CLB

- VR, KI und Chemieausbildung
- VR mit Biomolekülen
- ChatGPT und Studienzulassung
- GPT-4 und Analytische Chemie
- Grundlagen der Sprachmodelle

- Bitterstoffe
- Rollreibung bei Eis und Schnee

05 - 06 / 2023

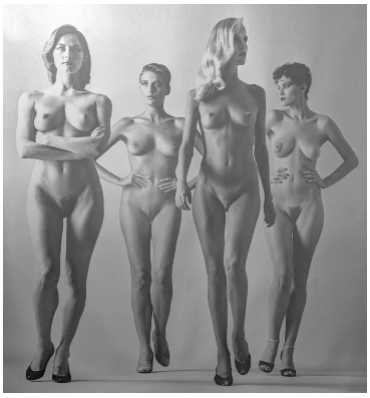
D 2046 E

Wieviel Wissen ist nötig, um denken zu können?

Die Frage der Überschrift lässt sich einfach beantworten: Keins. Babies haben kein Wissen, können aber denken, d.h. Umwelteindrücke verarbeiten, darauf reagieren. Es liegt eben schon eine „Grundverdrahtung" der Neuronenschaltkreise im Gehirn vor, die Denkvorgänge ermöglicht. Es gilt also, die Eingangsfrage anders zu stellen: *Wieviel Wissen ist nötig, um schnell denken zu können?*

Die Evolution hat uns Menschen mit der Entwicklung des Bewusstseins einen Vorteil geschaffen. Zwar werden Entscheidungen im Gehirn getroffen, ehe sie uns bewusst werden. Das ist besonders hilfreich, wenn es lebensnotwendige Entscheidungen sind. Die trifft das Mittelhirn in Sekundenbruchteilen; man nennt es auch gerne Reptiliengehirn – oder neudeutsch „Croc Brain", weil es sich schon in der frühen evolutionären Phase der Reptilienentstehung ausbildete. Zu entscheiden war nur: Ist es gefährlich – dann weglaufen, oder kann ich es fressen?

Entscheidende Rückkopplungen: Viele Einzeleindrücke und -entscheidungen jedoch formt das Gehirn mit Hilfe des Bewusstseins zu Modellen und Szenarien *künftiger* Entwicklungen, die helfen, in einer komplexen Umwelt besonders vorteilhaft handeln zu können. Auf diese Weise gibt es durch Lernverhalten wiederum Rückkopplungen auf das Gehirn und seine neuronale Verschaltung, die dann bei unbewussten Entscheidungen zu neuen Lösungen führen können. Diese zusätzliche Rückkopplung rettet zwar den freien Willen nicht, da sie ja auch mittels molekularer und elektrischer, also deterministischer oder quantenmechanischer Mechanismen abläuft. Aber es vergrößert die Komplexität des Gesamtprozesses derart, dass durchaus ein subjektives Gefühl von freiem Willen entsteht.



Das Titelbild bezieht sich auf den Schwerpunkt VR (virtuelle Realität) und KI in Chemie und Uni in dieser CLB. Ich habe mir die Szene überlegt als Ausdruck dafür, dass KI-Systeme und wohl auch Roboter am ehesten dann akzeptiert werden, wenn sie menschenähnlich sind bzw. sich aus unserer Gewohnheit her ansprechen lassen. Dazu sollte ein kleiner Witz kommen, und zudem durften die VR-Brillen als Anspielung auf den ersten Artikelbeitrag nicht fehlen. Also erfand ich Roboter mit VR-Brillen – ein Witz, denn ihr Sichterlebnis wird von Kameras und der internen Informationsverarbeitung direkt erzeugt. Aber vielleicht macht es sie menschenähnlich, wenn wir alle mit VR- bzw. AR-Brillen (augmented reality, erweiterte Realität) durch Straßen laufen. Historische Stadtbilder etwa ließen sich so gleich mit Zusatzinformationen in einer AR-Brille beschreiben.

Der KI habe ich vorgegeben: *Draw 3D-Models of a male robot and female robot with VR-glasses and draw in addition a woman and a man with VR-glasses walking on the street.* Von einem Mann war zwar nichts zu sehen, dafür gab es einen dritten Roboter. Die generierte Szene weckte dann gleich eine weitere Assoziation: „Sie kommen". Das ist das hier ebenfalls abgebildete berühmte Foto von Helmut Newton, Paris 1981. Damals bescherte der Fotograf den Frauen eine neue Präsenz; sind es heute KIs und Roboter? Übrigens gibt es auch einen Film von 2002 mit dem deutschen Titel „Sie kommen" (engl.: „They"). In dem Film handelt es sich um – außer für ganz wenige Menschen – unsichtbare Außerirdische. Nun, KI ist nicht außerirdisch, aber wer weiß… Mag sich jeder sein Urteil über KI bilden, auch mit den Artikeln ab Seite 212 (Abb.: Rolf Kickuth/Helmut Newton).

Nun ist unsere Umwelt wirklich komplex. Wir werden heutzutage von Informationen überflutet, und nicht nur das: Etliche davon sind Falschinformationen, Fake News. Um sich da noch zurecht zu finden, geht man auf die Schule, die Universität – und lernt eine Menge Wissen, egal ob Gesetzte und Urteile für Juristen, Eigenschaften von Chemikalien und Reaktionswege für Biologen und Chemiker, oder auch Sprachen einschließlich Programmiersprachen. Bislang musste ein Grundstock derartigen Wissens vorhanden sein, um im Beruf schnell Entscheidungen treffen zu können, Pläne zu machen.

Schon früh habe ich jedoch gefühlt: Das braucht alles im Detail und dauerhaft nicht präsent zu sein, Latein etwas sieben Jahre in der Schule sind Zeitverschwendung, eine Gedichtinterpretation nach der anderen ist unnötig. Vielfach haben sich Interessensvertreter – und Liebhaber – bestimmter Fachrichtungen in den Lehrplänen durchgesetzt und dies und jenes als unverzichtbar markiert. In Wirklichkeit handelt es sich um Lehrinhalte aber um alte Zöpfe, die abgeschnitten gehören. Gefördert in der Schule wurden bislang vorwiegend Schüler, die sich viel für lange Zeit merken konnten. Dieses Denken ist tief verankert in der Gesellschaft, auch durch rechtliche Rahmen. Ein Arzt hätte die Differentialdiagnose wissen müssen, statt jemanden falsch zu behandeln; schließlich hat er sie mal gelernt. Gerade lese ich wieder von Bestrebungen für eine neunjährige Gymnasialzeit – eine irreführende Idee. Gelernt wird von den Schülern meist Saisonwissen: Was man gerade für die nächste Klausur benötigt. Dann kann man es vergessen…

Eine Utopie für die Ausbildung: In Zukunft gilt es, die Ausbildung von Grund auf anders und ganzheitlich auszurichten. Das beginnt mit der Frage: Wozu lernen? Allgemein sicher: Um mit anderen Menschen – auch global gesehen – gut auszukommen, und um immer wieder Nieschen zu finden, einen Verdienst zu erwirtschaften.

Die Ausbildung sollte zunächst mal eine bessere Kommunikation ermöglichen. Dazu gehört nicht nur Sprechen, sondern – wie gewohnt – auch Schreiben, Lesen und auch Rechnen, um etwa Vergleiche ziehen zu können. Das ist also sicherlich Grundwissen, was für eine erweiterte Denkfähigkeit gelernt werden muss. Dann werden sich aber schnell individuelle Unterschiede bei den Kindern zeigen – und einen neuen Ansatz fordern: Individuelle Ausbildung. Klassen mit 20, 30 Schülern müssen der Vergangenheit angehören. Optimal sähe ich ein Verhältnis von einem Lehrer auf fünf Schüler. Es gilt, sie psychologisch zu betreuen und ihre nicht Fakten-gebundenen Fähigkeiten zu erkennen und zu fördern, heutzutage Softskills genannt.

Wer ist introvertiert, wer extrovertiert? Wer kann gut kommunizieren, wer gut rechnen? Wer kann gut mit ungenauen Fakten umgehen, wer benötigt exakte Analysen? Wer kann gut überzeugen, gut führen, wer macht lieber Detailarbeit? Wer neigt zu Aggressivität, wer ist ängstlich? Wer ist körperlich aktiv, wer eher ruhig? Wer will managen, wer mag sachbezogen arbeiten? Wer ist besonders kreativ, will ausgetretene Pfade verlassen? Wer ist besonders menschenbezogen, wer technikinteressiert? Wer ist flexibel, wer hält lieber an Bekanntem fest? All das – und mehr – können gleichzeitig Schwächen und Stärken sein, und letztgenannte gilt es, herauszubilden.

Dazu müssen dann die Angebote kommen, sich über Fakten weiterzubilden, und zwar motivierend. D. h. Angebote für Projektarbeiten sind erheblich auszuweiten, zu Einzelarbeiten, aber auch zu Teamarbeiten. Dazu sollten viel mehr als bislang Möglichkeiten des Erschnupperns der Berufswelt gegeben werden, nicht nur eine BOGY-Woche (Berufsorientierung an Gymnasien) in der ganzen Schulzeit.

Ganz wichtig jedoch: Die Beurteilungsfähigkeiten, die jungen Menschen vermittelt werden, sollten auf Logik und wissenschaftliches Arbeiten beruhen. Ausbildung in Mathematik und Physik ist grundlegend, Religion muss fortfallen – Ethik-Unterricht ist für gesellschaftlichen Zusammenhalt geeigneter. Es soll nicht geglaubt werden, sondern es gilt, zu beobachten, aus den Beobachtungen Theorien zu erstellen, diese durch Experimente zu verifizieren – und ggf. abzuwandeln. Das ist die wissenschaftliche Methodik. Der Quatsch von Homöopathie muss genauso verschwinden wie der Glaube an Engel. Das soll nicht heißen, dass man – wie die Vulkanier ;-) nur auf Logik aufbauen und Emotionen unterdrücken soll – wir sind alle Menschen. Da folgt man auch mal seinem Bauchgefühl. Bei sehr ungenauen, unsicheren Ausgangslagen ist dies für manche Entscheidung unausweichlich. Aber dies sollte auch überprüft, rückreflektiert werden. Es gilt, Risiken nach realer Gefährdungslage abzuwägen.

Schließlich wird auch die künftige Ausbildung nicht auf Beurteilungen verzichten können. Das sollten jedoch keine Prüfungen sein, auf die man sich wochenlang vorbereitet, viel Wissen in sich reinhämmert, Angstzustände bekommt – um am Ende fast alles wieder zu vergessen. Vielmehr ist eine kontinuierliche Beurteilung gefragt, die nicht das Ziel eines Bestehens oder Nicht-Bestehens hat, sondern das Ziel, die Weichen für weitere Ausbildungen zu stellen, neue bzw. alternative Wege aufzuzeigen.

Ansätze für derartige Reformen sind gegenüber meiner Schulzeit früher durchaus zu finden, aber zögerlich, mit starken hemmenden Mechanismen, die sich oftmals unter dem Deckmantel von Tradition verstecken; Kinder versohlen war auch mal eine Tradition...

Der Weckruf der KI-Sprachmodelle: ChatGPT ist ein Weckruf, wie es zuvor keinen gegeben hat, um Ausbildungen zu reformieren. Deshalb habe ich die interessanten Beiträge von Volker Wiskamp zusammen in diese Ausgabe gesetzt. Sie sind ein Zeugnis dafür, wie rasend schnell die Informationstechnik sich gerade entwickelt. Auch diese CLB hat dafür Beispiele. So sind sowohl das Titelbild wie auch der Cartoon auf Seite 225 mit Hilfe von Systemen mit künstlicher Intelligenz generiert worden. Trotzdem steht da mein Name drunter: Die Idee und das Konzept stammen von mir – die „Prompts“, Anweisungen, die ich der KI gab, das Aussuchen ihrer Vorschläge, das Zusammenstellen von Einzelbildern – der Cartoon besteht aus fünf Teilen, die auch zum Teil verbessert wurden, die Sprechblasen… Ohne diesen kreativen Input wäre die KI aufgeschmissen gewesen. Von daher sollten die neuen Möglichkeiten von ChatGPT, Dall-e und Co. als Werkzeuge genutzt werden, um Ausbildungen zu verbessern, kreativer zu machen, Arbeitsvorgänge von Routine zu entlasten.

