

Foto: Ute Voigt

## Molekulare Computer

### Rechnen im Reagenzglas

*In konventionellen Computern arbeitet ein Prozessor Rechenschritt nach Rechenschritt ab. Die Gentechnik ermöglicht dagegen, gleichzeitig viele DNA-Moleküle in einem Reagenzglas miteinander reagieren zu lassen. Da in DNA-Molekülen Informationen kodiert sind, liegt es für den Informatiker nahe, das Reagenzglas als Computer aufzufassen mit zahlreichen parallel arbeitenden Prozessoren. Die gezielte Manipulation informationstragender Makromoleküle könnte Lösungsansätze für Probleme liefern, von denen heute noch selbst die schnellsten Supercomputer in die Knie gezwungen werden. Und noch sind längst nicht alle grundlegenden Möglichkeiten des Rechnens mit DNA bekannt.*

**M**oderne Computer widersetzen sich bisher erfolgreich der Lösung vieler Probleme des täglichen Lebens. So bräuchte ein heute verfügbarer Supercomputer für die Berechnung des kürzesten Weges einer beliebigen Rundreise durch 64 Städte bereits einige Jahrzehnte. Käme jemand auf die Idee, seine Reise auf 9 weitere Städte auszudehnen, müßte er bis zu  $2^{28}$  Stunden auf seine Routenplanung warten, was in etwa der Zeitspanne bis zur nächsten Eiszeit entspricht.

Charakteristisch für derartige Probleme ist das exponentielle Wachstum ihrer Laufzeit, der sogenannten Zeitkomplexität. Vergrößert man die Eingabe um 1, wird die Laufzeit mit der Basis  $k$  der Zeitkomplexität multipliziert. Im einfachsten Fall  $k = 2$  bedeutet dies immer noch eine Verdoppelung. Vor diesem Hintergrund brächte selbst eine Beschleunigung der Prozessorgeschwindigkeit um den Faktor 10.000 praktisch keine wesentliche Verbesserung.

Auf der Suche nach neuen Wegen in der Informationsverarbeitung erkannte man schon früh, an welcher Stelle die Natur überlegen ist: Obwohl die „Schaltzeit“ von Neuronen im Bereich von etwa einer Millisekunde derjenigen von Mikroprozessoren hoffnungslos unterlegen ist, kann zum Beispiel ein potientielles Beutetier einen Angreifer blitzartig erkennen und entsprechend reagieren – eine Fähigkeit, die das Vermögen von Mikroprozessoren bei weitem übersteigt. Möglich wird sie durch Parallelverarbeitung.

Heutzutage sind Rechnerarchitekturen mit einigen zehntausend parallel arbeitenden Verarbeitungselementen möglich, die im besten Fall Geschwindigkeitssteigerungen proportional zur Anzahl der Prozessoren zulassen. Zur Lösung vieler hartnäckiger Probleme reicht dies allerdings noch immer nicht aus. Kann es also für ein imaginäres Compu-



ter-Beutetier keine Rettung vor seinem Angreifer geben?

### Molekulare Computer

In den dreißiger Jahren hat Alan Turing auf der Suche nach der Antwort darauf, was Automaten grundsätzlich zu leisten vermögen, eine mittlerweile nach ihm benannte (hypothetische) Maschine erfunden, die all das berechnen kann, was intuitiv als berechenbar gilt. Insbesondere vermag sie auch jede Rechnung irgendeines Supercomputers auszuführen. Eine Turingmaschine besteht aus einem potentiell unendlichen Band, das in Felder aufgeteilt ist. Jedes Feld kann ein Zeichen aus einer endlichen Menge als Inschrift tragen. Der Schreib-Lesekopf der Maschine befindet sich auf genau einem Feld des Bandes, liest den Inhalt und kann ihn gemäß einer endlichen Anzahl festgelegter Regeln überschreiben, unverändert lassen oder sich ein Feld nach rechts oder links bewegen.

