

Um die Ehre

Garri Kasparow ist ein unglaublich schlechter Verlierer. Als er, der amtierende Schachweltmeister, 1997 gegen IBMs Deep Blue Supercomputer antrat, sah *Newsweek* in Kasparow die letzte Verteidigungslinie gegen die intellektuelle Übermacht der Computer. Ein Jahr vorher hatte Kasparow noch gegen Deep Blue gewinnen können, aber dann entschloss sich IBM, sein ganzes Gewicht hinter das Prestigeprojekt zu legen. Kasparow verlor spektakulär – und beschuldigte IBM geschummelt zu haben. Sein Ego war stellvertretend für uns alle sichtbar angeknackst. Ohne Zweifel war das eine Niederlage mit hoher Symbolkraft. Aber gegen wen oder was hat Kasparow eigentlich verloren?¹

Beim Schach gibt es grob geschätzt 10^{43} verschiedene Spielzustände.² Eine Zahl mit 43 Nullen. Zum Vergleich: Im Gehirn haben wir so

1 Dass er ein schlechter Verlierer ist, gibt Kasparow selber freimütig zu: 30 Jahre nach seiner Niederlage gegen Deep Blue beschreibt er in seinem Buch seine Version der Geschehnisse und zieht daraus Lehren für unser Verhältnis zu KI (Kasparow, 2017). In einem kurzen Artikel macht das auch der Journalist, der das Spiel zwischen Kasparow und Deep Blue damals für *Newsweek* beobachtet hat (Levy, 2017).

2 Diese Abschätzung stammt von Shannon (1950). Er ignoriert dabei, dass nicht jede Figur bei Einhaltung der Spielregeln jede Brettposition erreichen kann und dass Figuren während des Spiels geschlagen werden und dann nicht mehr da sind. Zählt man aber einfach mal alle Möglichkeiten, wie man die 16 weißen und die 16 schwarzen Figuren aufstellen kann, erhält man ein Gefühl für die Größenordnung. Die erste Figur kann auf eines der 64 Felder gestellt werden. Die zweite auf die übrigen 63, und so weiter. Für die 32 Figuren gibt es bei 64 Feldern also $64 \cdot 63 \cdot 62 \cdot \dots \cdot 34 \cdot 33$ Möglichkeiten sie aufzustellen. Mit der Fakultät kann man das auch als $64!/32!$ schreiben. Da jeder acht Bauern hat, die identisch sind, spielt es keine Rolle, wo die Bauern stehen, also muss diese Zahl durch $8!$ geteilt werden. Für die 6 anderen Figuren, die jeweils paarweise auftreten (Läufer, Pferde, Türme für jede Farbe), muss man die Zahl weiter durch 2^6 teilen. So kommt man auf etwa 10^{43} (beim Nachrechnen lieber mit dem Logarithmus arbeiten, da Taschenrechner und Computer bei so großen Zahlen schnell überfordert sind).

etwa 10^{11} , also 100 Milliarden, Nervenzellen. So viele Transistoren finden sich auch auf dem M2 Ultra Prozessor von Apple, der 2023 auf den Markt kam. In der Mitte des Jahrhunderts werden auf der Erde etwa 10 Milliarden (10^{10}) Menschen leben. Alle Menschen zusammen kommen dann auf 10^{21} Gehirnzellen. Descartes stellte sich noch Maschinen vor, denen man für jede mögliche Situation einzeln Regeln vorgeben muss, wie sie sich zu verhalten haben. Es ist tatsächlich »praktisch unmöglich«, dass eine Schachmaschine für jede Spielsituation ein eigenes »Organ« hat. Diese Maschine müsste deutlich mehr interne Zustände haben, als alle Menschen zusammen Gehirnzellen.

