

Was leistet eine Computationale Theorie der Kognition, und was nicht?

§1 Definition und Gegenstand

Die Computationale Theorie des Geistes („Computational Theory of Mind“ [CTM]) basiert auf zwei bis drei Hauptthesen:

(Prozess-These) Geistige Prozesse (bewusste und nicht-bewusste) sind algorithmische Prozesse.

(Architektur-These) Der Geist besteht in einer Architektur algorithmischer Vermögen.

Diese Architektur baut sich in Schichten auf und regelt das Zusammenwirken der verschiedenen Vermögen, wobei höhere Schichten in weniger abstrakten Prozessen niedriger Schichten umgesetzt werden.¹ Die einzelnen Vermögen bilden Module in dieser Architektur.²

Varianten der CTM unterscheiden sich darin, welche der beiden Thesen im Vordergrund steht. Projekte wie SOAR³ betonen die Rolle der Architektur als Grundgerüst einer allgemeinen Theorie des Mentalen. Andere Theoretiker entwerfen eher Modelle für einzelne Vermögen.

Mit der Prozess- und der Architektur-These verbindet sich in der Regel ein materialistischer Funktionalismus:

(Materialismus-These) Die Architektur des Geistes wird vom Gehirn physisch implementiert.

Daher stammt der Slogan: „Der Geist verhält sich zum Gehirn wie die Computer-Software zur Computer-Hardware“. Die Computermetapher bietet sich oft auch für Details der Architektur des Geistes an (etwa Datenzugriff, Arbeits- und Langzeitspeicher etc.). An dieser Stelle spielt die Architektur realer Computer eine Rolle, nicht die abstrakter Modelle wie der

¹ Vgl. z.B. Zenon Pylyshyn, *Computation and Cognition*.

² Vgl. Jerry Fodor, *The Modularity of Mind*, oder: Peter Carruthers, *The Architecture of the Mind*.

³ Vgl. Jill Lehman & John Laird & Paul Rosenbloom, "A Gentle Introduction to SOAR, An Architecture for Human Cognition: 2006 Update". Ähnlich gelagert sind die Ansätze ICARUS (vgl. Dongkyu Choi & Pat Langley, "Evolution of the ICARUS Cognitive Architecture") und ACT (vgl. John Anderson et al., "An Integrated Theory of the Mind").

Turing-Maschine [TM]. Dies erstreckt sich bis hin zu physischen Kontrollmechanismen wie Interrupts und Transducern (zu Inputs und Outputs).⁴

Im Unterschied zu einer allgemeinen CTM, die sich auf alles, was das bewusste Erleben von Personen ausmacht und umsetzt, erstreckt, kann man eine ‚Computational Theory of Cognition‘ [CTC] ansetzen als die Einschränkung der CTM auf kognitive Vermögen im engeren Sinne.

Was sind kognitive Vermögen in Abgrenzung von geistigen? Die Trennung kann vor dem Besitz einer vollständigen Theorie nicht adäquat gezogen werden, doch kann man eine Eingrenzung versuchen.

Beispiele für kognitive Vorgänge sind:

- *Verstehensleistungen* bei Sätzen und Texten, allgemein Sprachverstehen (inklusive des Übersetzens).
- *Klassifikationsleistungen* sowohl genuin semantisch bezüglich von Objekten, Ereignissen und Eigenschaften, als auch wissensbasiert bezüglich dieser in einem Kontext (z.B. Kunstepochen).
- *Problemlösen* in seinen verschiedenen Formen, sowohl bezüglich interner Wissensverarbeitung als auch in externer Umgebungsveränderung; Beispiele sind: eine Glühbirne wechseln, eine Einkaufsliste besorgen, eine Rechnung durchführen.

Involviert in diese kognitiven Vorgänge bzw. das Umgehen mit deren Resultaten ist z.B.

- *Lernen* in verschiedenen Formen bei menschlichen Personen, nicht nur durch Bildung von Assoziationen und Kategorien oder Speichern von Episoden und extrahierter Bedingungen, sondern auch als Erwerb z.B. der Fähigkeit der Erläuterung von Zusammenhängen und dem Lernen an einer anderen Person als Modell.

