

Künstliche Intelligenz - Newell und Simon 1976

cagent
cagent@cognitiveagent.org

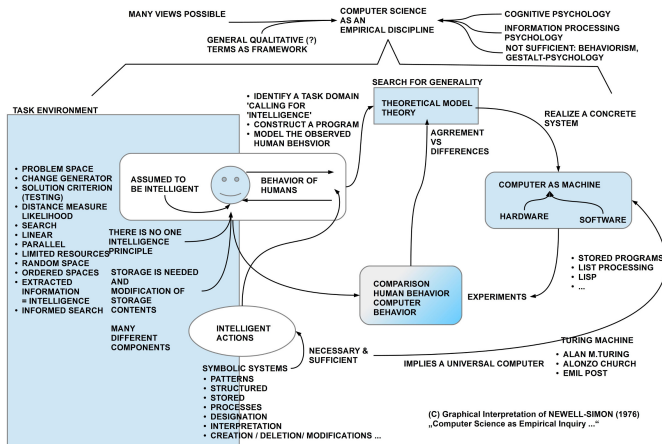


Fig. 1. Newell und Simon sehen die Informatik als eine empirische Disziplin

Abstract—Im Jahr 2017 nimmt die Erwähnung von sogenannter *Künstlicher Intelligenz* außerhalb der Wissenschaften, im öffentlichen Bereich, fast inflationär zu. Zugleich muss man feststellen, dass Erklärungen des Begriffs 'Künstliche Intelligenz' wie auch anderer Begriffe in seinem Umfeld Mangelware sind. Es wird daher ab jetzt mehr Blogbeiträge geben, die auf diese Thematik gezielter eingehen werden. Hier ein erster Beitrag mit Erinnerung an einen wichtigen Artikel von Newell and Simon 1976.

I. INFORMATIK ALS EMPIRISCHE WISSENSCHAFT

Im Jahr 1975 empfingen Allen Newell¹ und Herbert A. Simon² den angesehenen *ACM Turing Preis* von der ACM aufgrund ihrer vielen wichtigen Beiträge zur Künstlichen Intelligenzforschung in den vorausgehenden Jahren. Die Preisrede beider Autoren wurde in den *Communications of the ACM* 1976 abgedruckt (siehe: [NS76]).

In dieser Rede wagen die Autoren eine generelle Sicht auf die Informatik ('computer science'), die Akzente erkennen lässt, die von heutigen Auffassungen von Informatik – zumindest wie sie in Deutschland üblich sind – doch deutlich abweicht.

Für die beiden Autoren ist die Informatik eine *empirische Wissenschaft*, deren Interesse darin besteht, das Verhalten von Menschen, dort, wo es *Intelligenz* erkennen lässt, in ein *theoretisches Modell*

zu übersetzen, das sich dann als eine *konkrete Maschine* (ein Computer, ein Roboter) physikalisch realisieren lässt. Man kann dann diese konkrete Maschine dazu benutzen, *Tests* durchzuführen, durch die man *überprüfen* kann, ob sich die gebaute Maschine hinreichend *ähnlich* wie ein Mensch verhält oder aber davon deutlich *abweicht*. Liegen Abweichungen vor, dann muss man den Sachverhalt weiter ergründen und versuchen, ob man das theoretische Modell verbessern kann.

Für sie erscheint der *Mensch* als eine Art *Standardmodell* für *Intelligenz*, allerdings nicht so, dass man den Begriff 'Intelligenz' mit einer einzigen Verhaltensweise oder mit einem einzigen Prinzip identifizieren könnte. Das vielfältige menschliche Verhalten verweist nach den Autoren vielmehr auf eine *Vielzahl von Komponenten*, deren Zusammenwirken zu einem als 'intelligent' wirkenden Verhalten führt. Für das Erkennen einer möglichen 'Intelligenz' ist es ferner wichtig, dass man den ganzen *Kontext* berücksichtigt, in dem *spezifische Aufgaben* vorliegen, die *gelöst* werden sollten.

Durch ihre Forschungsarbeiten zu *symbolischen Systemen* und zur *heuristischen Suche* haben Newell und Simon herausgefunden, dass die Klärung eines *Problemraumes* nur dann *besser als zufällig* sein kann, wenn der Problemraum minimale Regelmäßigkeiten, eine minimale Ordnung aufweist, die – sofern sie erkannt wurde – dann in Form *spezifeller Informationen* *angesammelt* werden kann und dann, nach Bedarf, bei der Klärung des Problemraumes *genutzt* werden kann. Und es ist genau diese *spezifische angesammelte Information* die die Autoren mit *Intelligenz* gleichsetzen! Ein Mensch kann nur dann gezielter, schneller eine Aufgabe lösen, wenn er über spezielle Informationen (Wissen) verfügt, die ihn in die Lage versetzen, genau jene Verhaltensweisen zu zeigen, die schnell und effizient zum Ziel führen.

Überträgt man dies auf die theoretischen Modelle der Informatik, dann muss man Wege finden, *spezifisches Bereichswissen* (engl.: 'domain knowledge') für ein intelligentes Verhalten in geeignete Datenstrukturen und Prozesse zu übersetzen. Auf die vielen Beispiele und Details zu diesen Überlegungen wird hier verzichtet³.

