

#### 4 Algorithmen als Bestätigung intuitiver, heißt richtiger Entscheidungen

Das Gehirn ist bei Christian und Griffiths schon lange nicht mehr Ausführungsorgan algorithmischer Abläufe – hier werden weniger die neuronalen Prozesse beschrieben als vielmehr die Logik alltäglicher, sozialer Entscheidungen unter den Prämissen von Statistik, Mittelwerten und Wahrscheinlichkeiten. Das algorithmische Gehirn steht symbolisch für eine längst vollzogene Übertragung mathematischer, stochastischer Logik auf das Gehirn, nicht nur als konkrete Erklärung für die neuronale Funktionsweise, sondern auch als Metapher, als sinnstiftende Bezeichnung dessen, was den Menschen und seinen Geist ausmacht. Die Mathematisierung der Wahrnehmung beschreibt, wie durch die Implementierung von Stochastik und Wahrscheinlichkeit in die wissensproduzierenden Technologien diese Verfahren auf das Gehirn zurückübertragen werden: In der Folge sind es nun also Stochastik und Wahrscheinlichkeit, die als mathematische Grundgesetze auch über unsere bewussten und unbewussten Entscheidungen herrschen.

In der Vorstellung algorithmischer Gehirne steckt indes nur eine stark vereinfachte Definition algorithmischer Funktionsweisen. Mit dem Verweis, dass Algorithmen deutlich älter sind als ihre Anwendung im Computer, wird der Minimalkonsens einer einfachen schrittweisen Anleitung verifiziert und enttechnologisiert. Die stark reduzierte und für alle verständliche Algorithmusdefinition dient als Brücke, um die regelgeleitete Abfolge von Arbeitsschritten zu verallgemeinern, um sie als Vorgehensweise für – fast – alle nicht ›chaotischen‹ und effizienten Durchführungen von Prozessen und Entscheidungsfindungen zu beschreiben. Auf dieser Basis lassen sich alle Abläufe als natürlich und programmierbar fassen, und die enttechnologisierte und gleichsam naturalisierte Definition dient überdies als Metapher, um sie unter die Logik einer regelhaften Abfolge zu subsumieren und nachvollziehbar werden zu lassen. In einem nächsten Schritt soll der Algorithmus nicht mehr allein die Arbeitsweise neuronaler Prozesse einfangen, sondern dient dazu, algorithmische Vorgehensweisen als originär dem ›gesunden Menschenverstand‹ inhärent zu betrachten. So gelingt es, auch vermeintlich intuitive Entscheidungen als vernunftgeleitet zu beschreiben. Damit werden zum einen die Spezifik von in Computern verwendeten mathematisch verfassten Algorithmen und ihre epistemischen Auswirkungen negiert. Durch das Unsichtbarmachen ihrer historischen Gewordenheit werden Algorithmen essenzialisiert und zu einer mathematischen Regelhaftigkeit der Natur stilisiert. Zum

anderen wird mit dieser Universalisierung vergessen, dass die ›schrittweise Verarbeitung‹ in der Anwendungsdefinition computergestützter Berechnungen durch Algorithmen nicht vollständig beschrieben wird. Algorithmen, so die Kritik (Stiller 2015, 63), sind eben nicht mit einem Rezept und der schrittweisen Bearbeitung von Handlungsanweisungen vergleichbar.

