

# DIE WISSENSCHAFT DES UNBERECHENBAREN

Ist man schon Mozart, wenn man eine gute Sinfonie wertschätzen kann? Oder gibt es einen fundamentalen Unterschied zwischen einer großen kreativer Leistung und der Anerkennung solcher Leistung?

Die Antwort scheint offensichtlich, doch wenn man beweisen will, dass Mozart seinem Publikum etwas voraus hatte, dann muss man die Frage erst präzise genug machen. Überraschenderweise ist ein vielversprechender Ansatz in der theoretischen Informatik zu finden: Die Mozart-Frage ist eine Analogie für das sogenannte „P versus NP“ Problem, für dessen Lösung ein Preisgeld von einer Million Dollar ausgeschrieben ist.

Die Theorie von Berechenbarkeit, Information und Komplexität stellt fundamentale Fragen, und auch wenn die endgültigen Antworten noch nicht feststehen, haben wir schon einiges gelernt. Dieses Wissen ist nicht nur für Ingenieure relevant, sondern findet auch Anwendungen in der reinen Mathematik, der Philosophie, Ökonomie, Physik, Linguistik und Biologie. Das Phänomen Berechnung scheinen also nicht nur eine Eigenheiten der elektronischen Rechenmaschinen zu sein, sondern tiefere Bedeutung zu haben.

Sehr wichtig ist dabei auch, was *nicht* möglich ist. Trotz allen Fortschritten in der Automatisierung von Industrie und Gesellschaft gibt es fundamentale Schranken, die kein Computer jemals überwinden können wird. Die Grenzen des Berechenbaren zu kartieren, ist Kern der theoretischen Informatik und liefert Munition für mannigfaltige Interpretationen.

Dieses Autonome Tutorium zeigt einige Resultate dieser Forschung auf und diskutiert ihre Bedeutung. Es richtet sich ausdrücklich an Studierende aus allen Fachbereichen. Kein Vorwissen ist erforderlich.

## Zur Literatur

Die Texte sind fast ausschließlich englischsprachig. Aus Platzgründen werden hier nur Teile der geplanten Literatur aufgeführt. Die Lektüre, die vorausgesetzt wird, ist bewusst für ihre (relativ) leichte Zugänglichkeit ausgewählt. PDF-Versionen aller hier zitierten Werke sind mit Google leicht zu finden.

## Themenkatalog

Da kein Vorwissen vorausgesetzt wird, wird am Anfang des Tutoriums die *Turing-Maschine* eingeführt. Exemplarisch werden einige nicht berechenbare Probleme gezeigt, wie z.B. das klassische *Halteproblem* und der *Satz von Rice*.

Gibt es also Probleme, deren Lösung zwar (anders als bei klassischen Paradoxen) definitiv existiert, aber im Allgemeinen nicht gefunden werden kann? Das setzt voraus, dass Turing-Maschinen auch alle möglichen Algorithmen und Maschinen abdecken. Dagegen gibt es einige Einwände [CS99], die aber ihrerseits angezweifelt werden [Dav04]. Das

dient als Sprungbrett für eine Diskussion, wie relevant die (unendlichen) mathematischen Objekte für die (endliche) physikalische Realität sind.

Die gleiche Frage stellt sich bei der *Komplexitätstheorie*, die untersucht, wie viele Ressourcen für die Lösung eines Problem nötig sind. Auch wenn es grundsätzlich algorithmisch gelöst werden kann, ist die Berechnung oft nicht praktikabel, weil sie z.B. zu viel Zeit oder Speicher benötigt. Das ist eine Eigenschaft des Problems, nicht eines konkreten Computers oder Algorithmus. Deshalb hat die Komplexitätstheorie auch viele Verbindungen zu anderen Wissenschaften. Wir untersuchen, wie sie auf philosophische Fragestellungen [Aar13] angewandt werden kann.

*Quantencomputer* und andere hypothetische Rechnerarchitekturen ändern nichts an Berechenbarkeit, beeinflussen aber die Komplexität von Problemen. Die genauen Auswirkungen verschiedener Theorien, insbesondere der Quantenmechanik, sind Gegenstand aktueller Forschung. Wenn diese Modelle physikalisch umsetzbar sind, fordert das die Idee heraus, dass klassische Turing-Maschinen ein gutes Modell für real umsetzbare Berechnungen seien. Der Stand der Wissenschaft lässt aber auch eine andere Hypothese zu [Aar05]: NP-vollständige Probleme können nicht effizient gelöst werden und physikalische Theorien, die das erlauben würden, sind folglich falsch.

Die formale Logik ist die Mutter der theoretischen Informatik und hat zahllose Verbindungen mit den anderen Themen des Tutoriums. Wir holen wieder Vorwissen auf: Die Regel formaler Logik, ihre Beziehung zum „normalen“ informellen logischen Denken, Beweiskalküle [Wad15] und die Gödelschen *Unvollständigkeitssätze*. So gewinnen wir neue Perspektiven auf vorherigen Themen und wappnen uns gegen fragwürdige Interpretationen [Raa05] dieser Sätze.