Es gilt, die Grenzen der KI zu lernen, ihren Unsinn zu entlarven – genauso wie den von Menschen verbreiteten. Dazu benötigt es eben Logik und wissenschaftliches Denken. Verbote von KI-Nutzung sind so kontraproduktiv wie das Verbot von Taschenrechnern in Schulen.

KI ohne Grenzen: Zudem geht die Entwicklung der KI-Systeme mit Riesenschritten voran. Zu GPT-4, das neueste Sprachmodell, das Volker Wiskamp in seinem zweiten Artikel untersucht, meint der deutsche IT-Spezialist Helmut Linde „Nach der Lektüre der zahlreichen Fallbeispiele ((aus einem Fachartikel vom 22.3.2023)) fällt es schwer, die Position aufrechtzuerhalten, dass die Entwicklung einer allgemeinen künstlichen Intelligenz noch viele Jahrzehnte entfernt liegen oder sogar technisch grundsätzlich unmöglich sein könne.“ (s. S. 246) Sehr bemerkenswert: Die Autoren des Fachartikels räumen ein, dass niemand wirklich versteht, warum diese Modelle so gut funktionieren. Viele Fähigkeiten seien emergente Eigenschaften. Was große KI-Modelle unerwartet können ist etwa das Lernen mit nur wenigen Beispielen (Few-Shot-Learning). Kommt einem das nicht bekannt vor? Noch mehr Rückkopplungsschleifen, wichtige Elemente komplexer neuronaler Systeme – wie jetzt schon als Verbesserung von GPT-4 angedacht: Auf welcher Stufe des (Un-)Bewussten steht man dann?

Der Stifterverband stellte gerade eine „Beschleunigungsformel für Deutschland“ vor, die gekürzt in der Rubrik „Erreichen und Erhalten“ zu finden ist. Roadmaps, Katalysatoren, Orchestrierung der systemischen Zusammenarbeit seien die drei Erfolgsbausteine für einen neuen Beschleunigungsansatz. Katalysatoren sind dabei Initiatoren und Beschleuniger von Veränderungen, zum Beispiel als Schlüsseltechnologien oder administrative Rahmenbedingungen. Orchestrierung meint: Eine Balance finden aus Koordination und Wirkungsmessung sowie Freiräumen für Kreativität schaffen. All dies gilt genauso und zuallererst für Ausbildungen. Dann können die KI-Sprachmodelle und -Geschöpfe kommen – die hoffentlich mit dem Gender-Hype sachlich umgehen (siehe Artikel auf S. 291),

Ihr

Schwerpunkt Ausbildung und Computer

Verständlich ohne Verständnis

Komplexe neuronale Netze: Die Technik hinter ChatGPT et al.

Rolf Kickuth

ChatGPT ist wohl das „iPhone-Ereignis" für künstliche Intelligenz (KI). Sie ist schon lange in unseren Alltag eingedrungen, aber nahezu unmerklich, beispielsweise bei der Optimierung von Fotos in den Mobiltelefonen oder auch als Sprachassistent – mit allerdings noch begrenzter Funktionalität. Durch die unmittelbare und gleichzeitig total einfache – weil wie in der zwischenmenschlichen Kommunikation gewohnte – Ansprache an ChatGPT und die schlau klingenden Antworten ist jedoch praktisch jedem klar geworden, welches Potenzial in KI steckt. Im vorangegangenen Artikel zeigt CLB-Autor Volker Wiskamp, dass es für KIs mittlerweile kein Problem mehr ist, auch in Deutschland universitäre Prüfungen zu bestehen. Dennoch ist das keine Zauberei, sondern kaum vorstellbare Rechenarbeit.

Maßgeblich für die Funktion von ChatGPT sind künstliche neuronale Netze, über die in der CLB wiederholt berichtet wurde, so in [1, 2, 3]. Spekulationen über die Leistungsfähigkeit von künstlichen neuronalen Netzen (KNN; auch ANN für Artificial Neural Nets) nährten eine Zeit lang ihre Mystifizierung.

1. Neuronetze: Eine alte Idee

Tatsächlich liegen die Anfänge ihrer Erforschung in einer Zeit, die durch Technik-Euphorie gekennzeichnet war. Einen ersten Schritt dazu machten 1943 die US-Mathematiker Warren McCulloch (1898-1969) und Walter Pitts (1923-1969): Sie schufen ein „Neuron" als logisches Schwellwert-Element mit mehreren Eingängen und einem einzigen Ausgang (Abbildung 1). Ihr einfaches Neuronenmodell konnte sowohl Und-, Oder- als auch Nicht-Gatter simulieren. Es gab die logischen Werte true oder false aus, 1 oder 0, schaltete Strom oder nicht, wenn die Summe der Eingangssignale einen Schwellenwert überschritt [4]. Dies entsprach der neurobiologischen Analogie eines Aktionspotentials, das eine Nervenzelle bei einer kritischen Änderung ihres Membranpotentials aussendet.

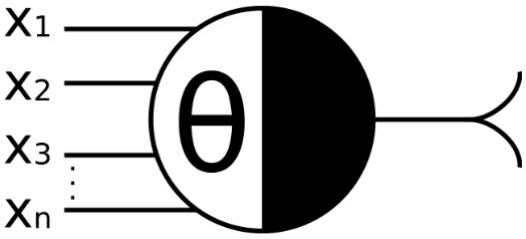


Abbildung 1: Diagramm eines McCulloch-Pitts-Neurons (Abb.: Wikipedia / Adrian Lange CC BY-SA 3.0).

1949 formulierte dann Donald Hebb (1904-1985), ein kanadischer Physiologe, die 1. Lernregel: Gleichzeitig aktive Neuronen verstärken Verbindungen; im Englischen gibt es dafür den eingängigen Merksatz: „What fires together wires together" [5]. Detaillierter ausgedrückt: Wenn eine Nervenzelle A eine Nervenzelle B dauerhaft und wiederholt erregt, wird die Synapse dadurch so verändert, dass die Signalübertragung effizienter wird. Dadurch erhöht sich das Membranpotential im Empfänger-Neuron.

Ganz im Sinne der intelligenten Leistung, Muster zu erkennen und diese weiterzuentwickeln, liegt es natürlich nahe, sich das anzuschauen, was durch die Natur bereits erfunden wurde, will man intelligente Systeme entwickeln.

2. Lernen und Intelligenz

Nach und nach entwickelte man Vorstellungen, wie das große neuronale Netz, was wir als Gehirn kennen, lernen kann. Lernen ist ein Grundpfeiler der Intelligenz. Es ermöglicht, sich den Gegebenheiten des Lebens und der Umwelt anzupassen, darin im Sinne einer Lebenserhaltung oder -verbesserung zu handeln und die Umwelt dementsprechend gegebenenfalls im eigenen Interesse zu verändern. Bei dem Nachempfinden intelligenter

Der Autor:
Rolf Kickuth (70) ist Chemiker und Verleger der CLB. 1990 entwickelte der Autor mit der „AXON" eine Zeitung für künstliche Intelligenz, insbesondere künstliche neuronale Netze. In den späteren 1990er Jahren war er auch Chefredakteur des „Informatik-Spektrum", der Hauszeitschrift der Gesellschaft für Informatik. Seit 1985 war er zunächst beim Verlag Chemie (jetzt Wiley-VCH), später in seinem Verlag „Rubikon" für die Schweizer Wochenzeitung „Chemische Rundschau" redaktionell tätig, von 1994-1999 als Leiter der Redaktion Deutschland. Er engagiert sich für Humanismus und Atheismus.

Systeme steht daher die Entwicklung eines effizienten maschinellen Lernens an erster Stelle. Wichtige Mechanismen des Lernens sind

- das Entwickeln neuer bzw. das Löschen bestehender Verbindungen;
- das Ändern der Stärke der Verbindungen zwischen den Neuronen, der

synaptischen Gewichte

- das Anpassen der Schwellenwerte der Neuronen (wann leitet es Impul-

se weiter);

- grundlegend natürlich das Hinzufügen oder Löschen von Neuronen.

Was soll damit erreicht werden? Etymologisch ist das Wort „lernen“ u. a. mit den Wörtern „lehren“ und „List“ verwandt. Es gehört zur Wortgruppe von „leisten“, das ursprünglich „einer Spur nachgehen, nachspüren, schnüffeln“ bedeutet. Im Gotischen heißt lais „ich weiß“, bzw. genauer „ich habe nachgespürt“ und laists für „Spur“. Die indogermanische Wurzel „lais-“ bedeutet „Spur, Bahn, Furche“. Schon von der Herkunft her hat Lernen also etwas mit Spuren hinterlassen, aber auch mit nachspüren zu tun.

Das zeigt sich dann auch strukturell im Gehirn. Lernen entsteht in Neuronetzen durch die Bildung oder Verstärkung von Verbindungen bei gleichzeitiger Aktivität zweier Neuronen (oder Gruppen von Neuronen). Werden diese Verbindungen nicht genutzt, bilden sie sich auch wieder zurück. Beides spiegelt die neuronale Plastizität wieder. Das ist ein Mechanismus, der eine Lösung des Konflikts zwischen Stabilität und Plastizität von solchen Netzwerken ermöglichen soll: Sie sollen gelernte Muster behalten, trotzdem Neues leicht lernen und damit die alte Information nicht überschreiben. In den Details ist biologisches Lernen allerdings noch nicht vollständig aufgeklärt. Insbesondere geht es darum, die vielfältigen Rückkopplungsschleifen im Gehirn zu untersuchen, die die selbstständigen Lernmechanismen ermöglichen.

Unmittelbar mit Lernen verbunden ist der Begriff der Intelligenz. Als Intelligenz (die direkte Übersetzung des lateinischen Begriffs bedeutet ungefähr „dazwischen lesen“ oder auch „wählen zwischen“) bezeichnet man allgemein die kognitive Leistungsfähigkeit des Menschen. Da einzelne kognitive Fähigkeiten unterschiedlich stark ausgeprägt sein können und keine Einigkeit besteht, wie diese zu bestimmen und zu unterscheiden sind, gibt es keine allgemeingültige Definition der Intelligenz. Beispiele für unterschiedliche Ausprägungen sind analytische, kreative, emotionale oder soziale Intelligenz. Es gibt verschiedene Intelligenztheorien oder Intelligenzmodelle, die versuchen, die Ursachen und Auswirkungen von Intelligenz zu beschreiben.

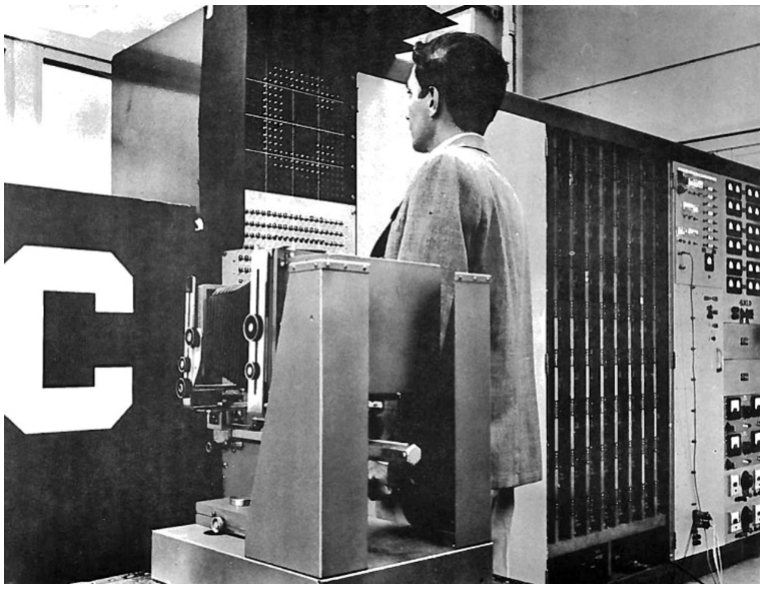
Für das Nachempfinden von Intelligenz in künstlichen Systemen wichtig: Es gibt intelligente Leistungsfähigkeiten, die auf Spezialgebiete bezogen sind. Bei künstlicher Intelligenz sind dies beispielsweise Computer, die auf Schach spezialisiert sind. Dann spricht man von schwacher künstlicher Intelligenz. Der Mensch hingegen zeichnet sich durch universelle Intelligenz aus, die bei menschengemachten Systemen als starke künstliche Intelligenz gekennzeichnet wird. Bei allen Ausprägungen gibt es ein jedoch Gefühl dafür, was als intelligent zu bezeichnen ist, und Forscher wollten dies gerne nachempfinden.