Sogenanntes ‚Maschinelles Lernen‘⁵ deckt die komplexeren Formen des Lernens nicht ab.

Nicht-kognitive Vorgänge im obigen Sinne sind mutmaßlich:

- ein Gespräch führen

⁴ Vgl. auch Aaron Sloman, "The Irrelevance of Turing Machines to Artificial Intelligence".

⁵ Vgl. z.B. Jörg Frochte, *Maschinelles Lernen*.

- *kreative Tätigkeiten* ohne Regelanleitung, sowohl beim Erstellen von Kunstwerken wie Gedichten und Bildern, als auch beim Entwerfen neuer Theorien
- emotionale Reaktionen

Daneben befinden sich Komponenten des Geistes, die – evtl. wie die gerade angeführten – in Beziehung zur Kognition stehen und deren Rolle zu klären ist. Fraglich ist jeweils

- Lassen sich diese im Symbol-Paradigma der Repräsentation darstellen?
- Lassen sie sich als Komponenten von Algorithmen verstehen?

Zu diesen Komponenten zählen:

- Emotionen (als objektbezogen) und Stimmungen (als nicht objektbezogen)
- Bewusstheit/Selbstbewusstsein als Sich-Innesein des Akteurs des bewussten Erlebens

Bewusstheit/Selbstbewusstsein ist etwas völlig anderes als ein operational definiertes Fokussiertsein-auf bei einem Computersystem mit Sensoren oder als das Verfügen über eine Repräsentation für das System selbst oder Metaprogrammierungsmöglichkeiten bei einigen Programmiersprachen.

- freies Entscheiden

Freies Entscheiden ist mehr als regelgeleitete Auswahl in einem Algorithmus. Architekturen wie SOAR definieren Problemlösen zwar durch u.a. ‚Entscheidungszyklen‘, doch handelt es sich hier nicht um ein Entscheiden im gehaltvollen Sinn eines freien Willens. Freier Wille ist mehr als zufälliges Auswählen (etwa einer Nicht-deterministischen Turing-Maschine [NTM]) und mehr als durch Regeln definiertes Auswählen. In CTC-Modellen (wie in SOAR oder ICARUS) erfolgt die Auswahl, nachdem die Handlungsoptionen mit Präferenzen sortiert oder metrisiert und auf ihre situative Anwendbarkeit geprüft wurden. Gegeben diese Werte oder Anordnung determiniert sich die Auswahl. Eine freie Entscheidung hingegen könnte sich sowohl daran orientieren als auch anders entscheiden. All das erinnert an das traditionelle Problem des ‚Freien Willens‘ und der Inkompatibilität desselben mit einem deterministischen Mechanismus.⁶ Wir haben keine *Theorie* des freien Entscheidens, die für solche Anwendungsfälle eine Regel oder ein Modell nahelegt.

⁶ Im Sinne eines Libertarismus, vgl. z.B. Peter van Inwagen, *An Essay on Free Will*.

Kann eine CTC auch eine Theorie des Versagens einschließen? Etwa:

- Irrationalität
- psychischen Störungen

Pauschal kann man hier Fehlfunktionen in den entsprechenden Prozessen erwähnen, doch haben Irrationalität und psychische Störungen nicht-zufällige Muster und oft eine Geschichte.

§2 Methodik

Eine CTC arbeitet mit Auffassungen von Hardware und Software.

Eine Einbeziehung der Hardware einer Architektur drückt z.T. die Berücksichtigung empirischer Constraints (z.B. der Bearbeitungszeit von Problemen und des Umfangs des Arbeitsspeichers) aus. Abgesehen von solchen allgemeinen Constraints, die funktionaler Natur sind, kann von den Implementierungsdetails einer Architektur im Allgemeinen abgesehen werden (im Sinne eines Funktionalismus).⁷

Die Software-Architektur drückt das Modell bzw. die Theorie aus. Methodisch liegt teilweise ein weiter und teilweise ein enger Begriff von algorithmischer Verarbeitung zugrunde. Im weiten Verständnis – das oft in allgemeinen Behauptungen entsprechender Ansätze auftritt – genügt es, die Bedingungen und Zielzustände der Verarbeitung exakt deklarativ zu erfassen oder in einer Flow Chart zu präsentieren, ohne eine genaue Angabe der Implementation. Gehaltvoller ist die Orientierung an einem präzisen, engen Begriff von Algorithmus. Ein Algorithmus ist eine (imperative) Sequenz von Schritten, diese sind jeweils (d.h. in der Durchführung eines Teil-Algorithmus):