¹https://en.wikipedia.org/wiki/Allen_Newell

²https://en.wikipedia.org/wiki/Herbert_A._Simon

³diese kann jeder in dem Artikel nachlesen

II. DISKURS

Hier einige Überlegungen im Anschluss an den Artikel von Newell und Simon.

A. Intelligenz

Zunächst ist interessant, dass die Verwendung des Begriffs 'Intelligenz' gebunden wird an das Verhalten von Menschen, wodurch der Mensch als undiskutierter Maßstab für mögliche Intelligenz gesetzt wird.

Daraus folgt nicht notwendigerweise, dass es jenseits des Menschen keine andere Formen von Intelligenz gibt, sondern nur, dass man *den Typ* von Intelligenz, der beim Menschen vorliegt und sichtbar wird, vorläufig als *Standard* benutzen möchte. Also eine Intelligenz mit Index: *Intelligenz_{human}*.

Das macht auch verständlich, dass man als wichtige empirische Wissenschaft in Begleitung der Informatik die *kognitive Psychologie* sieht, die sich u.a. auch mit der sogenannten 'Informationsverarbeitung im Menschen' beschäftigt.

Es wundert dann allerdings, dass die Autoren den im Rahmen der Psychologie eingeführten Begriff des *Intelligenz-Quotienten (IQ)* samt den dazugehörigen erprobten Messverfahren nicht benutzen, nicht einmal erwähnen. Dies würde die Möglichkeit eröffnen, die Verhaltensleistung von technischen Systemen entsprechend zu messen und direkt mit Menschen zu vergleichen. Der oft zitierte Turing-Test (nicht von den beiden Autoren) ist verglichen mit den Testbatterien des IQ-Quotienten mehr als dürftig und nahezu unbrauchbar.

Durch den Verzicht auf die sehr detailliert ausgearbeiteten Testbatterien der Psychologie bleibt die Charakterisierung des Begriffs 'Intelligenz' in der Informatik weitgehend vage, fast beliebig.

An dieser Stelle könnte man einwenden, dass in der Informatik andere Aufgabenstellungen untersucht werden als in der Psychologie üblich bzw. andere Aufgabenstellung, die Menschen in dieser Weise nicht angehen, dadurch wir die Verwendung des Begriffs 'Intelligenz' aber noch undurchsichtiger, geradezu ominös.

Ogleich Newell und Simon betonen, dass sie die Informatik als eine empirische Theorie sehen, bleibt der ganze Komplex des tatsächlichen *objektiven Messens* etwas vage. Zum objektiven Messen gehören zuvor *vereinbarte Standards*, die beim Messen als Referenzen benutzt werden, um ein *Zielobjekt* zu 'vermessen'. Wenn das zu messende Zielobjekt ein *Verhalten* sein soll (nämlich das Verhalten von Maschinen), dann muss zuvor sehr klar definiert sein, was denn das *Referenz-Verhalten* von Menschen ist, das in einem

(welchen?) Sinn als 'intelligent' angesehen wird und das dazu benutzt wird, um das Maschinenverhalten damit zu vergleichen. Es ist weder bei Newell und Simon klar zu sehen, wie sie ihr Referenzverhalten von Menschen zuvor klar definiert haben, noch sind die Messprozeduren klar.

Der grundsätzliche Ansatz von Newell und Simon mit der Informatik als empirischer Disziplin (zumindest für den Bereich 'Intelligenz') erscheint auch heute noch interessant. Allerdings ist das begriffliche Chaos im Kontext der Verwendung des Begriffs 'Intelligenz' heute zu einem notorischen Dauerzustand geworden, der es in der Regel unmöglich macht, den Begriff von 'künstlicher Intelligenz' in einem wissenschaftlichen Sinne zu benutzen. Jeder benutzt ihn heute gerade mal, wie es ihm passt, ohne dass man sich noch die Mühe macht, diese Verwendung irgendwie transparent zu machen.

B. Lernen

Während Newell und Simon im Fall des Begriffs 'Intelligenz' zumindest ansatzweise versuchen, zu erklären, was sie damit meinen, steht es um den Begriff 'Lernen' ganz schlecht.

Explizit kommt der Begriff 'Lernen' bei den Autoren nicht vor, nur indirekt. Dort wo heuristische Suchprozesse beschrieben werden, die mit Hilfe von symbolischen Systemen geleistet werden, stellen sie fest, dass man aufgrund ihrer empirischen Experimente wohl (in dem theoretischen Modell) annehmen muss, dass man Informationen speichern und verändern können muss, um zu jenem Wissen kommen zu können, das dann ein optimiertes = intelligentes Verhalten ermöglicht.

Aus psychologischen Lerntheorien wissen wir, dass 'Intelligenz' und 'Lernen' zwei unterschiedliche Eigenschaften eines Systems sind. Ein System kann wenig intelligent sein und doch lernfähig, und es kann sehr intelligent sein und doch nicht lernfähig.