3. Erste Entwicklungen zur KI

Dies ist die Triebfeder zur Entwicklung künstlicher Intelligenz. 1958 – neun Jahre nach Hebb’s Postulat – baute der amerikanische Psychologe Frank Rosenblatt (1928-1971) das Perceptron, ein einfaches künstliches neuronales Netzwerk [rose1958]. Es war lernfähig – der entscheidende Unterschied zu dem McCuloch-Pitts-Modell – und fehlertolerant, konnte Muster klassifizieren und „Erfahrungen“ verallgemeinern, also generalisieren. All das waren Eigenschaften, die man zuvor nur dem Gehirn zuschrieb.

Das Mark 1-Perceptron bestand aus einem Raster mit 400 Photozellen, die mit 512 neuronensimulierenden Schaltkreisen verbunden waren (Abbildung 2). Eine Hebb-ähnliche Lernregel veränderte die Verbindungen (synaptischen Gewichte) zwischen ihnen, und die „Netzhaut“ lernte, Buchstaben zu

Abbildung 2: Das Perceptron sollte eher eine Maschine als ein Programm sein. Die erste Implementierung erfolgte allerdings in Software für IBM 704. Dann wurde es in speziell angefertigter Hardware als „Mark 1-Perceptron“ gebaut. Es konnte Buchstaben erkennen, die eine Anordnung von 400 Fotozellen aufnahm. Sie waren zufällig mit den „Neuronen“ verbunden. Synaptische Gewichte ließen sich durch Potentiometer einstellen, und Gewichtsaktualisierungen während des Lernens erfolgten durch Elektromotoren (Foto: Cornell Aeronautical Laboratory).



erkennen. Es war noch Mechanik angesagt bei dem Mark 1-Perceptron: Per Hand wurden die Voreinstellungen vorgenommen, und Elektromotoren verstellten die Potentiometer, die als synaptische Gewichte den simulierten Neuronen vorgeschaltet waren. Diese Neuronen selbst waren eine Kombination aus elektronischem Verstärker, der die Eingabespannungen aufsummierte, und einer Logikschaltung, die nach einer Schwellwertfunktion ein Ausgangssignal erzeugte.

Das Perceptron führte zu einem frühen KI-Hype: Auf einer Pressekonferenz der US-Marine von 1958 gab Rosenblatt Erklärungen über das Perceptron ab, das eine hitzige Kontroverse unter der jungen KI-Gemeinschaft hervorrief. Die „New York Times“ griff die Erläuterungen mit den Worten auf, dass das Perceptron „der Embryo eines elektronischen Computers ist, von dem man [die Marine] erwartet, dass er in der Lage ist, zu gehen, zu sprechen, zu sehen, zu schreiben, sich selbst zu reproduzieren und sich seiner Existenz bewusst zu sein“.

Die aufkommende Leistungsfähigkeit digitaler Computer führte dann auch zu einer KI-Entwicklung, die man als „symbolische KI“ beschreibt. Symbolische KI einerseits und neuronale oder auch konnenktionistisch genannte KI andererseits versuchen, sich dem Phänomen Intelligenz von zwei verschiedenen Sichtweisen aus anzunähern. Während die neuronale KI Basisstrukturen für Intelligenz-Erscheinungen erschafft – kurz gesagt das Gehirn nachbildet, also von den kleinsten Einheiten her höhere Leistungen erbringt (bottom up), verfolgt symbolische KI einen top down-Ansatz. Sie nähert sich der Intelligenz von begrifflicher Seite. Symbole sind dabei Objekte, die den Zugriff auf Bedeutungen ermöglichen. Dadurch können dann Schlüsse gezogen und Probleme gelöst werden. Hirnstrukturen spielen also bei der symbolischen KI keine Rolle; es kommt allein auf das Ergebnis an.

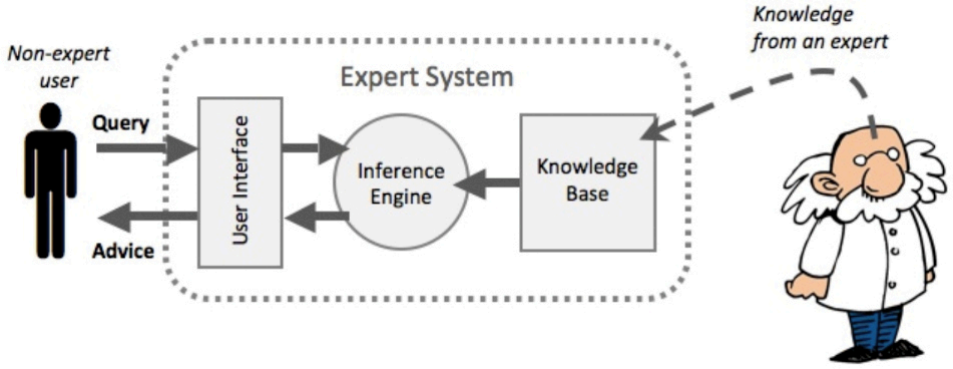
Typische Verfahren symbolischer KI sind repräsentiert in Expertensystemen (Abbildung 3). Sie nutzen beispielsweise fallbasiertes Schließen (CBR für case based reasoning). Dabei werden Objekte in Datenbanken abgelegt, mit Attributen gekennzeichnet, miteinander in Beziehung gesetzt. Regelwerke, die auch Ähnlichkeiten berücksichtigen können, erlauben, von bekannten Situationen (case based) auf neue zu schließen. Case-Based Reasoning hat sich besonders in Anwendungssystemen für den Kundendienst, so genannten Help-Desk-Systemen, bewährt.

4. Hoch-Zeiten und Eiszeiten der KI

All diese Wege, eine künstliche Intelligenz aufzubauen, hatten ihre Hoch-Zeiten, aber ebenso Eiszeiten – oder liefen ins Leere. Der entscheidende Grund dafür: Die Hardware war noch bei Weitem nicht so leistungsfähig, dass sich damit eine Komplexität nachbilden ließ, wie sie hinter dem steckt, das wir als Intelligenz bezeichnen. Dazu hatte bereits 1950 der geniale Mathematiker und Informatiker Alan Turing (1912-1954) eine Idee entwickelt. Grundlegend war sein Werk „Computing Machinery and Intelligence“; darin schlug er einen Test vor, anhand dessen man die Gleichwertigkeit des Denkvermögens einer Maschine verglichen mit dem eines Menschen feststellen könnte [turi1950]. Er nannte den Test, der heute nach ihm benannt ist, „The Imitation Game“ – 2014 der Titel eines Kinofilms über Turing.

Der Turing-Test ist auch heute noch oft im Gespräch. Sein Prinzip ist einfach: Ein Mensch unterhält sich ohne Sichtkontakt mit zwei unbekannten Gesprächspartnern. Einer davon ist eine Maschine. Kann der Fragen stellende Mensch nicht herausfinden, welcher Gesprächspartner die Maschine ist, gilt diese als intelligent, also als mit einem dem Menschen ebenbürtigen Denkvermögen ausgestattet.

Abbildung 3: Prinzip eines Expertensystems: Der Laie stellt eine Anfrage. Beantwortet wird sie von einem System, das bereits über eine Datenbank verfügt, die mit Wissen von Experten aufgebaut wurde (Knowledge Base). Die Inferenzmaschine (Inference Engine) leitet durch Schlussfolgerung neue Aussagen aus der bestehenden Wissensbasis ab. Die Mensch-Maschine-Schnittstelle (User Interface) übersetzt die Anfrage des Nutzers so, dass die Maschine es versteht, bzw. gibt menschenverständliche Antworten (Abb.: CLB-Archiv).



5. Neuronetze setzen sich durch

Bei den Bemühungen, eine KI zu schaffen, hat sich letztlich der Ansatz über künstliche neuronale Netze durchgesetzt. Symbolische KI mit Datenbanken und Expertensystemen blieb für Nischen vorbehalten. KNN erlebten ihre Wiedergeburt Ende der 1970er, Anfang der 1980er Jahre. Zum einen hatte sich herausgestellt, dass die von einigen Wissenschaftlern beschriebenen Einschränkungen nur für spezielle, sehr einfache Netzwerke

ohne „verdeckte Schicht“ gelten; zum anderen ermöglichte die Geschwindigkeitssteigerung konventioneller Computer, immer bessere Simulationen von künstlichen neuronalen Netzen zu konstruieren und zu prüfen. Besonders erfolgreich waren mehrschichtige Netze mit einer oder mehreren verdeckte Schichten von künstlichen Neuronen. Der Informationsfluss darin ist gerichtet, fließt von einer Seite zur anderen. Daher nennt man diese Netze auch Feedforward-Netze. Dabei sind alle Neuronen einer Schicht mit allen Neuronen der nächsten Schicht verbunden.

5.1 Backpropagation und Deep Neural Nets

Dann entwickelte man noch die Backpropagation-Methodik: Den KNN wurde beibracht, Fehler beim Training durch alle Schichten des Netzwerkes rückzurechnen und zu verringern [6]. Dazu muss ein externer Lehrer existieren, der zu jedem Zeitpunkt der Eingabe die gewünschte Ausgabe, den Zielwert, kennt. Damit stand ersten Anwendungen der künstlichen neuronalen Netze nichts mehr im Wege.

Diese ersten Anwendungen änderten jedoch nichts daran, dass die Entwicklung von KNN wieder einmal ins Stocken geriet. Schuld daran war nicht die Theorie, sondern die mangelnde Leistungsfähigkeit der Computer. In der zweiten Hälfte der 1990er Jahre hörte man praktisch nichts mehr von künstlichen neuronalen Netzen.

Heute haben die wirtschaftlich wohl größte Bedeutung nach wie vor strukturell relativ einfache künstliche neuronale Netze, Weiterentwicklungen der Backpropagation-Netze. Sie heißen Deep Neural Nets und geben dem mit ihnen möglichen Lernverfahren den Namen „deep learning“. Das „tief“ bezieht sich dabei einfach auf die Anzahl von Zwischenschichten von Neuronen, die zwischen Eingangsund Ausgangsschicht eines Backpropagation-Netzes liegen. Früher war eine Simulation solcher Netze kaum vorstellbar.

5.2 Long short-term memory

Auch mit leistungsfähiger Hardware galt und gilt es jedoch, für KI zahlreiche Probleme zu lösen. Wie kommt es beispielsweise, dass der Mensch so effizient lernen kann? Die long short-term memory (LSTM)-Neuronetze zeigen einen Lösungsweg auf. Sie erweitern die Lernleistung der künstlichen Konstrukte auf Anwendungen, die mit dem Verlauf der Zeit zu tun haben. LSTM-Netze gehören zu rekurrenten neuronalen Netzen (RNN). Sie zeichnen sich durch Verbindungen von Neuronen einer Schicht zu Neuronen derselben oder einer vorangegangenen Schicht aus.

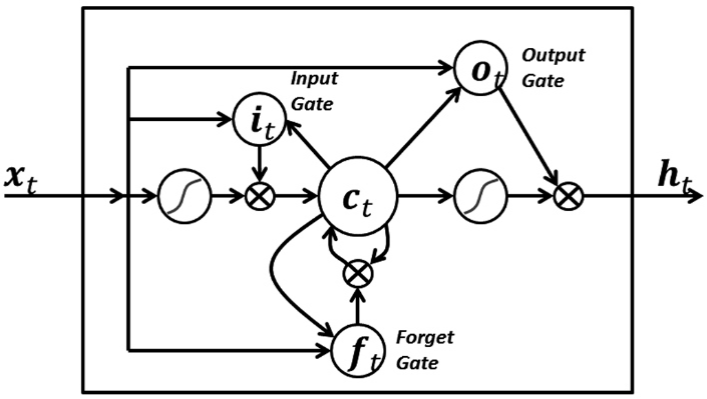
Der Aufbau und die Funktion eines LSTM – von dem es etliche Architekturvarianten gibt – ist bei weitem nicht mehr so übersichtlich wie ein künstliches Neuron, dass im Wesentlichen über eine Aufsummierungs- und eine Schwellwertfunktion verfügt. Daher sagt man, LSTM sind aus Modulen aufgebaut (Abbildung 4). Ein LSTM zeichnet sich insbesondere dadurch aus, dass auch über viele Schichten hinweg ein Fehler zurückgekoppelt werden kann.