Der Mehrwert des *algorithmischen Gehirns* liegt in seiner Verallgemeinerbarkeit: Das Prinzip Algorithmus wird durch die vorher erfolgte Verknüpfung von Geist und Mechanik beziehungsweise mathematischer Prozessverarbeitung auf die bewussten und unbewussten Entscheidungsfindungen des Menschen übertragen. Das in diesem Konzept verwendete, stark reduzierte Algorithmusverständnis, das als Hintereinanderarbeiten von Lösungsschritten zusammengefasst werden kann, gilt in den Neurowissenschaften jedoch inzwischen als überholt und wurde durch die Modellierung rekurrenter Algorithmen abgelöst. Die gegenwärtigen Neuronenmodelle zum Verständnis künstlicher wie biologischer neuronaler Netzwerke unterliegen der (mathematischen) Logik der Selbstorganisation: Die Selbstorganisation der neuronalen Netzwerke ist den Systemen immanent und läuft durch die Festlegung einer programmierten Kausalmatrix mit zugehörigen Kausalzusammenhängen, dem eigenständigen Erkennen von Fehlern und dem iterativen Wiedereinspeisen der Daten ohne Intervention von außen ab. Neuronale Netzwerke, so die Annahme, benötigen keine externen Lehrer mehr, alle Fehlererkennung, alle darauf basierenden Vorhersagen und Entscheidungen kann das System aus sich heraus generieren, und die Gesamtheit von Welt mit ihren kausalen Zusammenhängen ist in die (mathematische) Logik des Systems eingelagert. Dieses Verständnis von Gehirnfunktion wurde auf die Formel *predictive brain* gebracht, um das es im folgenden Unterkapitel gehen wird.

#### 4.1 Vorhersagendes Gehirn: Vorher wissen, was als Nächstes kommt

Die Mathematisierung der Wahrnehmung hat zu einer Universalisierung epistemologischer Anforderungen bei der Bewältigung der Datenanalyse geführt (s. hierzu Kap. 4). Problemlösungen, auch in den Computational Neurosciences, sind prinzipiell zu mathematischen Herausforderungen geworden. Meist handelt es sich dabei um Probleme, die sich allgemein auf Fragen der Stochastik, der Komplexität und der Machbarkeit algorithmischer Berechnungen in einer Polynominalzeit beziehen (die Polynominalzeit gibt den Zeitrahmen vor, in der sich Komplexitätsprobleme berechnen lassen müssen, siehe Kap. 3). Die größere Herausforderung für stochastische

Prozesse sind heute Probleme, für deren Lösungsweg die Regeln nicht klar festgelegt sind, einige der benötigten Informationen fehlen oder das Finden der genau richtigen Antwort die Berücksichtigung einer astronomischen Anzahl von Möglichkeiten erfordern würde. Für diese Fälle müssen Wahrscheinlichkeitsrechnungen gefunden werden, die so einfach wie möglich gestaltet sind (um ihre Berechnungszeit im durchführbaren Rahmen sicherzustellen) und gleichzeitig ausreichend Komplexität simulieren.

Die Algorithmen, die in den letzten Jahren entwickelt wurden, um die schwierigsten mathematisch-logischen Klassen von Problemen zu lösen (wie etwa die Frage: Ist die Berechnung eines Algorithmus auch für biologische neuronale Netze durchführbar?), haben dazu geführt, dass die Algorithmen zur rechnerischen Bewältigung realer Aufgaben mit Wahrscheinlichkeitswerten, dem Zufallsprinzip sowie mit Näherungswerten ausgestattet wurden (vgl. Christian/Griffiths 2016, 8). Diese Verknüpfung soll mithilfe stochastischer Berechnungen und Simulationen erreicht werden, denn, so die Hypothese, Stochastik oder Zufälligkeit ist in der Natur allgegenwärtig und scheint in der Biologie noch stärker ausgeprägt zu sein als in der Physik. Dieses in den Algorithmen verwendete Prinzip wird auf die Modelle und Simulationen (biologischer) neuronaler Netze übertragen und damit gleichgesetzt. Das algorithmische Gehirn wird durch die Verwendung von vorhersagenden Algorithmen zum vorhersagenden Gehirn (*predictive brain*, selbst ein Algorithmus). Die Computermodelle, die zunächst der Untersuchung Neuronaler Netzwerke dienten, werden somit zum allgemeingültigen Erklärungsmodell morphologischer Funktionsweise des Gehirns erklärt, sodass es zu einer Gleichsetzung der neuronalen Strukturen mit den Werkzeugen, mittels derer Wissen über diese Strukturen erzeugt wird, kommt.