- effektiv (d.h. ‚geistlos‘ in ihrer Durchführung)
- endlich (in Input und Output und Ressourcenzugriff)
- substratneutral

⁷ Wie in §1 erwähnt, geschieht dies zumeist im Rahmen der Unterstützung der Materialismus-These. Verbindet sich diese auch mit einem deterministischen Mechanismus, stellen sich die ebenfalls in §1 erwähnten Verständnisschwierigkeiten, wie sich dies zu unserem Selbstverständnis als bewussten freien Akteuren verhält oder damit kompatibel sein soll.

Dem entsprechen Prozeduren in der Computerprogrammierung. Eine gehaltvolle CTC behauptet solche Algorithmen der Kognition, angestoßen und ausgeführt in einer kognitiven Architektur.

Eine CTM oder CTC hat nur empirischen wissenschaftlichen Gehalt, wenn sie auf den Standardbegriffen der Berechenbarkeit basiert. Der Standardbegriff von Berechenbarkeit (Computation) wird von der Berechenbarkeitstheorie in der Theoretischen Informatik behandelt.⁸ Entscheidend sind hier zwei Befunde:

- (i) Die verschiedenen Erläuterungen des Berechenbarkeitsbegriffes (von TMs und Programmierbaren Random Access Machines [PRAMs] bis zu λ -Kalkül, partiell rekursiven Funktionen usw.) sind äquivalent zu einander (d.h. sie können dieselben Funktionen berechnen und sich wechselseitig simulieren). Dies ändert sich auch dann nicht, wenn Elemente wie Nichtdeterminiertheit [NTMs] oder Fehlerzulässigkeit und Wahrscheinlichkeit (Probabilistische TMs [PTMs]) oder neue Konstruktionsideen (wie Quantencomputer) einbezogen werden. Darin gründet *Churchs These*, dass dieser Begriff der Berechenbarkeit den intuitiven Begriff der Berechenbarkeit erfasst.
- (ii) Das Modell der PRAM findet sich in der von Neumann-Architektur realer Computer wieder.

Man kann in der Theorie weitere Begriffe von Berechenbarkeit definieren („Hyperberechenbarkeit“⁹ oder „Super-Rekursivität“¹⁰) anhand von Modellen abstrakter Maschinen, die jenseits des ‚Turing-Limits‘ (d.h. über den Leistungen all der Standardmodelle hinaus) Probleme lösen können. Beispiele sind Maschinen, die ‚Orakel‘ zu in den Standardmodellen nicht lösbaren Fragen (wie der Entscheidbarkeit allgemeiner prädikatenlogischer Folgerung) konsultieren können (Orakel-TMs) oder etwa über unendliche Zusatzinformationen verfügen, die sie in einem Schritt konsultieren können (Advice-TMs). Materiell lassen sich solche Maschinen *nicht* realisieren.

Eine CTM, die sich auf Hyperberechenbarkeit in der These „Der Geist ist ein Computer“ beruft, muss also den Materialismus aufgeben. Das Gehirn kann nicht eine solche Maschine sein. Die meisten Vertreter einer CTM behaupten hingegen die Materialismus-These.

⁸ Vgl. z.B. Nigel Cutland, *Computability*.

⁹ Vgl. Apostolos Syropoulos, *Hypercomputation*, oder: Jack Copeland, „Hypercomputation“.

¹⁰ Vgl. Mark Burgin, *Super-recursive Algorithms*

Eine CTM oder CTC, die sich weder auf (Standard-)Berechenbarkeit noch auf Hyperberechenbarkeit festlegt, bleibt eher unklar und immunisiert sich gegen Kritik, welche an den Grenzen dieses Berechenbarkeitsbegriffes ansetzt. Sie kann dazu dienen, kognitive Leistungen als regelbasiert zu erläutern und knüpft so an philosophische Analysen an. Sie kann eine (vagere) Heuristik sein, die allerdings weniger empirisch überprüfbare Aussagen über die Kognition machen kann. Sie entspricht – auch wenn *de facto* gelegentlich in dieser unklaren Weise verfahren wird – nicht den proklamierten Ansprüchen vieler Vertreter einer CTM oder CTC.