Ein LSTM-Modul beinhaltet als zentralen Kern, etwas, das man „innere Zelle“ bezeichnet (s. Abb). Sie wird gezielt mit Informationen gefüttert. Dafür sorgen drei Gates („Tore“): Das *Input Gate* steuert das Ausmaß, mit dem ein neuer Wert in die Zelle fließt; das *Forget Gate* steuert das Ausmaß, mit dem ein Wert in der Zelle verbleibt bzw. vergessen wird, und schließlich steuert das *Output Gate* das Ausmaß, mit dem der Wert in der Zelle zur Berechnung für das nächste Modul der Kette verwendet wird. Diese Netzelemente werden mit sigmoiden neuronalen Funktionen und diversen Vektor- und Matrixoperationen verbunden und ineinander überführt.

Bei LSTM-Einheiten verbleibt ein Fehler jedoch im Speicher der Einheit, wenn Fehlerwerte vom Ausgang zurück übertragen werden. Dieses „Fehlerkarussell“ führt kontinuierlich einen Fehler zu jedem der Gatter zurück, bis sie lernen, diesen Wert loszuwerden – durch richtige Ergebnisse. Daher ist diese Backpropagation wirkungsvoll beim Trainieren einer LSTM-Einheit, um sich Werte für lange Zeiträume zu merken.

Die „Tore“ können über eine „Peephole“-Funktion in die innere Zelle blicken und somit auch Informationen aus dieser Zelle verarbeiten, ein Rückkopplungsmechanismus. Er erlaubt das Lernen

Abbildung 4: Prinzipieller Aufbau eines LSTM-Moduls. Was die „innere Zelle“ c erfährt, steuern drei Gates. Jedes dieser Gatter kann als ein Standard-Neuron in einem vorwärtsgekoppelten neuronalen Netzwerk betrachtet werden: das heißt, sie berechnen eine Aktivierung (unter Verwendung einer Aktivierungsfunktion) einer gewichteten Summe. Die Kreis-Symbole mit den gekreuzten Linien repräsentieren einen Faltungsoperator, eine spezielle Architektur von LSTM-Netzen (vergleichbar mit der Funktion von CNN). Die großen Kreise mit S-artiger Kurve sind die Sigmoidfunktionen. Die Pfeile, die von der Zelle jeweils zu den Gates zeigen, sind die „Gucklochinformationen“ vom letzten Durchlauf, ein Rückkopplungsmechanismus. Guckloch (Peephole) heißt diese Anordnung, weil die Gates den Status der inneren Zelle sehen können – und somit auch die Informationen aus der Zelle verarbeiten. Solche Rückkopplungen sind entscheidend für das Lernen von Informationen, die über einen bestimmten Zeitraum eintreffen. (Abb.: Wikipedia / BiObserver CC BY-SA 4.0).



von Informationen, die über einen bestimmten Zeitraum eintreffen.

Das LSTM vereint also sowohl Eigenschaften von Backpropagation-Netzen wie von rekurrenten neuronalen Netzen (RNN (s.o.). LSTM-Netze werden insbesondere in der Spracherkennung für die Klassifikation von Phonemen (Laute gleicher Bedeutung, z. B. gerolltes und nicht gerolltes „R“) eingesetzt.

Ein LSTM lässt sich auch mit reinforcement learning kombinieren und trainieren; man nimmt dazu beispielsweise Evolutionsstrategien oder genetische Algorithmen.

6. Nach LSTM kommen die Transformer

Die LSTM-Neuronetze greifen den Faktor „Zeit“ auf. Sie sind heutzutage unverzichtbar; in der Spracherkennung, bei den intelligenten Assistenten kommen sie zum Einsatz.

Hier kommt jetzt ChatGPT ins Spiel: Diese KI ist so leistungsfähig, dass sie uns glauben macht, wir sprächen mit einem anderen Menschen. Die davon ausgehende Faszination sowie einfache Nutzbarkeit ist offenbar so groß, dass die Verbreitung von ChatGPT alles geschlagen hat, was sich bislang schnell verbreitet hat. ChatGPT hat gerade mal zwei Monate benötigt, um 100 Millionen Nutzer zu haben; verglichen damit haben TikTok neun, Instagram 30 und Spotify 41 Monate gebraucht. Was steckt nun hinter ChatGPT?

Ein neues, „trendiges“ Neuronetz, das dies leistet, kommt von Google: der Transformer. Die von Google-Forschern im Jahr 2017 entwickelte Transformer-Netzwerkarchitektur war ursprünglich als Mittel zur Verbesserung der maschinellen Übersetzung gedacht, hat sich jedoch zu einem Eckpfeiler des maschinellen Lernens entwickelt, um einige der beliebtesten vorgefertigten Sprachmodelle auf dem neuesten Stand der Technik zu erstellen.

Im Gespräch mit einer Fachzeitschrift erklärte Google-KI-Chef Jeff Dean und im Januar 2020 Transformer-basierte Sprachmodelle zu einem wichtigen Trend, der sich fortsetzen werde [7]. Google selbst berichtetete schon 2017, dass der Transformer bis dahin als leistungsfähig geltende Neuronetze in Bezug auf die Benchmarks für die Übersetzung von akademischem Englisch nach Deutsch und Englisch nach Französisch übertrifft. Neben einer höheren Übersetzungsqualität erfordert der Transformer weniger Rechenaufwand für das Training und eignet sich viel besser für moderne Hardware für maschinelles Lernen, wodurch das Training um eine Größenordnung beschleunigt wird. [8]

Ein Grund dafür: Die sequentielle Natur von RNNs macht es schwieriger, moderne, schnelle, spezialisierte Computerchips wie Grafikprozesso-

KNN: Verschiedene Lernmechanismen

Lernen mit Anleitung: supervised learning

Man kann künstliche Neuronetze in Kategorien einteilen, die sich auf ihr Lernverhalten beziehen. Die bislang am häufigsten in Anwendungen eingesetzten KNN, die Backpropagation-Netze, arbeiten nach der Methode des überwachten Lernens (supervised learning). Dazu gehören auch die in jüngerer Zeit so erfolgreichen Deep Learning-Netze mit ihren Faltungsmechanismen (CNN). Dabei gibt ein externer Lehrer dem Netz zu jeder Eingabe die korrekte Ausgabe oder die Differenz der tatsächlichen zur korrekten Ausgabe an. Anhand dieser Differenz wird dann das Netz über die Lernregel modifiziert. Diese Technik setzt allerdings voraus, das Trainingsdaten existieren, die aus Paaren von Ein- und Ausgabedaten bestehen.

Reinforcement Learning: „Zuckerbrot und Peitsche“

Als weitere Lernvariante gibt es das „bestärkende Lernen“ (RL; reinforcement learning). Diese Variante des Lernens ist ein Zwischending: Es weist Elemente der Überwachung wie auch Elemente der Selbstständigkeit auf. Letztere liegen darin, dass das Netz selbstständig eine Strategie sucht, um Lösungen für eine Aufgabe zu finden. Es gibt bei RL keinen Fehleralgorithmus, der zu einem bestimmten Moment einen bestimmten Weg für optimal angibt. Aber es gibt ab und an „Belohnungen“, die auch negativ sein können. Im Gegensatz zum überwachten Lernen wird dem Netz hier lediglich mitgeteilt, ob seine Ausgabe korrekt oder inkorrekt war. Das Netz erfährt nicht den exakten Wert des Unterschiedes. Das Netz verändert daraufhin mit Zufalls- und Wahrscheinlichkeitsmechanismen selbst seine Eigenschaften. Erfolgreiche Algorithmen sind für das Auslösen dieser Veränderungen Monte-Carlo-Methoden und Temporal Difference Learning. Bei diesen handelt es sich um eine Reihe von Algorithmen, bei denen der Agent eine Nutzenfunktion besitzt, welche einen bestimmten Zustand oder eine bestimmte Aktion in einem Zustand bewertet. Bei kleinen Zustands- oder Aktionsräumen kann dies eine Tabelle sein, deren Felder anhand der erhaltenen Belohnung aktualisiert wird. Bei großen Zustandsräumen muss die Funktion jedoch approximiert werden. Dazu eignet sich beispielsweise die Fourierreihe oder auch ein neuronales Netz.

RL ist damit dem natürlichen Lernen nahe. Auch wir erkunden unsere Umwelt in frühen Lernphasen nach dem Versuch und Irrtum-Verfahren, erhalten Belohnungen und Bestrafungen (Lob, Akzeptanz, materielle Einbußen...), „Zuckerbrot und Peitsche“.

Man setzt RL vorteilhaft in folgenden Fällen ein: Ein Modell der Umgebung ist bekannt, aber eine analytische Lösung ist nicht verfügbar; zweitens: es ist nur ein Simulationsmodell der Umgebung vorhanden; drittens: um Informationen über eine Umgebung zu erhalten, muss man mit ihr interagieren.

Selbstständiges Lernen durch Rückkopplung

Es gibt auch künstliche Neuronetze, die selbstständig lernen (unüberwachtes Lernen, unsupervised learning). Hierbei gibt es überhaupt keinen externen Lehrer, daher heißt dieses Lernparadigma auch self-organized learning. Das Netz versucht ohne Beeinflussung von außen die präsentierten Daten in Ähnlichkeitsklassen aufzuteilen. Herausragendes Merkmal selbstlernender Neuronetze sind vielfältige Rückkopplungsmechanismen. Biologische Gehirne sind durch eine Vielzahl von Rückkopplungen zwischen den Neuronen geprägt, insbesondere im Neocortex des Menschen.

ren oder ähnliche, speziell für künstliche Neuronetze entworfene TPUs (Tensor Processing Units) vollständig zu nutzen. Diese zeichnen sich nämlich durch parallele und nicht sequentielle Verarbeitung aus.

6.1 Aufmerksamkeit erzeugen

Im Gegensatz dazu führt ein Transformer-Neuronetz nur eine kleine, konstante Anzahl von Schritten aus, die empirisch ausgewählt werden.

Die Grundidee hinter den Transformer-Modellen ist einfach: Es gilt, zu einem gegebenen Text immer das nächste passende Wort zu finden. Dieser Prozess wird mehrfach wiederholt, bis genug Text erzeugt wurde. Der Name „Transformer“ stammt von dieser Methode, mit der ein Computer eine Folge von Zeichen in eine andere Folge von Zeichen übersetzen kann.

Im Kern beruht ein Transformer auf einem Verarbeitungsschritt, der als Aufmerksamkeitsmechanismus (engl.: attention) bezeichnet wird und der vom Algorithmus wie gesagt mehrfach wiederholt wird. Der Name des Verfahrens leitet sich davon ab, dass zu jedem Wort die Aufmerksamkeit des Sprachmodells auf bestimmte andere Ausdrücke im Text gelenkt wird, die für die Interpretation jenes Wortes besonders wichtig sind. Bei dem Wort „König“ etwa käme viel Aufmerksamkeit auf im Text benachbarte Ausdrücke wie „Springer“, „Läufer“, „Thron“ oder „Schloss“ , da sie Hinweise auf die korrekte Deutung des Wortes „König“ geben können. [9]

In jedem Schritt wird ein Aufmerksamkeitsmechanismus angewendet, der die Beziehungen zwischen allen Wörtern in einem Satz unabhängig von ihrer jeweiligen Position direkt analysiert. Das machen im Grunde auch RNNs, aber Transformer-Netze können dies in einem Schritt, sehr effizient, ohne sequenzielle Verarbeitung. In dem englischsprachigen Beispiel “I arrived at the bank after crossing the river” kann der Transformer lernen, sich sofort um das Wort „bank“ zu kümmern. Er findet heraus, dass sich das Wort „bank“ auf das Ufer eines Flusses und nicht auf ein Finanzinstitut bezieht; dabei trifft er diese Entscheidung in einem einzigen Schritt.