Jede mögliche Berechnung kann also durch das Manipulieren von Zeichenreihen (die Felder des Bandes)

ausgeführt werden. Es läßt sich sogar zeigen, daß hierfür nur zwei verschiedene Zeichen ausreichend sind. Hier liegt nun der Ansatzpunkt für die Entwicklung massiv paralleler Computer auf molekularer Ebene.

Die Entwicklung des Lebens ist verknüpft mit der Entwicklung von DNA (Desoxyribonukleinsäure). Eine lebende Zelle verarbeitet in DNA kodierte Information, wobei die Evolution in Milliarden von Jahren ihr „Handwerkszeug“ immer weiter verfeinert hat: Es entwickelte sich eine Vielfalt von Enzymen, Mechanismen zur Vervielfältigung etc. Die stürmische Entwicklung der Forschung auf diesem Gebiet hat dazu beigetragen, viele dieser Mechanismen zu verstehen und darüber hinaus auch im Reagenzglas erfolgreich anzuwenden. Die Idee des Rechnens mit DNA liegt darin, die durch Evolution entstandenen Werkzeuge zu benutzen, um die zuvor in DNA kodierten Probleme durch Manipulation der Moleküle – übertragen auf herkömmliche Computer bedeutet dies Manipulation von Zeichenrei-

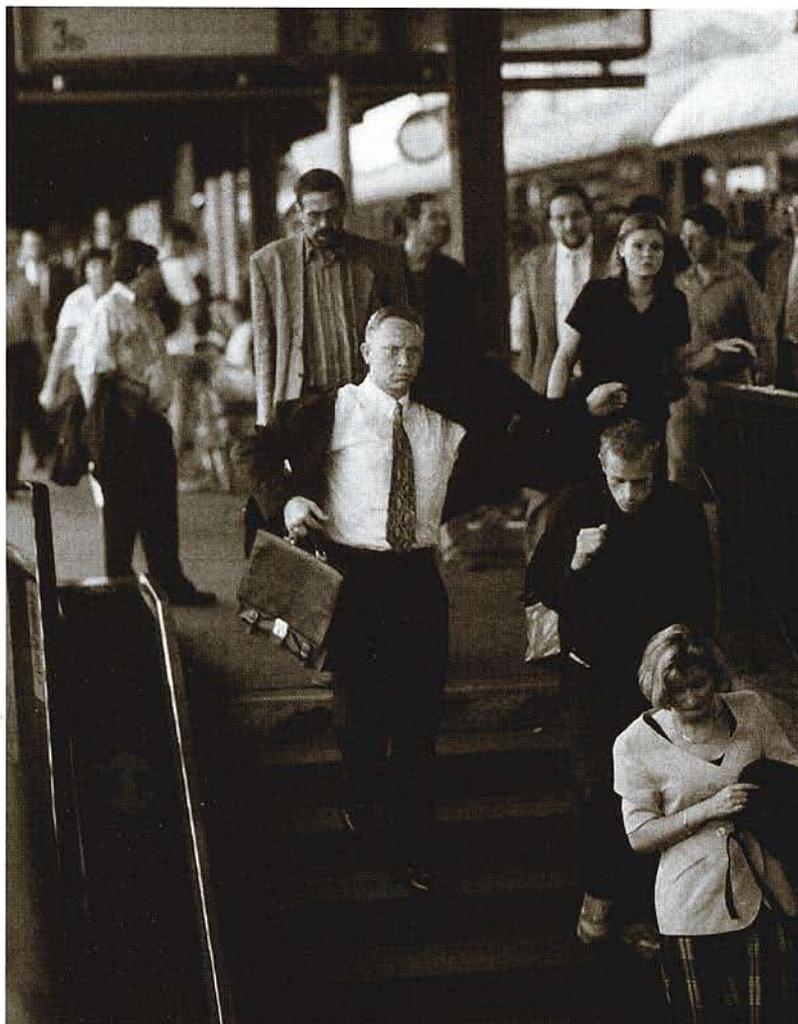
hen – zu bearbeiten. Aber worin besteht nun der Vorteil? In den in Labors verwendeten Lösungen kommen bis zu  $10^{21}$  DNA-Moleküle vor. Jedes einzelne von ihnen kann potentiell ungefähr 400 Bit kodieren. Dies übertrifft zum Beispiel die Speicherkapazität einer Ein-Gigabyte-Festplatte hunderttausendmilliardenfach. In gewissem Sinn ist jedes Molekül ein einzelner Prozessor in einem gigantischen Parallelrechner, wodurch sich auf den ersten Blick ungeahnte Möglichkeiten eröffnen – und die Überlebenschancen für das Computertier steigen.