Dann muss der Computer halt den nächsten Zug ausrechnen, könnte man meinen. Allerdings können auch die größten Supercomputer auf der Welt noch kein Schachspiel komplett im Voraus planen. Die Anzahl an möglichen Zügen in einem Schachspiel ist erstaunlich konstant über die Partie hinweg, ungefähr 30. Machen Schwarz und Weiß je einen Zug, gibt es also etwa 30·30, ungefähr 1000 Möglichkeiten, wie sich das Spiel in jeder Runde weiterentwickeln kann. In einem typischen Spiel macht jeder Spieler 40 Züge, das heißt, es gibt in etwa $1000^{40} = 10^{120}$ mögliche Schachpartien.³ Seit dem Urknall sind so etwa 10^{17} Sekunden vergangen und ein aktueller Supercomputer kann 10^{17} Berechnungen pro Sekunde durchführen. Um auszurechnen, was die beste Schachstrategie ist, war bisher einfach noch nicht genügend Zeit. Trotzdem können Menschen und Computer Schach spielen.

Wie spielen Menschen und Computer Schach?

Am Anfang des Spiels folgen beide größtenteils Descartes naiver Vorstellung: Sie merken sich für eine große Zahl an Spielzuständen den Zug, den sie in dieser Situation machen wollen. Schachanfänger lernen verschiedene Eröffnungen auswendig, die alle wenig originelle Namen tragen, zum Beispiel das Vierspringerspiel oder die Spanische Eröffnung. Für diese Eröffnungen gibt es etablierte Varianten und Verteidigungen, die sich über die Jahrhunderte, die Schach schon gespielt wird, als besonders erfolgreich herausgestellt haben. Erfahrene Schachspieler kennen wesentlich mehr Varianten als Anfänger und es ist nicht ungewöhnlich, dass Spieler mehr als 20 Züge lang die Eröffnung einer

3 Diese Überschlagsrechnung findet sich so auch bei Shannon (1950).

bekannten Partie nachspielen. Das hört sich langweiliger an als es ist, denn in der Vorbereitung auf ein Spiel analysieren Schachspieler frühere Partien ihrer Gegner und studieren ihre Vorlieben und Schwächen genau. Jeder Spieler versucht in der ersten Phase das Spiel auf eine Eröffnungsvariante zu lenken, die er für besonders vielversprechend hält. Vielleicht hat ein Spieler sich sogar extra für seinen Gegner eine neue Variante überlegt, die irgendwann von einer etablierten Variante abweicht.

Während der Eröffnungsphase denken menschliche Spieler daher kaum nach. Sie spielen zügig die etablierten und vorbereiteten Züge. Deep Blue hatte, genauso wie Kasparow, eine Datenbank an etablierten Eröffnungen zur Verfügung. (Deep Blues Gedächtnis war allerdings ein bisschen größer.) Diese Eröffnungen hat sich Deep Blue aber nicht selber ausgedacht. IBM hat dafür mehrere Großmeister eingestellt, die für Deep Blue diese Eröffnungen zusammengestellt haben. Das ist nicht unbedingt unfair, denn auch Kasparow hat Sekundanten, die ihn bei der Vorbereitung unterstützen, und auch er steht mit seiner Eröffnung auf den Schultern früherer Schachriesen. Aber während das Team bei IBM wusste, dass ihr Gegner Kasparow sein wird, wusste Kasparow nicht, gegen wen er eigentlich spielen wird. Deep Blue hatte bisher nicht viele öffentliche Spiele gespielt. Außerdem wurde das Programm fortwährend mit der Hilfe mehrerer Großmeister verbessert, Kasparow wusste allerdings nicht, wer diese Großmeister waren. Kein Wunder, dass Kasparow – der für seine sorgfältige Vorbereitung berühmt ist – im Vorfeld etwas nervös wirkte.⁴

Aus Sicht der KI-Forschung beginnt der interessante Teil des Schachspiels erst wirklich im Mittelspiel, sobald die Eröffnungsphase vorbei ist. Im Mittelspiel müssen die Spieler selbständig vollkommen neue Spielsituationen bewerten. Die Züge dauern dann auch wesentlich länger. Erst im Mittelspiel spielte Kasparow wirklich gegen eine auf sich gestellte KI. Einer der Großmeister, der IBM bei der Entwicklung geholfen hatte, berichtete Jahre später, dass sie auf bestimmte Eröffnungsvarianten spekuliert hatten, die extra für Kasparow vorbereitet waren. Damit Kasparow das aber nicht merkt, tat Deep Blue manchmal so, als ob er lange nachdenken muss, um den nächsten Zug zu wählen. Kasparow sollte denken, dass er schon gegen die KI spielt, obwohl er in Wirklichkeit noch gegen die von Menschen vorbereitete Eröffnungs-