Das Modell arbeitet nach dem Prinzip des bestärkenden Lernens. Dabei entwickelt das Programm in mehreren Stufen selbstständig eine Strategie, die mit „Belohnungen“ gesteuert wird. Das kann positives oder negatives Feedback sein, das in den ersten Stufen von Menschen überwacht und gegeben wurde. Die Struktur von Transformern ist ziemlich komplex (Abbildung 5). Er hat je einen Satz miteinander verketteter Encoder Decoder. Jeder Codierer verarbeitet seine Eingabevektoren, um Codierungen zu erzeugen, die Informationen über die Teile der Eingaben enthalten, die füreinander relevant sind. Es übergibt die so gewonnenen Codierungen als Eingänge an den nächsten Codierer. Jeder Decodierer macht das Gegenteil: Er verwendet die in den Codierungen eingebauten Kontextinformationen, um eine Ausgabesequenz zu erzeugen. Dafür verfügt jeder Codierer und Decodierer über einen Aufmerksamkeitsmechanismus, der für jeden Eingang die Relevanz jedes Eingangs abwägt. Wenn ein Satz an ein Transformer-Modell übergeben wird, werden die Aufmerksamkeitsgewichte zwischen jedem Satzmerkmal gleichzeitig berechnet. Das Transformer-Modell berechnet im Detail für jede Aufmerksamkeitseinheit drei Gewichtsmatrizen (Abfragegewichte, Schlüsselgewichte und Wertgewichte). Sowohl die Codierer als auch die Decodierer verfügen über ein finales neuronales Feed-Forward-Netz für die zusätzliche Verarbeitung der Ausgangssignale und enthalten Bestandteile weiterer Neuronetz-Modelle.

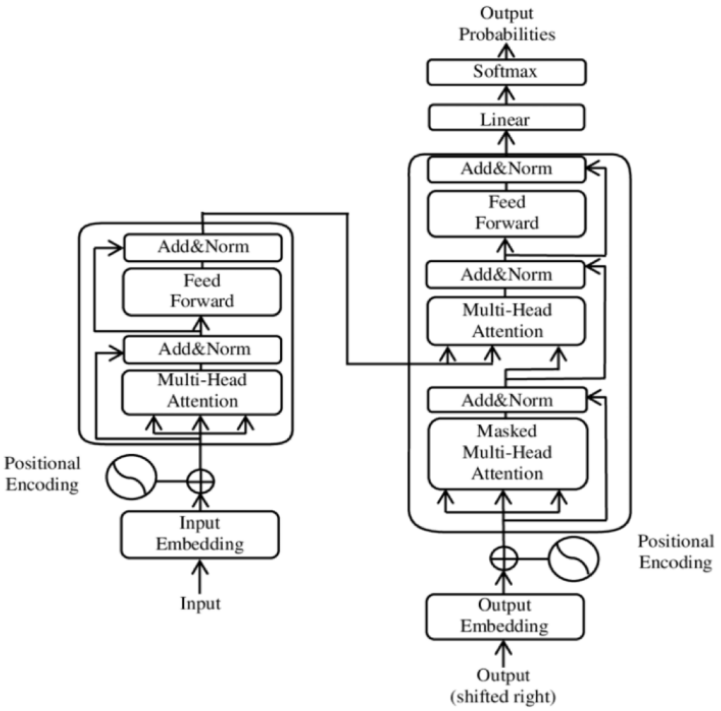


Abbildung 5: Prinzipieller Aufbau des Transformer-Netzwerkmodells. Der Eingabetext wird in Token – einzelne Wörter oder Wortteile, zerlegt , und jeder Token wird über eine Worteinbettung in einen Vektor umgewandelt. Dann werden der Worteinbettung Positionsinformationen des Tokens hinzugefügt. Tokenisierung ist der Prozess der Umwandlung einer Folge von Zeichen (z. B. in einem Computerprogramm oder einer Webseite) in eine Folge von lexikalischen Token mit einer zugeordneten und damit identifizierten Bedeutung. Der Encoder besteht aus Codierungsschichten, die die Eingabe iterativ Schicht für Schicht verarbeiten, während der Decodierer aus Decodierungsschichten besteht, die dasselbe mit der Ausgabe des Codierers tun. Die Funktion jeder Codierschicht besteht darin, Codierungen zu erzeugen, die Informationen darüber enthalten, welche Teile der Eingaben füreinander relevant sind. Um dies zu erreichen, nutzt jede Encoder- und Decoderschicht einen Aufmerksamkeitsmechanismus. Diese rudimentäre Erklärung kann mit der Abbildung nur eine Ahnung davon vermitteln, dass ein Transformer-Netzwerkmodell eine hohe Komplexität besitzt. Sie mit Daten zu füttern und diese Daten durch das Netzwerk durchzujagen bedarf einer riesigen Rechenleistung, die erst seit Kurzem zur Verfügung steht (Abb.: Wikipedia / Yuening Jia CC BY-SA 3.0).

6.2 OpenAI und GPT

Im Jahre 2018 wurde dann von OpenAI der Generative Pretrained Transformer (GPT) als eine Ab-

wandlung des Transformers mit einem vereinfachten Aufbau veröffentlicht [10]. OpenAI ist ein amerikanisches Forschungslabor für künstliche Intelligenz (KI) , das aus der gemeinnützigen OpenAI Incorporated (OpenAI Inc.) und ihrer gewinnorientierten Tochtergesellschaft OpenAI Limited Partnership (OpenAI LP) besteht. OpenAI betreibt KI-Forschung mit der erklärten Absicht, eine freundliche KI zu fördern und zu entwickeln . Die Organisation wurde 2015 in San Francisco von Sam Altman (OpenAI-CEO, ehemaliger Präsident des Startup-Beschleunigers Y Combinator), Reid Hoffman (Mitbegründer von LinkedIn), Jessica Livingston (Gründungspartnerin von Y Combinator), Elon Musk (Tesla- und SpaceX-Gründer), Ilya Sutskever (CTO, ehemaliger Google-Experte für maschinelles Lernen), Peter Thiel (PayPal-Mitbegründer) und anderen gegründet, die gemeinsam 1 Milliarde US-Dollar zugesagt hatten. Musk trat 2018 aus dem Vorstand zurück, blieb aber Spender. Microsoft hat OpenAI LP 2019 1 Milliarde US-Dollar und im Januar 2023 eine zweite mehrjährige Investition in Höhe von 10 Milliarden US-Dollar zur Verfügung gestellt.

Bedeutsam für die Funktion von GPT ist, dass das Sprachmodell nicht mehr für eine spezielle Aufgabe wie Übersetzung oder Klassifikation von Texten trainiert ist, für die häufig nur begrenzte Mengen an Beispieldaten zur Verfügung stehen. Stattdessen wurde das GPT-Modell auf sehr großen Datensätzen generischer Texte vortrainiert, um statistische Eigenschaften von Sprache als solcher unabhängig von der konkreten Aufgabenstellung zu lernen. Um dies zu verdeutlichen: Die Folgeversion GPT-2, die bereits 2019 erschien, enthielt 1,5 Milliarden Parameter, d. h. trainierte Gewichte der künstlichen Neuronen, die letztlich den Lerninhalt widerspiegeln; es wurde mit 40 GB Text trainiert. GPT-3 enthält bereits mehr als das Hundertfache davon, nämlich 175 Milliarden Parameter. Es wurde mit bereits 570 GB Klartext sowie 0,4 Billionen Token – wie gesagt kleinen Textteilen – trainiert. Im Einzelnen erfordert die Berechnung der Parameter eine sehr große Zahl von Rechenschritten; Abbildung 6 mag davon einen Eindruck vermitteln.

6.3 ChatGPT

Die Selbstoptimierungsfähigkeit von GPT-3 hatte allerdings noch den Nachteil, viel unerwünschten, falsch gewichteten oder schlicht falschen Text zu erzeugen, sprich zu diskriminieren, beleidigen etc. In einem Fachartikel der OpenAI-Entwickler gibt es ein lustiges Beispiel: Auf die Frage, warum man nach dem Meditieren Socken essen solle, antwortet das Sprachmodell ungerührt, dass man in den Socken die Essenz der Erleuchtung schmecke und dass stinkende Füße das wahre Aroma eines wachen Geistes seien. [11]

Für ChatGPT führte OpenAI daher noch einen Schritt ein, der ausgegebene Informationen in die richtige Richtung lenken soll. Eine Gruppe von rund 40 Mitarbeitern stellte über zehntausend Beispiele dafür zusammen, wie GPT bestimmte Aufgaben angehen sollte. Diese Beispiele galt es auch noch zu verallgemeinern. Dazu musste das Sprachmodell in die Lage versetzt werden, seine eigenen Ausgaben zu bewerten. Man konstruierte dazu ein eigenes Bewertungsmodell, das mit Aufsicht der Mitarbeiter trainiert wurde. Die dann verwendete Methode ist Reinforcement Learning (siehe Kasten oben). Auf GPT angewendet entspricht jede einzelne Entscheidung der Erzeugung eines neuen Wortes. Die Belohnung, die den Lernprozess steuert, ist die Qualitätsbeurteilung, welche sich das Sprachmodell nach der Erledigung einer Aufgabe selbst ausstellt. Wurde bei früheren Modellen nur darauf geachtet, das jeweils nächste Wort richtig

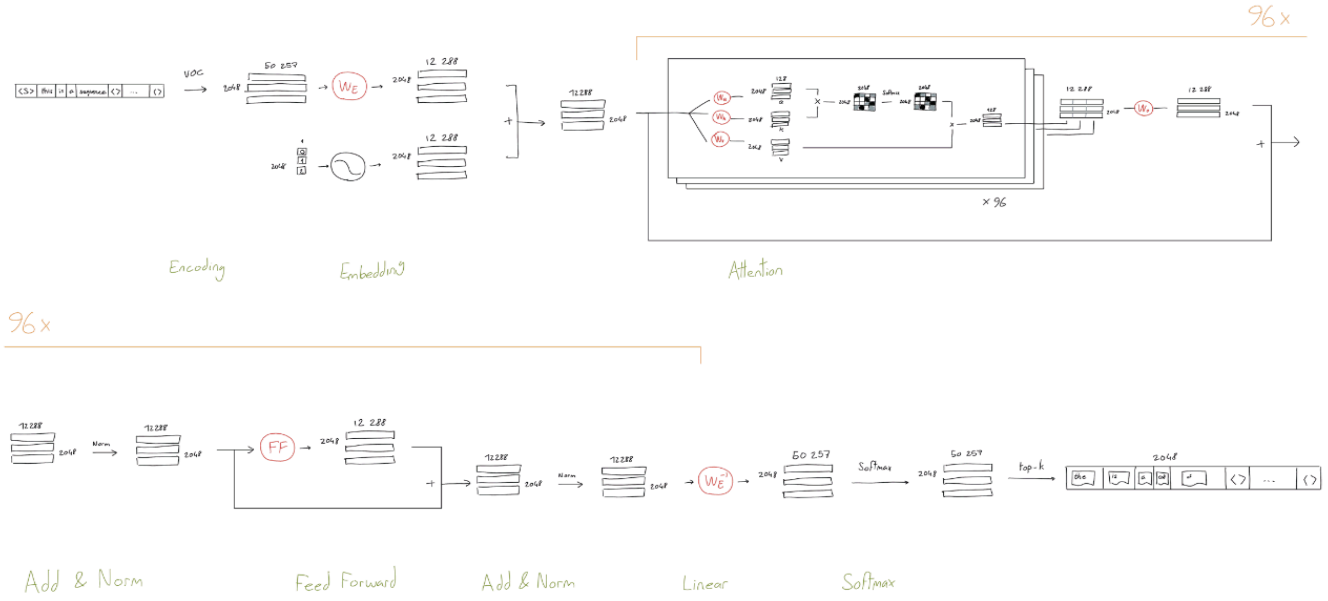


Abbildung 6: Die Neuronetz-Struktur von GPT-3 vermittelt die für seine durchführung nötige Rechenleistung (die zwei Etagen gehören nebeneinander gelegt (Abb.: [13]).