§3 Semantische Maschinen

Die Semantik von Symbolen im Allgemeinen geht zurück auf die Bedeutung von Sätzen, d.h. die Semantik einer natürlichen Sprache. Die Bedeutung von Sätzen hängt zusammen mit der Wahrheit von Behauptungssätzen und der Wahrheit der in ihnen ausgedrückten Meinungen. Originär ist dieser Wahrheitsbezug und -anspruch von Behauptungen/Behauptungssätzen und Meinungen. Darin liegt originärer Wirklichkeitsbezug, originäre Intentionalität, von der die Intentionalität von Repräsentationen abgeleitet ist. Eine Repräsentation hat Intentionalität/Semantik, weil sie Teil einer Meinung ist.

Meinungen zu haben, die wahr oder falsch sind und als solche unterschieden werden können, bedarf des Verfügens über den Begriff der Meinung, also einer höherstufigen Intentionalität (einer ‚Theory of Mind‘).¹¹ Damit einhergehen Normen der Rationalität beim Vorbringen und Überprüfen von Behauptungen und Meinungen. Meinungen im eigentlichen Sinne können daher nur rationale Wesen besitzen, was Vorstufen von Meinungen und Intentionalität bei Tieren nicht ausschließt. Insbesondere sind dies Normen einer Sprachgemeinschaft (d.h. sowohl Normen der sprachlichen Verständlichkeit und Wohlgeformtheit als auch epistemische Normen der Begründung). Bei diesen Leistungen geht – irgendwie – Bewusstheit und Selbstbezug ein, ebenso das Einnehmen von Sprecherrollen mit der Übernahme der entsprechenden Verantwortung für Äußerungen. Intentionalität haben Personen, die sich für ihre epistemischen und praktischen Urteile und entsprechende Handlungen rechtfertigen können und vor anderen müssen. In Urteilen, Deliberation, Reflektion und entsprechendes Handeln geht freies Entscheiden ein. Für solche Urteile –

¹¹ Vgl. z.B. Donald Davidson, „Rational Animals“, sowie allgemein: Michael Corballis, *The Recursive Mind*. Vgl. zum Folgenden auch: Manuel Bremer, *Philosophische Semantik*.

nicht für das, was einem zustößt – kann und muss man sich rechtfertigen, zumindest immer ‚im Prinzip‘, wenn entsprechender Begründungsbedarf (von anderen oder in der Selbstreflexion) angemeldet wird.

Die bis jetzt hergestellten Computer und Roboter sind keine Personen. Daher besitzen sie *für sich* auch keine Intentionalität oder Semantik. Sie sind jedoch *semantische Artefakte*.

Ein Algorithmus kann auf einer Ebene der Abstraktion beschrieben werden als eine Sequenz von Symbolmanipulationen, d.h. als ein syntaktisches Objekt. Sogar ein laufender Algorithmus kann so beschrieben werden. Ein laufender Algorithmus ist zugleich in einem Substrat implementiert und damit bezüglich einer solchen Beschreibung nicht nur eine Abstraktion, sondern eine syntaktische und kausale Sequenz. Ein laufender Algorithmus ändert die Wirklichkeit (zumindest im implementierenden Substrat) in einer Weise, die der definierten Implementation seiner Teile korrespondiert. Darin besitzt ein laufender Algorithmus eine (prozedurale) Semantik, nicht allein auf der abstrakten Ebene einer Modelltheorie, sondern im Verknüpfen seiner Teiloperationen mit Veränderung in der Wirklichkeit als deren Bezugnehmen auf die Wirklichkeit. Ein laufender Algorithmus besitzt dermaßen einer Semantik und der abstrakte Algorithmus wurde abstrahiert von Symbolen, die eine Semantik haben.