Das vorhersagende Gehirn kann als neurobiologische Aktualisierung des algorithmischen Gehirns unter der Ägide der Computational Neurosciences und der Neuroinformatik gefasst werden. Das algorithmische Gehirn steht noch unter dem Eindruck einer symbolisch getriebenen Kognitionsverarbeitung, das vorhersagende Gehirn geht dagegen von konnektionistischen Prämissen aus, von der Komplexitätstheorie und selbstlernenden, vorhersagenden Neuronalen Netzwerken, die aus der KI-Forschung in die Vorstellung vom denkenden Gehirn übertragen werden. Diese Entwicklung lässt sich knapp zusammenfassen mit: vom Reduktionismus zur Komplexität, zur Simulation durch Wahrscheinlichkeit. Das Gehirn prozessiert demnach Information, indem es Vorhersagen vornimmt. Der Ansatz des *predictive brain* versucht in seinen Modellen verstärkt die Biologie mit den Vorannahmen aus

der Physik und Physiologie zu verbinden und für die Computational Neurosciences und die Neuroinformatik nutzbar zu machen. Diese Verknüpfung soll mithilfe stochastischer Berechnungen in den Computern und Simulationen erreicht werden. Gleichzeitig werden funktionsmorphologische Modelle, das heißt die neuronalen Netzwerke des menschlichen Gehirns, gleichgestellt mit den Eigenschaften ihrer Namensvetter, den künstlichen neuronalen Netzwerken, die dadurch fit gemacht werden sollen für die technischen Herausforderungen der kommenden Jahre. Computational Neurosciences und die Neuroinformatik sind die vielversprechendsten Neurodisziplinen, um die immense Datenflut unter Kontrolle zu bekommen und gleichzeitig mittels dieser Daten Bedeutung zu generieren:

Neuroscience is a multidisciplinary effort to understand the structures and functions of the brain and brain-mind relations. This effort results in an increasing amount of data, generated by sophisticated technologies. However, these data enhance our descriptive knowledge, rather than improve our understanding of brain functions. This is caused by methodological gaps both within and between subdisciplines constituting neuroscience, and the atomistic approach that limits the study of macro- and mesoscopic issues. Whole-brain measurement technologies do not resolve these issues, but rather aggravate them by the complexity problem. (Kotchoubey et al. 2016, 1)

Der Kognitionsphilosoph Andy Clark, ehemaliger Professor für Logik und Metaphysik an der Universität von Edinburgh, beschreibt in seinem 2013 erschienenen Paper *Whatever next? Predictive brains, situated agents, and the future of cognitive science* einen neuen Ansatz kortikaler Berechnungen, dem er den Namen »action-oriented predictive processing« (handlungs- oder aktionsorientiertes vorausdeutendes Prozessverfahren) gibt. Anhand dieses Konzepts, so die Hoffnung, sollen einerseits die Computational Neurosciences und die Neuroinformatik mit der Kognitionswissenschaft und der Philosophie des Geistes verbunden werden und andererseits sollen damit mehrere Bereiche der kortikalen Verarbeitung von Wahrnehmung, Kognition und Motorik erklärt werden. Clarks 2013 erschienenes Paper hat viel Zuspruch, aber auch Widerspruch hervorgerufen, beides jeweils aus dem eigenen Feld. Gleichzeitig spiegeln sich in Clarks Konzept des *predictive processing framework* eine Vielzahl der Vorannahmen wider, die sich im Feld der Computational Neurosciences und der Neuroinformatik finden lassen.