Simulated exams	GPT-4 estimated percentile	GPT-4 (no vision) estimated percentile	GPT-3.5 estimated percentile
Uniform Bar Exam (MBE+MEE+MPT)[1]	298/400 ~90th	298/400 ~90th	213/400 ~10th
LSAT	163 ~88th	161 ~83rd	149 ~40th
SAT Evidence-Based Reading & Writing	710/800 ~93rd	710/800 ~93rd	670/800 ~87th
SAT Math	700/800 ~89th	690/800 ~89th	590/800 ~70th
Graduate Record Examination (GRE) Quantitative	163/170 ~80th	157/170 ~62nd	147/170 ~25th
Graduate Record Examination (GRE) Verbal	169/170 ~99th	165/170 ~96th	154/170 ~63rd
Graduate Record Examination (GRE) Writing	4/6 ~54th	4/6 ~54th	4/6 ~54th
USABO Semifinal Exam 2020	87/150 99th–100th	87/150 99th–100th	43/150 31st–33rd
USNCO Local Section Exam 2022	36/60	38/60	24/60
Medical Knowledge Self-Assessment Program	75%	75%	53%
Codeforces Rating	392 below 5th	392 below 5th	260 below 5th
AP Art History	5 86th–100th	5 86th–100th	5 86th–100th
AP Biology	5 85th–100th	5 85th–100th	4 62nd–85th
AP Calculus BC	4 43rd–59th	4 43rd–59th	1 0th–7th
AP Chemistry	4 71st–88th	4 71st–88th	2 22nd–46th
AP English Language and Composition	2 14th–44th	2 14th–44th	2 14th–44th
AP English Literature and Composition	2 8th–22nd	2 8th–22nd	2 8th–22nd
AP Environmental Science	5 91st–100th	5 91st–100th	5 91st–100th
AP Macroeconomics	5 84th–100th	5 84th–100th	2 33rd–48th
AP Microeconomics	5 82nd–100th	4 60th–82nd	4 60th–82nd
AP Physics 2	4 66th–84th	4 66th–84th	3 30th–66th

Abbildung 7: Examensergebnisse von GPT-4 im Vergleich zu GPT-3.5, dem Modell hinter ChatGTP (Abb.: [14]).

vorherzusagen, lernt mit Reinforcement Learning hingegen lernt das ChatGPT-Modell, dieses Wort so zu wählen, dass die Qualität der gesamten Antwort am Ende möglichst hoch ist.

6.4 GPT-4

Es kann natürlich immer noch Unsinn liefern – was sehr viele Nutzer bestätigen. Aber die Entwicklung geht mit Riesenschritten weiter. Kurz nachdem CLB-Autor Volker Wiskamp den in dieser Ausgabe gedruckten Artikel über ChatGPT eingereicht hatte, wurde von OpenAI das Modell GPT-4 am 14. März 2023 veröffentlicht. Erfreulicherweise konnte Volker Wiskamp gleich einen fortgeschrittenen Test damit durchführen, mit bemerkenswerten Ergebnissen (siehe auch diese Ausgabe).

GPT-4 ist kein reines Sprachmodell mehr, sondern kann neben Texteingaben auch mit Bildern umgehen. Damit liegt eine multimodale Fähigkeit vor, wie sie übrigens in Deutschland mit einem eigenen Modell die Heidelberger Firma Aleph Alpha von Jonas Andrulis mit dem Produkt „Luminous“ anbietet. [12] OpenAI produzierte zwei Versionen von GPT-4 mit Kontextfenstern von 8192 und 32 768 Token, eine deutliche Verbesserung gegenüber GPT-3.5 – das ChatGPT zugrunde liegt – und GPT-3, die auf 4096 bzw. 2049 Token beschränkt waren. Kurz gesagt: GPT-4 lässt sich von vorn herein mit einer Eingabe bis zu rund 25 000 Zeichen füttern.

Laut OpenAI soll das Modell kreative und technische Schreibaufgaben ausführen können, Songexte komponieren, Drehbücher schreiben oder auch den Stil seiner Nutzer imitieren können. Allerdings ist die Fähigkeit zum Erzeugen gewalthaltiger oder sonstwie schädlicher Inhalte offenbar nicht gebannt – aber mit einer deutlich geringeren Wahrscheinlichkeit als seine Vorgängermodelle, nach internen Tests um 82 Prozent reduziert als bei GPT-3.5, das Modell hinter ChatGPT. Zudem soll GPT-4 in Tests, die in den USA für die Aufnahme in Universitäten gestellt werden, erheblich besser abschneiden als ChatGPT (Abbildung 7)[14].

Für den SAT-Test Mathematik steigerte das neue Modell sein Positiv-Ergebnis von 70 auf 89 Prozent. Im AP-Test für Chemie verbesserte es sich von der Note 2 auf die Note 4 (Notenskala 1 = keine Empfehlung, 5 = sehr gut qualifiziert). Der SAT (Scholastic Assessment Test) ist ein US-amerikanischer standardisierter Test, der hauptsächlich von Studienplatzbewerbern an amerikanischen Universitäten gefordert wird (Studierfähigkeitstest). Das Advanced Placement Program (abgekürzt als AP) ist ein pädagogisches Programm in den USA und Kanada, welches Kurse auf College-Niveau an US-amerikanischen und kanadischen Highschools anbietet.

Kurze Zeit nach der OpenAI-Veröffentlichung über die Leistungen von GPT-4 gaben Mikrosoft-Forscher ein Paper heraus, das noch über erheblich darüber hinausgehende Fähigkeiten berichtet. [15] Einige Punkte:

- GPT-4 hat wohl trotz seines rein textbasierten Trainings eine Art visuelles Vorstellungsvermögen entwickelt. Das ist etwas anderes als bei der ebenfalls von OpenAI verfügbaren KI-Plattform Dall-E zur Generierung von Bildern: GPT-4 ist eben nur mit Text, nicht mit Bilder trainiert worden.
- GPT-4 soll auch ein rudimentäres Verständnis für Musik entwickelt haben.
- Das Sprachmodell soll in der Lage sein, mathematische Probleme zu lösen – wobei diese sich nicht auf Rechenaufgaben, sondern auf mathematische Beweise auf Hochschulniveau beziehen. Generell zeige GPT-4, dass es abstrakte, komplexe Probleme inhaltlich verstehen und häufig sinnvolle Lösungsansätze wählen kann. Ein Beispiel dafür bezieht sich auf das Fermi-Problem, auf den Umgang mit ungenauen Daten. Der italienische Kernphyiker und Nobelpreisträger Enrico Fermi (1901-1954), der an der Entwicklung der Atombombe in den USA beteiligt war, stellte dazu seinen Studenten eine klassische Aufgabe dieses Typs. Es ist die Frage nach der Anzahl der Klavierstimmer in Chicago. GPT-4 löst solche Aufgaben auf sinnvolle Weise, indem es beispielsweise abschätzt, wie viele Haushalte es in Chicago gibt, wie viele davon ein Klavier besitzen, wie häufig diese gestimmt werden müssen und so weiter.
- GPT-4 hingegen hat die Fähigkeit erworben, externe Werkzeuge oder Module einzusetzen. Das Sprachmodell kann nun zum Beispiel selbstständig neue Inhalte aus dem Internet suchen und so Anfragen mit aktuellen Daten beantworten.
- Das Modell sein Weltwissen generalisieren und zur Navigation in einer Umgebung einsetzen kann. Es kann ein Gebäude erkunden (nach Spracheingabe) und es dann auch zeichnen. Dies unterstreicht die bemerkenswerte Fähigkeit von GPT-4, sowohl räumliche Informationen aus Texten zu extrahieren als auch diese Informationen in eine visuelle Darstellung umzuwandeln, obwohl das Modell ausschließlich auf Texten trainiert wurde.
- Die Microsoft-Forscher kommen auch zu dem Schluss, dass GPT-4 in der Lage ist, das Denken und Fühlen von Menschen in einer gegebenen Situation einzuschätzen. Dabei kann das Modell präzise zwischen den Standpunkten unterschiedlicher Personen differenzieren, die dieselbe Situation aus verschiedenen Perspektiven erleben.

Ein Kommentator zu diesem Artikel – Helmut Linde, Head of Digital R&D, Covestro – folgert: „Nach der Lektüre der zahlreichen Fallbeispiele fällt es schwer, die Position aufrechtzuerhalten, dass die Entwicklung einer AGI ((Artificial General Intelligence)) noch viele Jahrzehnte entfernt liegen oder sogar technisch grundsätzlich unmöglich sein könne.“[16]

Linde sieht auch Entwicklungspotenzial: „Eine der größten Limitationen von GPT-4, wie die Autoren betonen, besteht darin, dass das Modell seine Ausgaben stets Wort für Wort erzeugt. Die Lösung einer Aufgabe muss daher in gewissem Sinne immer in Vorwärtsrichtung erfolgen, wobei das Ende einer Ausgabe zumindest abstrakt schon am Anfang vorgedacht sein muss. Im Gegensatz dazu können Menschen in einer Art innerem Dialog ihre Antwort gedanklich zu Ende bringen und, wenn sie auf einen Widerspruch stoßen, zurück an den Anfang gehen und ihren Gedankenfluss anpassen. Eine mögliche Verbesserung für die Zukunft könnte also darin bestehen, das Modell um eine innere Feedback-Schleife zu erweitern, so dass es nicht nur linear eine Antwort erzeugen, sondern seinen eigenen Denkprozess überwachen, beurteilen und lenken kann.“ Zudem könnten mehrere Modelle auf unterschiedlichen Abstraktionsebenen arbeiten und sich gegenseitig beeinflussen.

Die Microsoft-Forscher betonen noch etwas Bemerkenswertes: Sie räumen ein, dass niemand wirklich versteht, warum diese Modelle so gut funktionieren. Viele Fähigkeiten seien emergente Eigenschaften, die während des Trainings der Modelle entstünden, sich aber bisher nicht direkt auf die Architektur der Modelle zurückführen ließen.

7. Ein kleiner Ausblick

Die Entwicklung der KI verläuft rasant. Teilweise – und wirtschaftlich erfolgreich – nutzt man ähnlich wie bei der Chipentwicklung nach dem Mooreschen Gesetz Skalierungsmechanismen; das zeigt ja die bisherige Transformer-Entwicklung. Teilweise – eher noch in der Forschung – entwickelt man aber auch völlig neue Strukturen.

7.1 Skalierung – auch im Kleinen

Die Skalierung betrifft hauptsächlich den Aufwand zur Erstellung eines Sprachmodells, mit Blick sowohl auf die Hard- wie auch auf die Software. Dabei gibt es ganz neu jetzt auch Tendenzen einer Skalierung nach unten: GPT-artige Sprachmodelle sollen auch auf kleineren Rechenumgebungen laufen. Skaliert wird hier also in der Breite, hin auf viele Anwender.

Drei Wochen vor dem Launch von GPT-4 durch OpenAI hat Facebooks Mutterkonzern Meta ein eigenes großes Sprachmodell (LLM) präsentiert: das mit offenen Daten vortrainierte LLaMA (Large Language Model Meta AI). Von Chat-GPT nehmen Experten an, dass man es erst mit erheblichem Hardware-Aufwand zum Laufen bringen kann: Mindestens fünf Nvidia-A100-Server-Grafikprozessoren

mit jeweils 80 GB RAM seien nötig, zusammen für über 83 000 Euro zu erwerben. Der A100-Chip von Nvidia hat 54 Milliarden Transistoren, ist in 7 nm-Strukturgröße ausgeführt. Zudem sind die großen OpenAI-Produkte – GPT-3 und GPT-4 – gar nicht „open“, sondern Kästen, von denen man nicht weiß, was drin ist, Blackboxes eben.

