viel vom Geschehen in der Umwelt vorzusagen. Das ist also das Bedürfnis nach Bestimmtheit, übrigens auch Bestimmtheit der Innenwelt – ich muss mich selbst auch verstehen. Das andere kognitive Motiv ist das Bedürfnis nach Kompetenz, oder, altertümlich gesprochen: nach Macht. Macht ist die Fähigkeit, etwas bewirken, etwas machen, die

Welt so verändern zu können, dass ich meine Bedürfnisse befriedigen kann. Das sind die beiden kognitiven Bedürfnisse, sie führen nämlich beide zu Erkenntnissen.

Kognition geht nicht ohne Motivation

Nebenbei: deshalb kann man Kognitionstheorie nicht ohne Motivationstheorie betreiben; wenn man diese kognitiven Motive weglässt, dann fehlt etwas ganz Entscheidendes.

KI: Wie funktioniert der Bezug zur Umwelt bei den Computermodellen Ihrer Theorie?

Das Wahrnehmungssystem darf nicht fest vorprogrammiert sein. Mit fest vorgegebenen Wahrnehmungsschemata kann man einen Roboter bauen, der eine festgelegte Aufgabe ausführt, zum Beispiel Autos lackiert. Beim Menschen, und bei Systemen die sich an eine beliebige Umwelt anpassen können, muss es ganz anders sein: die müssen auch neue Sachverhalte erkennen können. Sie müssen explorieren, sie müssen mit den Dingen herumspielen und auf diese Weise deren Eigenschaften kennenlernen. Sie müssen die Bedeutung dieses Gegenstandes erfassen können. Dieser Bereich der Wahrnehmung – der Bereich *Perceptual Learning* – ist die Voraussetzung, damit sie sich in beliebigen Umwelten zurechtfinden können. Und: sie müssen sich umprogrammieren können. Sie müssen aus basalen Verhaltensweisen, wie gehen, greifen, und so weiter, neue Verhaltensweisen *ad hoc* zusammenbauen können.

KI: Viele Forscher im Bereich der Kognitiven Robotik vertreten die Auffassung, dass Intelligenz sich erst im Wechselspiel mit der Umgebung ergibt. Sie bewegen sich weg von symbolischen Modellen und internen Repräsentationen und konzentrieren sich stattdessen auf die Interaktion mit der Umwelt...

Klar: erst in der Auseinandersetzung mit einer Umgebung zeigt sich Intelligenz. Intelligenz bedeutet ja, dass ein System in der Lage ist, sich sehr gut mit seiner Umgebung auseinanderzusetzen, und zwar auch mit einer Umgebung, die es erst

noch kennen lernen muss. Intelligenz ist aber auch noch ein bisschen mehr als Sensor-Motor-Kopplung. Das wäre, in Karl Bühlers alter Terminologie, erst Dressur, und nicht Intellekt. Ein System, das nur Sensor-Motor-Kopplungen lernen kann, bleibt relativ primitiv. Wichtig ist, dass Systeme auch so etwas entwickeln wie Bewusstsein. Bewusstsein bedeutet einfach die Fähigkeit zur Selbstreflexion. Systeme, die sich selbst reflektieren und kritisieren, sich umprogrammieren, müssen nicht nur aus der Umgebung lernen. Sondern sie können auch neue Ideen an die Umgebung herantragen.

Systeme müssen sich selbst reflektieren

Ein Beispiel ist das Rückwärtsplanen. In der Umwelt gibt es so etwas nicht: ich kann nicht rückwärts handeln. Im Geiste kann ich sehr jedoch wohl rückwärts handeln, also vom Ziel aus.

KI: Können Sie uns beschreiben, wie Ihre Modelle konkret aussehen? Wie kann man sich das vorstellen?

Wir arbeiten mit einer Art von neuronalen Netzen, eigentlich Schemata-Netzen. Die bestehen im Wesentlichen aus Partonomen, also aus Teil-Ganzes-Zusammenfügungen, und aus Abstrakt-Konkret-Hierarchien. Letztere sind beispielsweise wichtig für die Analogiebildung. – Das ist das Gedächtnissystem. Es gibt nur diese zwei Ordnungsstrukturen – die Linguisten nennen sie Meronymien und Hyponymien. Dass diese beiden Hierarchien ausreichen, hat beispielsweise Friedhard Klix schon früh erkannt; seine innerbegrifflichen und zwischenbegrifflichen Relationen entsprechen diesen beiden Hierarchien.

Dann ist es wichtig, dass man in das Gedächtnis Verbindungen zu Motiven hineinbringt. Das ist meines Erachtens sehr einfach, es geschieht einfach dadurch, dass man bestimmte sensorische Schemata als Ziele ausgezeichnet werden. Wenn ich beispielsweise erfahre, dass man ein Rosinenbrötchen essen kann, und es schmeckt ganz gut, und hinterher ist der Hunger weg, dann ist das Schema „Rosinenbrötchen“

mit dem Motivsystem verbunden. Umgekehrt kann das Motivsystem das Schema aktivieren, und wenn ich Hunger habe, taucht vor meinem geistigen Auge unter anderem ein Rosinenbrötchen auf. Das wird dann zu einer Zielvorstellung, die ich anstreben kann.