4 Kasparow (2017), S. 163 und S. 167f.

strategie spielt. Ein umgekehrter Turing-Test: Der Computer stellt sich dumm, damit der Mensch denkt, er ist ein Computer. IBM hatte offenbar die harten Bandagen angelegt und scheute auch vor psychologischen Tricks nicht zurück, um Kasparow zu täuschen.⁵

Aber auch vom Mittelspiel aus ist der Suchraum beim Schach noch zu groß, als dass ein Computer das Spiel bis zum Ende durchplanen könnte. Stattdessen spielen Menschen und Computer nur ein paar der möglichen zukünftigen Züge durch. Unter diesen Zügen schätzen sie ab, welche am vielversprechendsten aussehen und fokussieren sich auf diese. Sie nutzen, so wie es im vorherigen Kapitel erklärt wurde, bei ihrer Suche nach dem vielversprechendsten Zug eine Heuristik. Diese Heuristik bewertet für jeden Zug, wie nah man damit an das Ziel herankommt, den Gegner matt zu setzen.

Heuristiken beim Schach nutzen zum Beispiel den Materialgewinn. Wenn ich einen Zug machen kann, in dem ich die Dame des Gegners schlage, dann verbessert das vermutlich meine Gewinnchancen. Sicher weiß ich das nicht, aber es ist eine recht gute Daumenregel. Es wird wärmer. Manchmal muss ich auch eine eigene Spielfigur opfern und damit erst scheinbar meine eigene Position schwächen, bevor ich eine wertvollere Figur des Gegners schlagen kann. Um zu entscheiden, ob sich ein Bauernopfer lohnt, muss ich ein paar Züge in die Zukunft schauen. Falls dadurch ein Zugzwang beim Gegner entsteht, und ich anschließend sicher eine wertvollere Figur des Gegners schlagen kann, lohnt sich der Umweg, der mich zunächst von meinem Ziel wegführt.

Um ihr Mittelspiel zu verbessern haben Menschen und Computer also im Wesentlichen zwei Möglichkeiten: Entweder braucht man mehr Rechenleistung, damit man mehr Züge und weiter in die Zukunft planen kann, oder man verbessert die Heuristik und lernt, Spielzustände besser einzuschätzen, damit man dumme Pläne gar nicht erst in Betracht ziehen muss.

Durch die fleißige Chipindustrie verdoppelt sich die Rechenleistung von Computern bisher noch knapp alle zwei Jahre. Dass Computer immer mehr Aufgaben besser erledigen können, liegt auch daran, dass sie einfach schneller werden. Das bedeutet nicht unbedingt, dass die Rechner intelligenter werden. Ein schnellerer Computer kann mehr Züge beim Schach im Voraus berechnen und mehr Spielzustände bewerten. Ein Computer, der einfach stumpf mehr Möglichkeiten durch-

⁵ Kasparow (2017), S. 185.

rechnet, spielt zwar besseres Schach, aber wir sollten dem Computer dann nicht unbedingt mehr Intelligenz zusprechen.

Die Intelligenz steckt in der Heuristik

Gute menschliche Spieler schauen üblicherweise nur drei bis vier eigene Züge – Großmeister mehr, aber nicht dramatisch mehr – in die Zukunft. Dabei analysieren sie vielleicht ein paar hundert Spielzustände. Eine lächerlich kleine Zahl, sogar im Vergleich zu einfachen, frühen Schachcomputern. Trotzdem spielen Menschen verdammt gut und erst 1997 konnte der beste menschliche Spieler von einem Computer geschlagen werden. Die Rechenleistung scheint nicht das Geheimnis der menschlichen Intelligenz zu sein. Schachspieler unterscheiden sich untereinander kaum darin, wie schnell sie Züge im Kopf durchspielen können. Und eine Großmeisterin, die unter Zeitdruck spielt – also viel weniger Züge untersucht als normalerweise – spielt nur unwesentlich schlechter als sonst.⁶