Mit diesem Wissen und dem der vorangegangenen Themen widerlegen wir nun verschiedene Argumente gegen die prinzipielle Möglichkeit künstlicher Intelligenz [Tur50; Aar13]. Außerdem diskutieren wir die (Schreckens-)Vision einer „technologischen Singularität“ oder „Superintelligenz“. Obwohl diese Positionen kritisch zu sehen sind, ist Vorsicht in der KI-Forschung [Fut15] durchaus angebracht.

## Literatur

- [Tur50] A. M. Turing. „Computing Machinery and Intelligence“. In: *Mind* LIX.236 (1950), S. 433–460.
- [CS99] B. Jack Copeland und Richard Sylvan. „Beyond the Universal Turing Machine“. In: *Australasian Journal of Philosophy* 77.1 (1999), S. 46–66.
- [Dav04] Martin Davis. „The Myth of Hypercomputation“. In: *Alan Turing: Life and Legacy of a Great Thinker*. 2004, S. 195–211.
- [Aar05] Scott Aaronson. „NP-complete Problems and Physical Reality“. In: *SIGACT News* 36.1 (März 2005), S. 30–52.
- [Raa05] Panu Raatikainen. „On the Philosophical Relevance of Gödel’s Incompleteness Theorems“. In: *Revue Internationale de Philosophie* 59.4 (2005), S. 513–534.

- [Aar13] Scott Aaronson. „Why Philosophers Should Care About Computational Complexity“. In: *Computability : Turing, Gödel, Church, and beyond*. Hrsg. von B. Jack Copeland. MIT Press, 2013, S. 261–328.
- [Fut15] Future of Life Institute. *Research Priorities for Robust and Beneficial Artificial Intelligence: an Open Letter*. 2015. URL: [futureoflife.org/AI/open\\_letter](http://futureoflife.org/AI/open_letter) (besucht am 23.08.2015).
- [Wad15] Philip Wadler. „Propositions as Types“. In: *Communications of the ACM* (2015).

## MEINE ROLLE ALS TUTOR VON „DIE WISSENSCHAFT DES UNBERECHENBAREN“

Das Kernanliegen des Tutoriums ist, die Denkweise und Konzepte der theoretischen Informatik an Menschen aus möglichst vielen verschiedenen Fachbereichen zu vermitteln. Das funktioniert nur, wenn alle Teilnehmer\*innen ein Verständnis für diese Ideen entwickeln, indem sie intensiv darüber nachdenken und alle Aspekte betrachten. Mein Ziel als Tutor ist also, diesen Denkprozess anzustoßen und zu unterstützen, statt nur eine lange Liste von Dogmen zu präsentieren. Offene Gruppendiskussionen sind dafür das beste Medium. Als Moderator werde ich auf einen produktiven und respektvollen Diskurs achten und auf eine ausgeglichene Verteilung der Redezeit Wert legen.

Die Themen dieser Diskussionen werden teilweise durch die Literatur vorgegeben. Davon ab ist mein Ziel aber auch, zu eigenen Thesen anzuregen: Aus anderem Hintergrundwissen resultieren hoffentlich Perspektiven, dir mir nie in den Sinn gekommen wären! Die kritische Auseinandersetzung mit all diesen Aussagen (die der Literatur, meine, und die der Anderen) ist mir dabei das höchste Gut. Ich übernehme also nicht nur die Moderation, sondern werde mich auch als „Anwalt des Teufels“ betätigen.

Gleichzeitig ist eine gewisse Menge an grauer Theorie nötig, damit es überhaupt etwas zu interpretieren gibt. Da das erklärte Ziel ist, dass man ohne jedes Vorwissen teilnehmen kann, müssen diese Grundlagen im Laufe des Tutoriums geschaffen werden. Somit muss ich zeitweise, notgedrungen, die Rolle eines Dozenten einnehmen.

Ich habe einige Ideen, um diese Segmente interaktiv und spielerisch zu machen. Trotzdem möchte ich davon möglichst wenig Gebrauch machen. Deshalb peile ich folgende Kompromiss an: Die argumentative und kritische Lektüre setzt gewisse Konzepte und Definitionen voraus, ohne die man nicht einmal *die These verstehen* kann. Einzig und allein dieses Minimum erarbeite ich mit den Teilnehmer\*innen, damit sie dem Text zumindest grundsätzlich folgen können — auch wenn viele Einzelheiten vielleicht noch geklärt werden könnten.

In den Diskussionen stehe ich dann immer bereit, um Fragen zur Materie zu beantworten und Missverständnisse zu korrigieren. Darin habe ich als Übungsgruppenleiter einige Erfahrung. Lieber ist mir aber, wenn die Teilnehmer\*innen das untereinander klären können.

Last but not least sollte zu jedem Zeitpunkt eine Atmosphäre des freien Ideenaustauschs herrschen. Mehr noch, die Treffen des Tutoriums sollten allgemein eine positive Umgebung sein. Ich will dafür mit gutem Beispiel voran gehen und wenn nötig unsozialem, diskriminierendem und anti-intellektuellem Verhalten entgegentreten.

Zusammenfassend sehe ich meine Rolle darin, ein Forum für den interdisziplinären Austausch rund um Interpretationen und Anwendungen der theoretischen Informatik zu schaffen.