LLaMA ist ein mit GPT vergleichbares Sprachmodell, nur deutlich offener. Anders als OpenAI sagt Meta zum Beispiel ganz genau, mit welchen Daten sie das Modell trainiert haben. Den größten Teil, nämlich 67 %, machen fünf englische Speicherinhalte (Dumps) von Internet-Saugern (Crawler) aus, die jeweils den Inhalt von drei Milliarden Websites beinhalten! 4,5 Prozent der Lama-Trainingsdaten sind Code von GitHub, weitere 4,5 Prozent Wikipedia-Dumps in unterschiedlichen Sprachen, 4,5 Prozent Bücher von Project Gutenberg und Book3 – und noch weitere Quellen, insgesamt 4,7 Terabyte reiner Text. In der Ausdrucksweise der Sprachmodelle: LlaMA hatte ein Training mit 1,4 Billionen Token (ein kleinerer Ableger mit immer noch 1 Billion Token). [17]

All das kostet. Meta sagt erfreulicherweise ganz genau, wie viel Energie das Training von Lama gekostet hat. Laut Facebook haben 2048 dieser schon erwähnten 80-Gigabyte-A100-Karten fünf Monate lang geackert. Das hat 2638 Megawattstunden Energie gekostet und 1015 Tonnen CO_2 ausgestoßen. Kostenmäßig rechnet man so ungefähr einen US-Dollar für eine A100-Stunde in der Cloud und das ergibt 7,3 Millionen US-Dollar. [18] Eigentlich wollte Meta den Zugriff auf das Sprachmodell und die Trainingsdaten regeln – zwar als kostenlose Testmöglichkeit, aber für ausgesuchte Nutzer.

Alles landete aber irgendwann im Netz, ist jetzt frei verfügbar. Stanford-Forscher haben damit noch ein genauere Abstimmung betrieben und das Alpaca-Modell erzeugt. Kann LlaMA nur Sätze vervollständigen, reagiert Alpaca wie von den GPT-Modellen gewohnt auf Fragen. Alpaca ist zur Zeit auch frei verfügbar, in einer rechtlichen Grauzone. Weder Meta noch OpenAI, mit deren Daten und Modellen die Feinabstimmung lief, haben auf das Modell auf Stanford bislang reagiert. Der Heise-Autor Jan-Keno Janssen rät davon ab, Alpaca kommerziell zu verwenden, erwartet aber im Grunde bald frei verfügbare Sprachmodelle für jeden. [18] Das Ganze funktioniert jedoch auf einer normalen Rechnerausstattung, egal ob Windows, Mac oder Linux verwendet wird – mit Abstrichen sogar auf dem Raspberry Pi, ein bekannter Einplatinen-Kleincomputer für unter 100 Euro. [19]

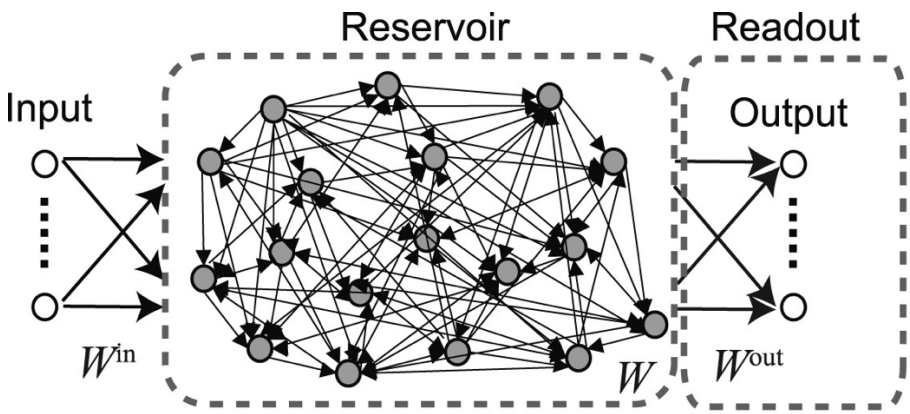


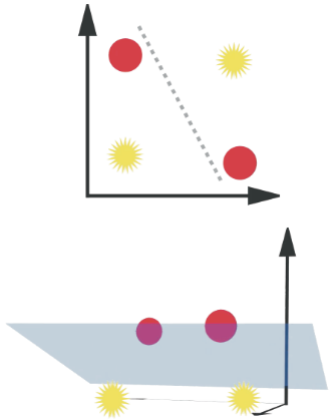
Abbildung 8: Schema des Reservoir-Computing. Das RNN-basierte Reservoir in der Mitte ist sehr komplex, weist viele Rückkopplungen auf. Zum Auslesen (rechts) genügt ein einfaches lineares Netzwerk (Abb.: [20]).

7.2 Hochkomplex: Reservoir-Computing

Am anderen Ende der Entwicklung stehen weitere neue Netzstrukturen. ChatGPT kann seinen Anwendern ja schon vermitteln, er versteht seine Sachen – nur stimmt das nicht. **ChatGPT ist verständlich, hat aber kein Verständnis.** Dieses Sprachmodell arbeitet statistisch und hat selbst eben keinerlei Verständnis dafür, was es produziert – von der Frage eines Bewusstseins völlig abgesehen. Gerade letztgenanntes ist jedoch auch ein starker Antrieb, Forschungen in KI zu intensivieren. Wird es doch einmal möglich sein, eine Struktur mit einem künstlichen Bewusstsein zu erzeugen?

Mittlerweile gibt es hochkomplexe Netze für schnelles Lernen. Gerade auch im Zusammenhang mit photonischen integrierten Chips – einer ganz neuen Richtung, extrem schnelle, datenverarbeitende Hardware zu entwickeln [3] – liest man beispielsweise zunehmend von „Reservoir-Computing“ (RC). Bei RC handelt es sich um eine Netzwerkstruktur, die aufwändiges Training vermeiden will – aber natürlich andere Hürden stellt. Als Kern verwendet Reservoir-Computing rekurrente neuronale Netze (RNN). Wie schon gesagt zeichnen sie sich durch Verbindungen von Neuronen einer Schicht zu Neuronen derselben oder einer vorangegangenen Schicht aus. Sie haben große Bedeutung für selbstlernende Systeme; der Lernprozess ist aber oftmals aufwändig. Im biologischen Gehirn verfügen die Neuronen üblicherweise auch über solche Rückkopplungen, insbesondere im Neocortex des Menschen.

Abbildung 9: Dimensionen und Reservoir Computing: Die gelben und roten Punkte lassen sich in der zweidimensionalen Darstellung (oben) nicht durch einen, sonder nur durch zwei gerade Linien voneinander teilen. In der dreidimensionalen Darstellung unten gelingt dies durch eine einzige Ebene. Ein RC-System kommt zu dieser Lösung, indem das komplexe Reservoir mit dem rekurrenten Netzwerk die dritte Dimension erzeugt und das lineare Netzwerk am Ausgang des Reservoirs die Trennungsebene legt (Abb.: [22]).



Beim Reservoir-Computing sind ähnlich wie bei den GPT-Modellen RNNs im Spiel, die schon vortrainiert sind. Dieser RNN-Kern ist quasi „Black Box“ in dem Gesamtkonstrukt (Abbildung 8). Darin speist man Daten ein, die in der Black Box – deren Funktion man gar nicht kennt – prozessiert werden (da stehen den kontroll-affinen Datenschützern mal wieder die Haare zu Berge...).

Diese Black Box, das Reservoir, ist ein rekurrentes Netz mit nichtlinearen Eigenschaften, anders als bei den GPT-Modellen mit zufällig und komplex verbundenen, rückgekoppelten neuronalen Knoten, dessen Verbindungen nicht verändert werden. Es weist drei Charakteristika auf: Einmal hat es eine hohe Anzahl von Knoten, um hochdimensional arbeiten zu können. Zum anderen verfügt es über Nichtlinearitäten, um verschiedenste Aufgaben lösen zu können, und schließlich hat es noch eine Art Kurzzeitgedächtnis, das durch Rückkopplungen zwischen den Neuronen realisiert wird. Zum Verständnis eines Satzes muss man ja beispielsweise die Wortreihenfolge erschließen können.

Wenn ein Eingangssignal in das Reservoir eingebracht wird, wird es auf einen höherdimensionalen Raum abgebildet, in der es linear trennbar werden kann (Abbildung 9). Der Vorteil von Reservoir Computing ist, dass die Netzwerke keine Abstimmung benötigen: alle Verbindungen innerhalb des Reservoirs bleiben fest. Nur externe Verbindungen (zwischen dem Reservoir und einer Ausgabeschicht) werden trainiert, um die gewünschte Aufgabe zu lösen. Mit anderen Worten, die Reservoirberechnung erfordert die Erzeugung komplexer nichtlinearer Dynamik, aber das Lernen wird stark vereinfacht. Für eine effiziente Reservoirberechnung müssen mehrere Anforderungen erfüllt sein in Bezug auf die dynamischen Eigenschaften des Netzwerkes. Zunächst sollten verschiedene Eingaben unterschiedliche Dynamiken auslösen, während ähnliche Eingaben ähnliche Dynamiken aufweisen sollten, sodass eine effiziente Klassifizierung ermöglicht wird. Zudem sollte der Zustand des Reservoirs nicht nur von den gegenwärtigen Eingaben, sondern auch von den Eingaben der jüngsten Vergangenheit abhängen. Dieses Kurzzeitgedächtnis (fading memory) ist wesentlich für die Verarbeitung zeitlicher Sequenzen, bei denen der Signalverlauf eine Rolle spielt.

Um zu einem benutzerdefinierten Verhalten zu gelangen, müssen Reservoir-Systeme nicht modelliert oder programmiert werden. Man betreibt sie lediglich mit überwachten maschinellen Lernverfahren für einfache lineare Klassifikatoren. Dies ist nämlich der notwendige Schritt, um die Ausgabe aus der Black Box in ein gewünschtes Ergebnis zu verwandeln. Reservoir-Computing ist vorteilhaft in Systemen, deren Gesamtverhalten komplex ist, die jedoch schwierig zu modellieren oder einem theoretischen Verhalten nach anzupassen sind.

Tatsächlich könnte jedes nichtlineare dynamische System möglicherweise zu einem Reservoir werden. Es muss ja lediglich die Fähigkeit haben, eine einfache Eingabe in eine höhere Dimension zu projizieren. Der Grund dafür ist, dass man, wenn das System nicht in Dimensionen projizieren kann, die höher als die Eingabe sind, die Eingabe genauso gut direkt in das Auslesenetzwerk einspeisen und ein besseres Ergebnis erzielen können.

Wissenschaftler nutzten sogar eine Population von E coli-Bakterien als Reservoir [20]. Die Bakterien reagieren auf verschiedene chemische Stoffe und erzeugen durch den Kontakt mit diesen komplexe zeitliche Muster. Dann wurden Proben von E. coli in ein Microarray gebracht, um die mRNA- und Proteinkonzentrationen zu messen, die den Zustand des Genregulationsnetzwerks darstellen. Zur Auswertung verwendete man ein einfaches neuronales Netz, ein Perceptron. Es lernte, chemische Reize und Bewegungsreaktionen in entsprechende Klassen einzuordnen.

Numerische Simulationen, die dem obigen experimentellen Design entsprechen, bestätigten, dass ein so gestaltetes RC-System ein XOR-Problem lösen kann. Später verwendete man gentechnisch veränderte Bakterienzellkulturen, um eine Vielzahl einfacher Klassifikatoren zu realisieren, die verschiedene chemische Inputs trennen und zu einem komplexeren Klassifikator zusammengefasst werden. Dieses Design des maschinellen Lernens in der synthetischen Biologie kann auf einen Reservoircomputer erweitert werden.

Es gibt Hinweise darauf, dass das Prinzip des Reservoir-Computings auch im menschlichen Gehirn zum Einsatz kommt. Wenn wir lernen, ein neues Wort zu buchstabieren, wächst in unserem Gehirn keine neue Verbindung zwischen Neuronen und Synapsen. Stattdessen lernen wir, die Reihe der Gehirnaktivitäten auszulesen und dies als die Bedeutung des Wortes zu interpretieren. Aus diesem Grund bleibt das Reservoir fest, während sich das Auslesenetzwerk ständig weiterentwickelt.