Motive ins Gedächtnis bringen

Es gibt also im Grunde drei Prinzipien: die Partonomen, die Abstrakt-Konkret-Relationen, und dann die Aversionen und die Appetite, also die Verbindungen der sensorischen Schemata mit Motivationssystemen, so dass man weiß, was man vermeiden und was man anstreben sollte.

Weil immer viele Motive da sind, muss das System in die Situation gebracht werden, immer mit einem einzelnen Motiv oder einem Motivamalgam, also einer Verbindung vieler Motive zugleich, zu arbeiten. Dafür muss ein Motiv ausgewählt werden, es braucht also ein System der Absichtsauswahl, sozusagen ein Willenssystem. Dieses System kann man recht einfach bauen, indem es einfach nur das stärkste Motiv auswählt. Tatsächlich handelt es sich um ein viel raffinierteres System, bei dem zum Beispiel Selbstbetrachtung, im Falle von Konflikten zwischen den Motiven, eine Rolle spielt.

KI: Woher weiß Ihr Modell, was ein sensorisches Schema für Eigenschaften hat? Muss der Programmierer das vorher „verdrahten“?

Nein, das „verdrahtet“ sich selber. Das geht ganz einfach: Beispielsweise springt das Auge die Konturlinien entlang und identifiziert die verschiedenen Konturlinien. Die Konturlinien sind dabei basal. Das Protokoll dieser Sprünge, der Abtastung, wird festgelegt und ist dann das Schema. Das ist im Grunde fürchterlich primitiv – und auch fürchterlich effektiv. Der Programmierer legt nur eine bestimmte Anzahl von Konturmustern fest, die das System kennt. Und die werden dann zu immer neuen Gebilden zusammengestrickt und bilden immer neue Schemata.

KI: Wie können Sie die Hypothese, dass es wirklich so funktioniert, überprüfen?

Vor allem frage ich: macht das System das, was es soll? Wenn wir feststellen, dass unser Modell einen Prozess in der Natur erklärt, brauchen wir gar keine Experimente. So ungefähr wie Einstein, der bekanntlich niemals Experimente machte, um seine Theorien zu überprüfen. Nachdem er die Relativitätstheorie gefunden hatte, sah er: so muss das sein, so schön und einleuchtend und so klar.

Wir bauen also Theorien nach dem Prinzip der Einfachheit, und bauen auf diese Art und Weise Systeme, die das tun, was man beim Menschen auch beobachten kann. Das ist ein Beleg dafür, dass diese Systeme als Theorien zur Erklärung für die Phänomene, die wir betrachten, in Frage kommen – und das ist schon eine ganze Menge.

Menschliches Verhalten erklären

Man kann aber auch noch mehr machen: Man kann darangehen, menschliches Verhalten vorherzusagen. Zum Beispiel, indem man Menschen in simulierte Abenteuerwelten versetzt und schaut: wie verhalten sie sich darin. Dann kann man einen künstlichen Agenten, den man entsprechend den theoretischen Annahmen über die menschliche Seelenstruktur gebaut hat, in dieselbe Welt setzen, und gucken, wie der sich verhält. Das ist aber sehr schwierig, weil sehr viele Faktoren eine Rolle spielen. Wir haben trotzdem auch mit diesem Ansatz gute Erfahrungen gemacht.

KI: Wie sind Sie dabei vorgegangen?

Wir haben große Anzahlen von Versuchspersonen unter verschiedenen Bedingungen arbeiten lassen und untersucht, wie diese Bedingungen sich im Einzelnen auswirken. Wir hatten diese Methode von den experimentellen Psychologen übernommen, und es gab durchaus einen Ertrag, aber meines Erachtens war das Verhältnis zum Aufwand nicht gut. Heute würde ich sagen: wir bauen eine Theorie und dann schauen wir, was wir damit erklären

können. So wie beispielsweise Darwin vorgegangen ist. Der hatte eine Theorie, die Evolutionstheorie, und hat sich damit einzelne Fälle, zum Beispiel ein konkretes Ökotoptop, angeguckt, und geschaut: kann ich das damit erklären? Er hat nicht damit angefangen, alle Finken, alle Ameisen und Organismenarten auf der Welt zu untersuchen, bevor er wusste, dass er Recht hatte. Er brachte die Theorie in Anschlag und akkumulierte damit erst einmal Wissen. Dabei verändert sich die Theorie – es handelt sich um ein fortschreitendes Theoriebildungsverfahren.