Damit ein Algorithmus aus Symbolen besteht, benötigen diese seine Teile eine Semantik. Von dieser Semantik wird in einer syntaktischen Beschreibung oder Untersuchung des Algorithmus abstrahiert. Man kann nicht von ‚Computation‘ oder ‚Berechnung‘ reden ohne Semantik, denn der Begriff der Repräsentation ist für den Begriff des Symbols unverzichtbar. Man kann allerdings Berechenbarkeit in Abstraktion von der Semantik der Symbole untersuchen. Dies geschieht auch in der meta-logischen Untersuchung formaler Sprachen und Kalküle und der theoretischen Informatik. Das Interesse an den entsprechenden meta-logischen Theoreme über ein Logiksystem speist sich indessen aus deren Anbindung an eine intendierte Semantik (informell anvisiert oder über entsprechende Korrektheitsbeweise angebunden).

Die unmittelbare Implementation eines Algorithmus kann sich, des Weiteren, fokussieren auf die kausale Implementation der syntaktischen Strukturen, wobei diese Implementation die Semantik der implementierten Symbole erbt.

Schreibt man einem System zu, einen Algorithmus auszuführen, werden diese Schritte in umgekehrter Reihenfolge betrachtet: man extrahiert eine syntaktische Beschreibung aus einer Interaktion von System und Wirklichkeit.

Ein Computersystem ist so konstruiert, dass dessen interne Struktur und die Operationen, die über diese definiert sind, eine Bedeutung haben (seien es numerische Resultate oder beispielsweise die Herstellung einer Netzwerkverbindung zum Bilderaustausch). In diesem Sinne sind Computer *semantische Artefakte*.

Diese Bedeutung existiert – natürlich – nur für die konstruierenden Personen und nicht für den Computer selbst.¹² Die kausalen Interaktionen des Computers sind Bestandteile der Wirklichkeit und selbst zunächst ohne Bedeutung. Die Arbeitsweise des Computers kann allerdings beschrieben und interpretiert werden auf abstrakten Ebenen als Sequenz semantischer und syntaktischer Prozeduren. Die Bedeutung ‚in‘ Computer leitet sich genauso von der Semantik der sprechenden Personen ab, die den Computer konstruieren, wie die Bedeutung ‚in‘ einem Buch. Beide sind artefizielle Bedeutungsträger. Im Unterschied zum passiven Buch, dessen Bedeutung nur im Lesen durch und für eine Person realisiert wird, realisiert der Computer Bedeutung in seiner eigenen Aktivität. Darin geht die Konstruktion und Implementation von Algorithmen über das Schreiben von Büchern hinaus.

Eine Idee hinter einer CTC lässt sich verstehen als der Versuch, menschliche Kognition in derselben Weise zu betrachten: Es gibt eine erhellende *abstrakte* Beschreibung der Kognition als Computation. Ob menschliche Kognition Computation im Sinne einer solchen CTC ist, hängt ab davon (i) wie sich das Modell an empirische Forschung zur menschlichen Kognition anbinden lässt (etwa bezüglich von Verarbeitungszeiten im Kurzzeitgedächtnis oder Phänomenen wie ‚Prompting‘)¹³ und (ii) wie sich Modelle höherer kognitiver Leistungen (d.h. solcher, die bewusst und von der allgemeinen Kognition her zugänglich sind) an introspektive Befunde anbinden lassen.

Eine solche Betrachtung ist zunächst eine *Interpretation* und eine entsprechende Modellbildung. Aber eine Modellbildung, die aus einer Interpretation hervorgeht, kann trotzdem das Betrachtete adäquat erfassen. ‚Interpretation‘ heißt nicht ‚Verfälschung‘. Die

¹² Das ist m.E. der richtige Teil von John Searles Kritik jeglicher CTM oder CTC (vgl. John Searle, *The Rediscovery of the Mind*).

¹³ Vgl. z.B. Allen Newell, *Unified Theories of Cognition*, S.222-34, zu einer eher positiven Einschätzung diesbezüglich und SOAR, kritisch zu SOAR z.B. Richard Cooper, "Cognitive Architectures as Lakatosian Research Programmes: Two Case Studies".