Aus diesem Grund stelle ich das Konzept des *predictive brain* hier vor mit dem Verweis, dass es als Chiffre für aktuelle Sichtweisen auf die Funktionsweise Neuroner Netzwerke in den Computational Neurosciences und weniger als die genaue Ausformulierung der abgeleiteten Erkenntnisse verstanden werden sollte. Einige der von Clark beschriebenen Begriffe und Funktionsweisen Neuroner Netzwerke, etwa neuronale Kaskaden, versteckte Entscheidungsebenen, Backpropagation sowie Top-down- und Bottom-up-Verfahren, sind der Programmierung artifizieller neuronaler Netze entlehnt. Erneut kommt es hier zu einem Zirkelschluss, in dem Mensch und Maschine beziehungsweise heute eher Algorithmen sich wechselseitig inspirieren und zum Modell füreinander werden. Das *predictive brain* kann somit als Übertragung algorithmischer Deep-Learning-Strukturen auf biologische neuronale Reizverarbeitung verstanden werden. Das Ziel, das mit dem Konzept des *predictive brain* verbunden wird, ist es, möglichst allgemeine Vorgehensweisen der Fehlerkorrekturen zu ermitteln und anhand dieses Modells die Wahrnehmung, die Kognition und den Bereich motorischer Bewegungen zu erklären.

Das Gehirn, so die neueste, von Deep-Learning-Algorithmen beeinflusste Erkenntnis, verfügt über die Möglichkeit, Fehlerkorrekturen mittels versteckter Prozesse in die Vorhersagen unterer neuronaler Ebenen zu reimplementieren. Clark bezeichnet diese Prozesse auch als *dumb processes*.<sup>1</sup> Systeme höherer Ebenen versuchen, die Eingaben für Systeme niedrigerer Ebenen auf der Grundlage ihrer eigenen entstehenden Modelle der kausalen Struktur der Welt (d.h. der Signalquelle) vorherzusagen. Fehler bei der Vorhersage von Eingaben der unteren Ebene führen dazu, dass die Modelle der höheren Ebene angepasst werden, um die Diskrepanz zu verringern. Als Vorlage für dieses Modell der »versteckten« Fehlerkorrektur zieht Clark die Helmholtz-Maschine heran, die programmiert wurde, um neue Repräsentationen mithilfe eines mehrstufigen Systems zu erlernen, indem nach tieferen Regelmäßigkeiten innerhalb der einzelnen Bereiche gesucht wird, ohne dass eine Vielzahl klassifizierter Muster der gewünschten Input-output-Zuordnung erforderlich ist.

Such models follow Helmholtz (1860) in depicting perception as a process of probabilistic, knowledge-driven inference. From Helmholtz comes the key idea that sensory systems are in the tricky business of inferring sensory causes from their bodily effects. This in turn involves computing multiple

---

1 Interessant an dem Ausdruck ist die Nähe zu Marvin Minskys unintelligenten Agenten. Der Ausdruck lässt sich außerdem als »unbewusste Prozesse« verstehen.

probability distributions, since a single such effect will be consistent with many different sets of causes distinguished only by their relative (and context dependent) probability of occurrence. (Clark 2013, 18f.)

Die Helmholtz-Maschine tat dies, indem sie eigene Top-down-Verbindungen nutzte, um die gewünschten Zustände für die versteckten Einheiten vorherzusagen, und so die Entwicklung seines Wahrnehmungs- sowie gleichermaßen Erkennungsmodells selbst überwachte.

Ein sich selbst erzeugendes, generatives Modell zielt darauf ab, die statistischen Muster einer Reihe von beobachteten Eingaben zu erfassen, indem es die für diese Struktur verantwortliche Kausalmatrix schematisch wiederholt:

In practice, this means that top-down connections within a multilevel (hierarchical and bidirectional) system come to encode a probabilistic model of the activities of units and groups of units within lower levels, thus tracking (as we shall shortly see in more detail) interacting causes in the signal source, which might be the body or the external world. (Ebd., 182)