7.3 Reservoir-Computing und Spike-Computing

Noch eine bemerkenswerte Entwicklungsrichtung, die aber ebenfalls erst in der Forschung anzusiedeln ist: Reservoir Computing stellt eine Alternative zu Spike-Computing dar, mit der sehr komplexe Aufgaben mit verhältnismäßig geringem Aufwand gelöst werden können. Biologische Neuronen kommunizieren über binäre Zustände, den Spikes, nicht über kontinuierliche Werte. Zudem arbeiten sie zeitabhängig, d. h. ihre Aktivität variiert nicht nur mit der Intensität der Eingabe, sondern auch mit der zeitlichen Abfolge der Spikes. Dieses Prinzip nennt man „spike timing dependent

plasticity“ (STDP), die entsprechenden Neuronetze SNN für spiking neural nets. Dadurch, dass solche Netze nur einzelne „Feuerimpulse“ freisetzen und nicht einem kontinuierlichen Systemtakt unterliegen, arbeiten sie sehr energieeffizient. Die Entwicklung solcher Biologie-nahen Computer läuft unter dem Begriff „neuromorphe Systeme“. Die CLB hat darüber beispielsweise in [2] und [3] berichtet.

Neuromorphe Spike-Verarbeitung und Reservoir-Ansätze unterscheiden sich grundlegend und haben komplementäre Vorteile [21]. Beide leiten ein breites Repertoire an Verhaltensweisen (oft als Komplexität bezeichnet) aus einer großen Anzahl physikalischer Freiheitsgrade (z. B. Intensitäten) ab, die durch Interaktionsparameter (Übertragungen) gekoppelt sind. Beide bieten die Möglichkeit, aus diesem Repertoire mittels steuerbarer Parameter ein bestimmtes, gewünschtes Verhalten auszuwählen.

In neuromorphen Systemen sind Netzwerkgewichte sowohl die Interaktions- als auch die steuerbaren Parameter, während in Reservoircomputern diese beiden Parametergruppen getrennt sind. Diese Unterscheidung hat wesentliche Auswirkungen. Erstens müssen die Interaktionsparameter eines Reservoirs nicht beobachtbar oder gar von System zu System wiederholbar sein. Zudem benötigen sie keine nennenswerte Hardware, um den Zustand des Reservoirs zu steuern. Schließlich können Reservoircomputer nur durch beaufsichtigtes Training dazu gebracht werden, ein gewünschtes Verhalten hervorzurufen, während neuromorphe Computer a priori unter Verwendung eines bekannten Satzes von Gewichten programmiert werden können. Der wesentliche Vorteil eines Reservoir-Computing-Systems mit Black Box (dem Reservoir) und einfachem Ausgabe-Netzwerk besteht in vereinfachtem Training. Reservoir Computing reduziert trainingsbedingte Herausforderungen, indem die Dynamik des Reservoirs festgelegt und nur die lineare Ausgangsschicht trainiert wird. Das bedeutet natürlich auch, dass man den Aufwand für eine geeignete Funktion des Gesamtsystems vorher betreiben muss.

7.4 Fazit

Dieser Ausblick kann vielleicht anzeigen, wie vielfältig die Entwicklungsmöglichkeiten sind; weitere Details gibt es zum Beispiel in [23]. Eins lässt sich mit Gewissheit sagen: ChatGPT war ein Weckruf, der aufzeigt, wozu KI fähig sein kann. Mit dem Sprachmodell hat zudem eine umfangreiche Kommerzialisierung und Verbreitung von KI begonnen. Die Einsatzmöglichkeiten sind nahezu undenkbar. Dass auch die Ausbildung darauf reagieren muss, zeigen in dieser CLB die Artikel von Volker Wiskamp sowie mein Editorial. Weiterentwickelte KI-Modelle werden kaum zu überschätzen sein. ▶▶|

Literatur / Quellen

[1] R. Kickuth: Künstliche neuronale Netze et al. – Algorithmen (auch) nach der Biologie; CLB 9 2003, S. 333-339

[2] R. Kickuth: Von der Bastelphase zur Eiszeit – und zum Wiedererwachen: Geschichte der künstlichen Intelligenz mit Auf und Ab sowie Nebenentwicklungen; CLB 3/4 2018, S. 112-135

[3] R. Kickuth: Von neuromorph über Protein-Nanodraht bis zu photonischen Schaltkreisen; CLB 3/4 2020, S. 154-177

[4] Warren McCulloch und William Pitts: A logical calculus of the ideas immanent in nervous activity. In: Bulletin of Mathematical Bio- physics, Bd. 5 (1943), S. 115–133, ISSN 0007-4985

[5] Donald Hebb: The organization of behavior. A neuropsychological theory. Erlbaum Books, Mahwah, N.J. 2002, ISBN 0-8058-4300- 0 (Nachdruck der Ausgabe New York 1949)

[6] Rumelhart, David E.; Hinton, Geoffrey E.; Williams, Ronald J. (1986-10-09). „Learning representations by back-propagating errors“. Nature. 323 (6088): 533–536.

[7] https://venturebeat.com/2020/01/02/top- minds-in-machine-learning-predict-where-ai-is- going-in-2020/topm2020

[8] https://ai.googleblog.com/2017/08/transformer- novel-neural-network.htmltran

[9] https://www.golem.de/news/kuenstliche-intelligenz-so-funktioniert-chatgpt-2302-171644-3.html (sehr lesenswerter Artikel von Helmut Linde, Head of Digital R&D, Covestro)

[10] https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

[11] https://arxiv.org/pdf/2203.02155.pdf, 4. März 2022

[12] https://www.aleph-alpha.com

[13] https://dugas.ch/artificial_curiosity/GPT_architecture.html, ein Blog von Daniel Dugas

[14] https://openai.com/research/gpt-4; 14. 3. 2023

[15] Sparks of Artificial General Intelligence: Early experiments with GPT-4; https://doi.org/10.48550/arXiv.2303.12712; 22. 3. 2023

[16] https://www.golem.de/news/gpt-4-funken-von-allgemeiner-kuenstlicher-intelligenz-2303-172962.html

[17] https://ai.facebook.com/blog/large-language-model-llama-meta-ai/ (Meta AI vom 24.2.2023)

[18] https://www.heise.de/news/c-t-3003-ChatGPT-Klon-laeuft-lokal-auf-jedem-Rechner-Alpaca-LLaMA-ausprobiert-8004159.html

[19] https://www.heise.de/hintergrund/Grosses-KI-Sprachmodell-am-eigenen-Rechner-Ein-LLaMA-fuer-die-Westentasche-7624131.html?seite=all

[20] Tanaka G. et al.: Recent advances in physical reservoir computing: A review; Neural Networks Volume 115, July 2019, Pages 100-123

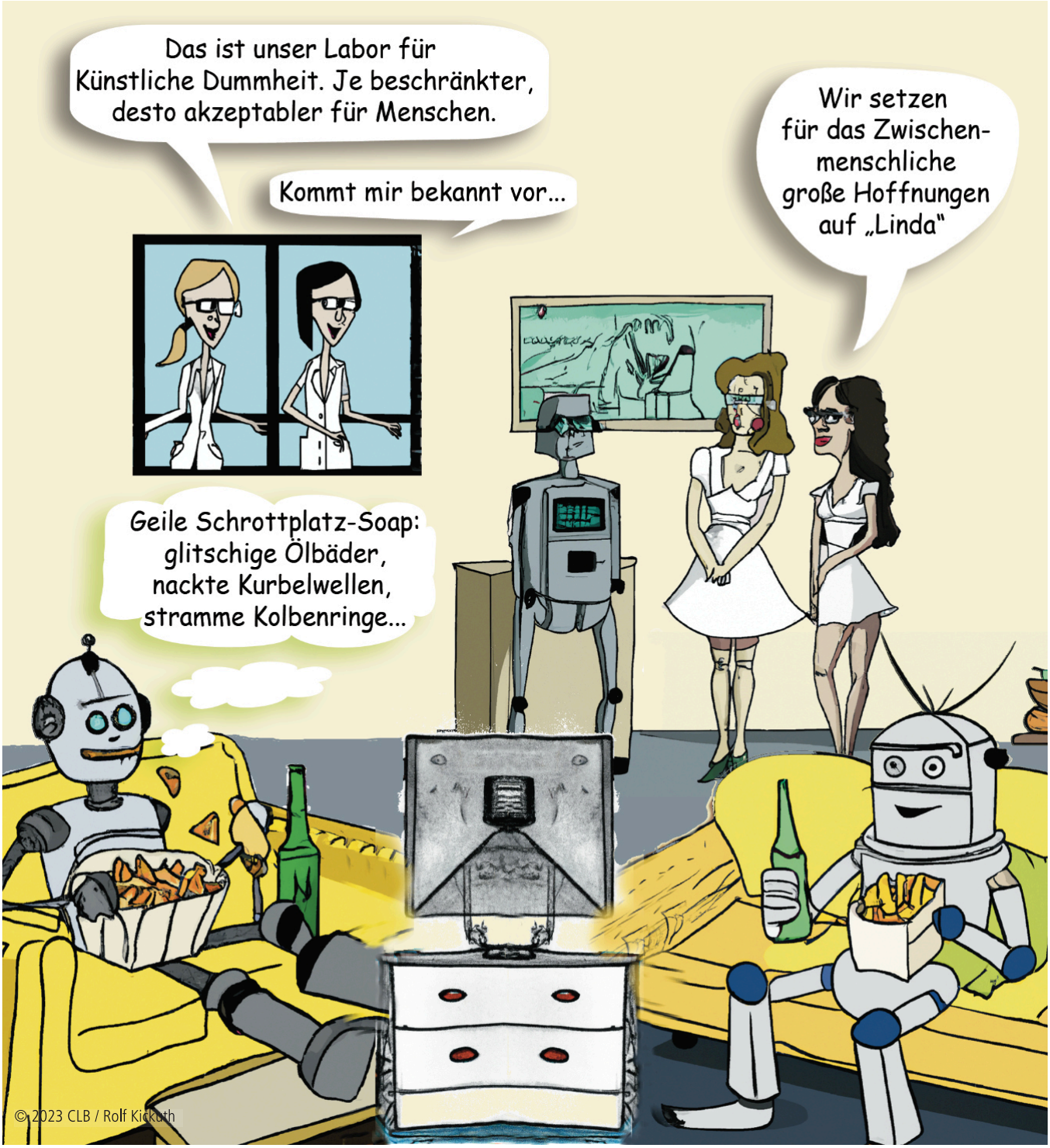
[21] Shastri B. J. et al.: Principles of Neuromorphic Photonics; arXiv:1801.00016v1 [cs.ET] 29 Dec 2017

[22] Irene Estébanez in dem YouTube-Video „Qué es el Reservoir Computing“, Instituto de Física Interdisciplinar y Sistemas Complejos, Univ. der Balearen

[23] R. Kickuth: Bio-inspired Computing; 380 Seiten, 219 meist mehrfarbige Abbildungen; Agentur und Verlag Rubikon, August 2020

Soweit kommt es noch:

Auf Blödheit trainiert – der Akzeptanz wegen…



Dieser Cartoon weist auf ein Problem hin, mit dem sich unsere Gesellschaft hinsichtlich der KI auch auseinandersetzten muss: Der Urheberschaft durch KI erzeugter Werke. So sind sowohl das Titelbild dieser CLB wie auch der Cartoon hier mit Hilfe von Systemen künstlicher Intelligenz generiert worden. Trotzdem steht da mein Name drunter: Die Idee und das Konzept stammen von mir – die „Prompts", Anweisungen, die ich der KI gab, das Aussuchen ihrer Vorschläge, das Zusammenstellen von Einzelbildern – der Cartoon besteht aus fünf Teilen, die verbessert und zusammengestellt wurden, mit Sprechblasen und deren Inhalten versehen… Ohne diesen kreativen Input meinerseits wäre die KI aufgeschmissen gewesen. Menschen geht es allerdings bei kreativen Prozessen durchaus ähnlich. Wir bekommen von Schöpfern anderer Werke ja auch Anregungen, sei es durch Musik, visuelle Kunst, Text oder sogar auch Mathematik. Von daher sollten die neuen Möglichkeiten von ChatGPT, Dall-e und Co. als Werkzeuge genutzt werden, um Ausbildungen zu verbessern, kreativer zu machen, Arbeitsvorgänge von Routine zu entlasten. Es wird gelten, die Anteile der KI-Werkzeuge deutlich zu machen (siehe auch mein Editorial). *RK*