Erfassung von Algorithmen durch eine Interpretation eines komplexen Systems heißt nicht, dass es sich nicht um ein Algorithmen abarbeitendes System handelt.¹⁴

§4 Modulare Prozesse

Die genuine Intentionalität involviert das Bewusstsein einer Person („Geist“ im engeren Sinne). Aber nicht alle mentalen Zustände und Prozesse sind bewusst. Auch die nicht bewusste Mentalität nutzt die genuine Mentalität, die sich im Bewusstsein gründet. Solche Prozesse kann – und evtl. muss – man modellieren unter Absehung von Bewusstsein und dessen Rolle. Hier können algorithmische Modelle eine Rolle spielen. Ungeklärt muss dabei bleiben – zumindest im Detail – wie dann die Übergänge zwischen bewussten und nicht-bewussten Vorgängen erfolgen. Eine CTC kann hier eine Option einer partiellen Theoriebildung sein.

Auch im Bewusstsein werden Regeln verwendet und Algorithmen abgearbeitet. Regelbefolgen und Begründen involvieren Prozeduren, für die es algorithmische Modelle geben kann. Insofern kann eine CTC Teil der Explikation und Erläuterung der kognitiven Vermögen von Personen sein.

Eine CTC hat die besten Erfolgsaussichten bei der Modellierung von Fähigkeiten, die nicht mit Bewusstheit und Deliberation verbunden sind (wie dem unmittelbaren Klassifizieren von Entitäten). Eine Reihe solcher „modularen“ Fähigkeiten hängen wenig bis gar nicht von allgemeinen kognitiven Leistungen wie der Aufrechterhaltung eines kohärenten Meinungssystems ab. Damit stellt sich bei diesen nicht das paradigmatische „frame problem“, was alles zu berücksichtigen und im Kontext zu reflektieren ist.

Eine solche Konzentration auf solche sehr speziellen Fähigkeiten bietet sich methodisch an, muss dann indessen mit dem Eingeständnis verbunden werden, keine allgemeine computationale Theorie des Rasonierens und der Kognition zu liefern.¹⁵

Allen Newell als Pionier der KI und Theoretiker einer paradigmatischen CTM gebührt der Verdienst, seine konkrete Modellierung und Implementation einer CTM (deren letzte Variante die Frühformen des SOAR-Projekt waren) an psychologische experimentelle Daten

¹⁴ Auch für eine CTC widerspricht eine Theoriebildung mittels nicht (von außen eines Untersuchungsbereiches) direkt beobachtbarer Entitäten nicht einer realistischen (an Stelle einer instrumentalistischen) Auffassung von Theorien.

¹⁵ Vgl. Jerry Fodor, *The Mind Doesn't Work That Way*.

anzubinden.¹⁶ Solche Daten allerdings stammen aus Laborsituationen, die kleinteilige kognitive Leistungen messbar operationalisieren (etwa das Abtippen eines Textes, das Reagieren auf assoziative Verknüpfungen von Symbolen und Bildern, usw.). Die Operationalisierung solcher Experimente legt eine Übersetzung in Flow Charts und Algorithmen nahe, was bei alltäglichen geistigen Leistungen in Situation nur unter Abstraktion von vielen Umständen möglich ist. Ein Modell des Problemlösens, das an den experimentell operationalisierten und wohl definierten Problemen ansetzt, hat seine Anwendungen, erstreckt sich jedoch nicht zwangsläufig auf das allgemeine Umgehen mit den Herausforderungen des Alltags und dem Gestalten der eigenen Biographie.

Ein ebenso der Modellierung besser zugänglicher Bereich sind alle Fertigkeiten, die in der Form entweder deduktiven Folgerns oder der Suche nach Modellen (im Sinne von Variablenbelegungen) dargestellt, formalisiert und (z.B. in der Logikprogrammierung) implementiert werden können.¹⁷ Solche Fähigkeiten sind wesentlich algorithmisch umsetzbar, sie machen den Kern des intuitiven Begriffs der Berechenbarkeit aus. Damit sind sie zugleich allerdings – haben wir einmal einen Algorithmus entdeckt – auch wesentlich geistlos durchführbar.

Interessanter bezüglich der allgemeinen Intelligenz ist nicht das Befolgen, sondern das Entdecken und Entwerfen solcher Algorithmen und die Überprüfung ihrer Korrektheit/Angemessenheit bezüglich der betrachteten Aufgabenstellung. Diese Leistungen sind nicht algorithmisch.