Fortschreitende Theoriebildung

KI: Darwin hatte seine Finken, und Einstein verfügte über Daten, die die Theorien seiner Vorgänger nicht erklären konnten. Was kann Ihre Theorie erklären, das andere nicht erklären können?

Beispielsweise die Theorie des Absichtswechsels – das ist etwas, das in der Psychologie erst sehr zaghafte als Forschungsgebiet aufgenommen wird, und das können wir meines Erachtens sehr elegant erklären, unter anderem im Hinblick darauf, dass es nur fünf Motivgruppen gibt...

KI: Welche sind diese fünf Motivgruppen?

Das sind zunächst die Motive der Existenzhaltung. Da könnte man genau genommen über hundert aufzählen, aber man kann Hunger, Durst, Schmerz darunter subsumieren. Als nächstes kommt die Sexualität. Dann die Affiliation – ein Motiv, das oft irrtümlich mit der Sexualität zusammengeworfen wird, ein Erbe der Psychoanalyse! – Denn Affiliation, den Geselligkeitstrieb, gibt es auch ohne Sexualität. Und schließlich gibt es noch Bestimmtheit und Kompetenz, die beiden kognitiven Bedürfnisse. Natürlich muss man diese Gruppen weiter ausdifferenzieren, zum Beispiel gibt es für die Affiliation ganz verschiedene Mechanismen. Aber wir können sagen, alles, was es noch gibt, setzt sich aus diesen fünf Sachen zusammen.

Eine wichtige Besonderheit unseres Ansatzes sind außerdem die Emotionen. Anders als der in der

Psychologie allgemein übliche Ansatz betrachten wir Emotionen nicht als eigenständige Gebilde, die zum Beispiel neben den Motiven und Kognitionen stehen. Emotionen sind Einfärbungen dieser anderen Prozesse. Wenn man von Angst redet, so ist dies nichts Selbständiges. Angst ist eine bestimmte Form des Wahrnehmens, eine bestimmte Form des Reagierens auf Dinge, eine bestimmte Form der Gedächtnisbildung, des Gedächtnisabrufs und so fort. So dass, wenn man das alles wegnimmt, nicht etwa die pure Angst übrig bleibt – sondern gar nichts. Dieses Konzept findet sich in Ansätzen schon bei Wilhelm Wundt, aber es ist den Leuten schwer begreiflich zu machen.

Emotionen sind nicht eigenständig, sondern Einfärbungen anderer Prozesse

KI: Wenn Sie Ihre Modelle mit menschlichem Verhalten vergleichen, was können Sie schon gut nachbilden, und was können Sie noch nicht modellieren?

Das ganz große Loch ist Bewusstsein, also Selbstreflexion, und damit hängt Sprache eng zusammen. Natürliche Sprache und Sprachentwicklung sind für die menschliche Kognition ganz wichtig und völlig vernachlässigt, nachdem Leute in der Kognitionswissenschaft gesagt haben, die Sprache wäre nur eine Art Gedankentransportsystem, aber hätte mit dem Denken nichts zu tun. Mit der Sprache hängt die Selbstbetrachtung zusammen. Aus unserer empirischen Forschung folgt, dass Menschen gegenüber den Modellen eine größere Varianz, und damit Flexibilität, in ihren Strategien haben, wenn sie Selbstreflexion einsetzen.

Sprache erhöht die strategische Flexibilität

Was mich am meisten reizt, neben den Emotionen, ist die Frage der Realisierung von Bewusstsein. Ich glaube, man muss dazu zuvörderst an natürliche Sprache ran.

KI: Auf welchem Weg sollten wir die Sprache erforschen?

Wir sollten darüber nachdenken, was Sprache eigentlich ist. Durch Chomsky ist uns die Frage der Semantik leider aus dem Griff geraten. Die Leute verwenden denselben Begriff im selben Satz oftmals in ganz verschiedenen Bedeutungen, wenn sie zum Beispiel sagen: „Das Institut ist heute geschlossen, es macht einen Ausflug.“ – Das versteht jeder sofort, obwohl mit „Institut“ jedes Mal etwas anderes gemeint ist; die Bedeutung wechselt sofort. Ich finde, Sprache ist unglaublich raffiniert. Wenn man sich diesem Problem nähern will, ist das Buch „Sprachtheorie“ von Karl Bühler von 1934, das in den Neunzigern neu aufgelegt wurde, ein guter Ausgangspunkt. Es ist wichtig, zu verstehen, wie plastisch Sprache mit Realität umgeht, und wie man durch Wortneubildungen und Umdeutungen immer neue Realitäten, und immer neue Aspekte von Realitäten einfängt. Das ist etwas, das die heutige Sprachforschung kaum untersucht. Es wäre aber ein wichtiger Anfang.

KI: Herr Professor Dörner, es war schön, dieses Gespräch mit Ihnen zu führen! Vielen Dank.

Das Interview führte Joscha Bach