Da also bessere Spieler beim Schach gar nicht viel mehr mögliche Fortsetzungen einer Partie analysieren, muss ihre Leistung durch eine bessere Heuristik erklärt werden. Sie können einfach besser einschätzen, welche Spielzustände für sie vorteilhaft sind und welche nicht. Materialgewinn ist nicht alles. Dazu kommen Fragen wie: Ist der König gut geschützt? Werden Figuren geblockt? Steht der Bauer frei? All diese Aspekte müssen gegeneinander abgewogen werden. Während eine Anfängerin nicht genau weiß, was sie mit einer Position anfangen soll, erfasst die Expertin mit einem Blick, was gut oder schlecht läuft, und weiß, wie man darauf reagieren sollte.

Abgesehen von der Rechenleistung spielt auch bei Computern die Heuristik die entscheidende Rolle. Ein Schachcomputer, der Materialgewinn wichtiger findet als den Schutz des eigenen Königs, spielt aggressiv. Ist es andersherum, spielt der Computer eher defensiv. Durch viel Herumprobieren mit verschiedenen Heuristiken versuchen die Programmierer herauszufinden, welche Heuristik am besten funkto-

6 Siehe Charness (1981) und Gobet & Simon (1996a). Im Endspiel ist die Suchtiefe größer, außerdem hängt die Suchtiefe natürlich von der Stellung ab. Eine spätere Studie fand heraus, dass Experten tiefer suchen als zuvor angenommen (Campitelli & Gobet, 2004). Im Vergleich zu Computern sind aber auch diese Schätzungen klein.

niert. Da Kasparow als Mensch viel weniger mögliche Züge analysierte als Deep Blue und trotzdem mithalten konnte, musste seine Heuristik immer noch deutlich besser sein als die Heuristik, die die IBM-Ingenieure entwickelt haben.

Wie entwickeln erfahrene menschliche Schachspieler also eine so gute Heuristik? Eines ist klar: Profispieler wie Kasparow haben sehr, sehr viel Übung. Ohne systematisches Training geht es nicht. Durch dieses Training können sie handlungsrelevante Muster auf dem Spielbrett erkennen, die Anfängern verborgen bleiben. Wo eine Anfängerin zum Beispiel nur eine Reihe von Bauern sieht, sieht die Expertin sofort eine Festung, die sie leicht verteidigen kann. Solche Denkstrukturen erlauben es Expertinnen und Experten, die entscheidenden Informationen effizient zu verarbeiten.⁷

Dass erfahrene Spieler auf Schach bezogene Informationen extrem effizient verarbeiten können, sieht man auch daran, dass sie nicht selten blind spielen können. Oder daran, dass Schachmeister eine Spielsituation nur kurz sehen müssen, um die Spielfiguren anschließend aus dem Gedächtnis wieder aufstellen zu können. Das liegt aber nicht etwa daran, dass sie einfach ein besseres Gedächtnis haben. Werden die Spielfiguren zufällig auf das Brett gestellt, sodass sie keine sinnvolle Schachposition ergeben, dann erinnern sich Schachmeister auch nicht viel besser als andere Menschen an die Position der Figuren. Sie haben schlicht durch ihr Training eine große Zahl an relevanten Schachmustern gelernt, die sie schnell wiedererkennen können.⁸

Diese Schachmuster bilden die Grundlage für die Heuristik, mit der Schachspieler eine Position bewerten. Wenn wir wüssten, welche Schachmuster Experten sehen, und auf welche Weise sie damit abschätzen, welche Züge besser sind, dann könnten wir diese Heuristik einem Suchprogramm vorgeben – und so ein Schachprogramm entwickeln, das wie ein menschlicher Experte spielt. Doch fragt man Schachexpertinnen und -experten, wie sie das machen, sind die Antworten selten hilfreich. Zu erklären, wie man Schach spielt, ist ein bisschen so wie zu erklären, wie man Fahrrad fährt. Entscheidende Aspekte weiß man nur intuitiv und kann man nicht wirklich erklären. In der Psycho-