§5 Einschätzungen und Beurteilungen

Die meisten Klassifikationsleistungen erfolgen problemlos im Hintergrund, etwa wenn wir uns in einer Situation bewegen und deren Objekte und Ereignisse in Kategorien einordnen. Auch sprachliche Fassungen solcher Urteile sind unproblematisch: „Der Tisch wackelt“, „Es ist kurz vor 3“, „Die Straßenbahn hat ein neues Werbebanner“ etc.

Wichtig sind allerdings ebenso Prozesse des Beurteilens und Einschätzens, die Überlegung und Erfahrung einschließen und keine bloße Klassifikation im obigen Sinne liefern. Solche Prozesse können in mehreren Arten von Fällen angestoßen werden:

¹⁶ Vgl. Allen Newell, *Unified Theories of Cognition*.

¹⁷ Vgl. Robert Kowalski, *Computational Logic and Human Thinking*.

- im Falle vager Kategorien: „Ist das noch Hard-Rock oder schon Heavy Metal?“
- im Falle von Erwartungen des Handelns anderer Personen: „Sind die Anhänger von Mike McConnell eine Gefahr für die nächsten demokratischen Wahlen?“
- im Falle des Verstehens des Handelns anderer Personen: „Hat Susanne wirklich keine Lust mehr auf die Reise?“
- im Falle von ästhetischen Evaluationen: „Soll die Vase hier stehen oder weiter rechts?“
- im Falle der Anwendung komplexer Normen: „Weisen die Indizien darauf hin, dass es sich um einen Mord aus niederen Motiven handelt?“
- im Falle der Einschätzung von Entwicklungen: „Ein allgemeines algorithmisches Modell zur Herstellung von Kohärenz einer Menge von Meinungen ist bis auf Weiteres nicht zu erwarten.“

und anderer mehr.

Die Problematik erinnert an die traditionelle Problematik der ‚Urteilkraft‘.¹⁸ Für den Erwerb von Urteilkraft in diesem Sinne besitzen wir kein Rezept. Eine Simulation solcher Beurteilungen in einem Computerprogramm (etwa von einer Liste der Indizien bis zu einer Operationalisierung des Ablesens ‚niederer Motive‘) persifliert die Beurteilungssituationen, in denen sich Personen wiederfinden. Dies beginnt schon mit der für die Simulation nötigen Einschränkung der Situationsmerkmale und des benötigten Kontextwissens über Situationen dieses Typs.

Genauso wenig wie im Falle des freien Entscheidens besitzen wir eine (detaillierte) Theorie, was dieses Einschätzen und Beurteilen ausmacht: es verlangt eine umfassende Situationsbeschreibung, bei verfügen über hinreichende sprachliche Kompetenz bezüglich solcher Beschreibungen, unter Einbeziehung relevanten Kontextwissens – aber schon hier treten wieder Einschätzungen auf, was ‚relevant‘ und ‚hinreichend‘ im betrachteten Kontext ist. Die Beurteilung einer Situation kann weit auf Wissensbestände und Erinnerungen zugreifen. Es gibt keine allgemeinen formalen Bedingungen, ein Beurteilungsproblem zu spezifizieren. Beurteilen macht indessen einen wesentlichen Teil des geistigen, sozialen und

¹⁸ Vgl. z.B. Klaus Günther, *Der Sinn für Angemessenheit*.

kognitiven Lebens von Personen aus. Diesen Bereich der allgemeinen Kognition und menschlicher Intelligenz kann sich eine CTC nicht erschließen.

§6 Einschränkungen einer CTM/CTC

Die vorangehenden Überlegungen können zu drei Hypothesen führen:

(Hypothese 1) Eher ist eine Computationale Theorie der Kognition möglich, aber nicht eine des Geistes im Allgemeinen.

(Hypothese 2) Eine Computationale Theorie der Kognition kann eine Heuristik für eine Theorie der Kognition sein, auch wenn die Kognition kein Computer ist.

(Hypothese 3) Eine Computationale Theorie der Kognition lässt die Stellen hervortreten, an denen der Geist über eine algorithmische Kognition hinausgeht (bzw. wie weit der Geist mit Algorithmen arbeitet).

Manuel Bremer, 2021.