Nimmt man die Charakterisierung von Newell und Simon für 'Intelligenz' dann handelt es sich um ein 'spezielles Wissen' zum Aufgabenraum, der das System in die Lage versetzt, durch ein 'gezieltes Verhalten' schneller ans Ziel zu kommen als durch rein zufälliges Verhalten. Eine solche Intelligenz kann einem System zur Verfügung stehen, auch ohne Lernen, z.B. (i) bei biologischen Systemen als eine genetisch vererbte Verhaltensstruktur; (ii) bei technischen Systemen durch eine volle Konfiguration durch Ingenieure. Bei biologischen Systemen tritt allerdings 'Intelligenz' nie isoliert auf sondern immer in Nachbarschaft zu einer Lernfähigkeit, weil die dynamische Umwelt biologischer Systeme beständig neue Anpassungen verlangt, die nicht alle vorher gesehen werden können. Im Fall technischer Systeme mit begrenzter Lebensdauer und

definiertem Einsatz war dies (und ist dies) weitgehend möglich.

Wenn man von 'Künstlicher Intelligenz' spricht sollte man daher die beiden Strukturen 'Intelligenz' und 'Lernen' sehr klar auseinander halten. Die Fähigkeit, etwas zu lernen, erfordert völlig andere Eigenschaften als die Struktur eines Wissens, durch das ein System sich 'intelligent' statt 'zufällig' verhalten kann.

C. Theorie

Die Forderung von Newell und Simon, die Informatik als eine 'empirische Wissenschaft' zu betrachten, die richtige theoretische Modelle (= Theorien) konstruiert und diese über realisierte Modelle dann empirisch überprüft, hat im Rahmen des allgemeinen Systems Engineerings auch heute noch einen möglichen Prozess-Rahmen, der alle diese Forderungen einlösen kann. Allerdings hat es sich in der Informatik eingebürgert, dass die Informatik einen Sonderweg kreiert hat, der unter der Überschrift *Softwareengineering* zwar Teilaspekte des generellen Systemsengineerings abbildet, aber eben nur Teilaspekte; außerdem ist durch die Beschränkung auf die Software ohne die Hardware ein wesentlicher Aspekt des Gesamtkonzepts Computer ausgeklammert. Ferner ist der harte Aspekt einer vollen empirischen Theorie durch die Phasenbildungen 'logisches Design' nur unvollständig abgebildet. Designmodelle sind kein Ersatz für eine richtige Theorie. Für das sogenannte 'modellgetriebene Entwickeln' gilt das Gleiche.

D. Mensch als Maßstab?

War es noch für Newell und Simon offensichtlich klar, dass man für den Begriff 'Intelligenz' den Menschen als Referenzmodell benutzt, so ist dies in der heutigen Informatik weitgehend abhanden gekommen. Dies hat einmal damit zu tun, dass der Wissenschaftsbegriff der Informatik samt der meisten Methoden bislang nicht in den allgemeinen, üblichen Wissenschaftsbegriff integriert ist, zum anderen dadurch, dass die Aufgaben, die die intelligenten Maschinen lösen sollen, aus allen möglichen ad-hoc Situationen ausgewählt werden, diese keinen systematischen Zusammenhang bilden, und man in den meisten Fällen gar nicht weiß, wie man den Bezug zum Menschen herstellen könnte. Dadurch entsteht der vage Eindruck, dass die 'Intelligenz' der künstlichen Intelligenzforschung irgendwie etwas anderes ist als die menschliche Intelligenz, von der menschlichen Intelligenz möglicherweise sogar ganz unabhängig ist. Man macht sich allerdings nicht die Mühe, systematisch und zusammenhängend die Verwendung des Begriffs der 'Intelligenz' in der Informatik zu klären. Aktuell hat man den Eindruck, dass jeder gerade mal das behauptet, was ihm gerade gefällt. Auch eine Art von *Fake News*. Die Marketingabteilungen der großen Konzerne freut es, weil sie nach Belieben alles Versprechen können, was

sie wollen, ohne dass irgendjemand sinnvoll nachprüfen kann, was das genau ist, ob das überhaupt geht.

Doch sollte man sich durch diese terminologische Unklarheiten, die auf eine fehlende wissenschaftliche Methodik der Informatik zurück zu führen sind, nicht davon ablenken lassen, zu konstatieren, dass trotz chaotischer Begrifflichkeit im Konkreten und Im Detail seit Newell und Simon 1976 extreme Fortschritte gemacht wurden in speziellen Algorithmen und spezieller Hardware, mit der man heute viel mehr realisieren kann, als sich Newell und Simon damals hätten träumen lassen. Die Vorteile, die man sich durch diese Fortschritte im Konkreten erarbeitet hat, werden aber weitgehend verspielt durch große Theoriedefizite, die zu gerade kuriosen Anschauungen und kuriosen Forschungsprojekten führen. In der KI ist dies aber nichts Neues: es gab sehr unterschiedliche Phasen; alle Fehler führten irgendwann dann doch zu verbesserten Einsichten.

REFERENCES

- [NS76] Allen Newell and Herbert A. Simon. Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, 19(3):113–126, 1976.