7 Ein gutes Buch zum Thema Expertise und zur Wichtigkeit von systematischem Training ist das von Ericsson & Pool (2016).

8 Die klassische Studie dazu wurde von Chase & Simon (1973) durchgeführt. Siehe auch Gobet & Simon (1996b).

logie nennt man das »implizites Wissen«. Nicht nur bei Schach ist die Entwicklung von KI-Systemen in der Vergangenheit daran gescheitert, solch implizites Wissen von Experten auf Computer zu übertragen. Stattdessen wurde der meiste Fortschritt bei Anwendungen von KI bis vor kurzem dadurch erzielt, dass die Computer einfach nur schneller geworden sind und dümmere Algorithmen dadurch besser funktionierten. Wirklich intelligenter werden sie dadurch nicht. Ich habe es jetzt schon mehrfach gesagt und ich kann nicht versprechen, dass ich es nicht nochmal sagen werde: Wir sind es, die Computern mehr Intelligenz zuschreiben, als unter der Haube vorhanden ist.

Kasparow verliert nicht gegen Deep Blue

Viele KI-Forscherinnen und -Forscher fanden damals den Wettkampf zwischen Garri Kasparow und Deep Blue ungefähr so interessant wie ein Tauziehen zwischen Arnold Schwarzenegger und einem Traktor. Für viele Schachfans hingegen spielte Deep Blue manchmal überraschende Züge, die ihn sehr menschlich aussehen ließen. Solche Züge trauten sie Computern, die das Spiel ja nicht genauso wie Menschen verstehen, einfach nicht zu.

Auch Kasparow war verwundert darüber, dass Deep Blue manchmal dumm wie Brot war und ihn im nächsten Moment mit brillanten Zügen in große Schwierigkeiten brachte. Während der sechs Spiele, in denen er gegen Deep Blue antrat, blieb Deep Blue für Kasparow unberechenbar. Während der Computer trotzdem einfach stoisch weiter rechnete, wurde Kasparow im Verlauf des Turniers zunehmend gereizt. Kasparows großes Ego war sichtbar angeknackst. Ein Kommentator scherzte damals: »Sie denken vielleicht, das ist keine große Sache, aber warten Sie, bis das Ding hinter *Ihrem* Job her ist!«⁹ Als IBM sich weigerte die internen Protokolle von Deep Blue schon während des Turniers zur Analyse zur Verfügung zu stellen, witterte Kasparow Betrug. Bei einer Pressekonferenz wurde er konkret gefragt, ob er glaubt, dass bei Deep Blues entscheidenden Zügen Menschen interveniert hätten. Er antwortete, dass er sich an Maradonnas irreguläres Tor gegen England

9 Kasparow (2017), S. 134.

bei der WM 1986 erinnert fühlte und ergänzte: »[Maradonna] sagte, es war die Hand Gottes!«¹⁰

In seinem Buch, das Kasparow 30 Jahre später über den historischen Wettkampf veröffentlichte, entschuldigt er sich dafür bei den Ingenieuren, die Großes geleistet hatten. Gleichzeitig kann er den Groll, den er immer noch gegen IBM hegt, nicht verbergen, und er wirft IBM durchweg unsportliches Verhalten vor.¹¹ So verständlich Kasparows persönliche Kränkung vielleicht ist, für den Rest von uns war IBMs symbolischer Sieg keineswegs die finale Niederlage der menschlichen gegenüber der künstlichen Intelligenz. Hinter IBM stand eine Heerschar an Großmeistern und Ingenieuren, die all ihr Wissen über KI und Schach in die Entwicklung von Deep Blue gesteckt hatten. Gegen sie hatte Kasparow eigentlich verloren.

¹⁰ Kasparow (2017), S. 195f.

¹¹ Kasparow (2017), S. 214-220; Levy (2017).