Diese Modelle kombinieren ein generatives, hierarchisches Top-down-Modell, in dem die Signalverarbeitung für das gesamte System mittels einer Kausalmatrix vorhergesagt wird, mit einem Modell, das unerwartete Abweichungen vom ersten Modell, also Vorhersagefehler der erwarteten Signalverarbeitung, zurückmeldet, »leaving only any residual ›prediction errors‹ to propagate information forward within the system« (ebd., 181). Das zunächst für die Komprimierung von Bilddaten erarbeitete Verfahren verringert den Informationsgehalt, für den Vorhersagen vorgenommen werden sollen, bei Bildern auf ein Pixel und stellt dieses in Abhängigkeit zu der benachbarten Information, also dem Nachbarpixel, in dem der Wert eines Pixels ins Wahrscheinlichkeitsverhältnis zum Wert seines nächsten Nachbarn gesetzt wird. Nur wenn der Wert des benachbarten Pixels nicht dem erwarteten Ergebnis entspricht, kommt es zu einer Fehlermeldung, die in das System zurückgemeldet wird. Das bedeutet, dass der Code für ein reichhaltiges Bild komprimiert werden kann – zumindest wenn ein ausreichend informierter Empfänger den Code wieder zurückübersetzen kann –, indem nur die »unerwarteten« Abweichungen kodiert werden, also die Fälle, in denen der tatsächliche Wert vom vorhergesagten abweicht (vgl. ebd., 182). Das Modell des *predictive brain* ähnelt von seiner Grundidee her sehr stark den Modellen künstlicher Deep-Learning-Algorithmen. Diese werden auf die neuronale Reizverarbeitung eins zu eins übertragen und zu einem universellen Prinzip erklärt. Clark fasst die neue

Theorie der vorhersagenden Strukturen als allgemeingültiges Programm zusammen, in das verschiedene neuronale Strukturen des Gehirns involviert sind und dessen Vorhersagen unsere Wahrnehmung, unser Empfinden und unsere Handlungen bestimmen: »As strange as it sounds, when your own behaviour is involved, your predictions not only precede sensation, they determine sensation. [...] Thinking, predicting, and doing are all part of the same unfolding of sequences moving down the cortical hierarchy.« (Clark 2013, 186)

Hier schließt sich der Kreis: Das vorhersagende Gehirn lässt sich hervorragend modellieren und simulieren, baut es doch auf denselben Prämissen auf wie Deep-Learning-Programme, nämlich denen der zufälligen, aber vorhersagbaren Ereignisse. Und solange ein Ereignis nicht linear, sondern zufällig ist, braucht es nur einen Auslöser, um zum nächsten Ereignis zu führen. Um welchen Auslöser es sich dabei handelt, ist in einem System, in dem Anfangs-, Randbestimmungen und Ähnlichkeitswerte durch eine Kausalmatrix definiert sind, unwichtig: »Because as long as it's random, I just need a thing which jumps me into the next thing. I don't really think it matters much. And random is all the same. I suspect actually simulating exactly the right sort of chaos doesn't really matter very much.« (Hutton 2022 [in Druck])

## 5 Schlummernde Vorlieben – von der Automatisierung des Denkens zum informierten Fühlen von Fakten

Im Anschluss an die kritischen Betrachtungen der instrumentellen Vernunft (s. Kap. 3) fängt das Konzept der Mathematisierung der Wahrnehmung die Entwicklungen der Neurowissenschaft der letzten 20 Jahre ein. Mit Mathematisierung ist die Übertragung neuronaler Abläufe in die Logik mathematischer Computermodele, Algorithmen und Simulationen gemeint, ebenso wie die metaphorische Verwendung einer mathematischen Regelmäßigkeit und algorithmischer Prozessualität von Denkabläufen. Auch wenn das algorithmische Gehirn, zumindest in der Auslegung von Christian und Griffiths (2016), kein originäres Beispiel aus den Computational Neurosciences ist, verdeutlicht es die Ausstrahlung der Mathematisierung der Wahrnehmung auf andere wissenschaftliche Disziplinen und Herangehensweisen. Die gegenwärtigen Neuronenmodelle können mit dem Konzept der Mathematisierung der Wahrnehmung auf die mathematische Verfasstheit ihrer epistemischen Bedingungen hin befragt werden. Das Konzept der Mathematisie-