### Kodierung in DNA

Das Prinzip des molekularen Computers beruht auf der Manipulation von informationstragenden Makromolekülen, etwa der DNA. Wie aber lassen sich die zu lösenden Aufgaben in DNA kodieren?

Ein DNA-Molekül ist ein zu einer Doppelhelix gewundener Doppelstrang, dessen Einzelstränge die unverzweigte Verkettung einfacher Bausteine (Nukleotide) sind. An-

**Alan Turing**, Jahrgang 1912, untersuchte in seiner 1937 erschienen Arbeit „On Computable Numbers“ das Prinzip eines abstrakten Universalrechners schon bevor der erste Computer realisiert war. Im Zweiten Weltkrieg hatte er den Hauptanteil am Knacken der deutschen Chiffriermaschine „Enigma“. Der Weltkrieg in Europa wurde dadurch so sehr abgekürzt, daß Deutschland der Einsatz der Atombombe erspart blieb. 1952 wurde er wegen seiner Homosexualität vor Gericht gestellt und zu einer Hormontherapie gezwungen. Er beging zwei Jahre später Selbstmord.



Heute Gießen, morgen Oberammergau: Beim Rundreiseproblem muß ein Vertreter die kürzeste Verbindung zwischen den Orten finden, die er besuchen will. Das Problem zeichnet sich dadurch aus, daß schon bei einer geringen Zahl von Orten selbst die besten Computer damit überfordert sind, die optimale Lösung zu finden.

Foto: Ute Voigt

hand von in ihnen enthaltenen Basen lassen sich vier verschiedene Nukleotide unterscheiden, die mit den Buchstaben A, C, G, und T bezeichnet werden. Jeweils 10 von ihnen kommen auf eine Windung der Doppelhelix. Jeder Einzelstrang hat zwei verschiedene Enden (das sogenannte 5'- und 3'-Ende), wodurch sich eine Richtung ergibt. Die Doppelhelix besteht aus zwei gegenüberliegenden Einzelsträngen.

Die Paarungen jeweils gegenüberliegender Nukleotide sind nicht beliebig, es fügen sich vielmehr stets G und C sowie A und T zusammen. Man spricht auch von Basenpaaren.

Wegen der Gegenläufigkeit kann etwa der Einzelstrang 5'-TAGCC-3' mit 3'-ATCGG-5' einen Doppelstrang bilden, mit 5'-ATCGG-3' aber nicht.

DNA als Träger der Erbinformation speichert in Form der Basensequenz den gesamten Bauplan eines Organismus. Da die verschiedenen Nukleotide (jeweils eines Einzelstranges) in beliebiger Reihenfolge verkettbar sind, können so auch die Aufgaben des molekularen Computers mit dem vierelementigen Alphabet in molekularer Schrift geschrieben und als DNA-Molekül gespeichert werden.

Die Komplexität im Zusammenhang mit natürlicher DNA sei durch folgende Zahlen verdeutlicht: Die nur etwa zwei Nanometer (zwei millionstel Millimeter) dicke DNA be-

steht beim Darmbakterium *Escherichia coli* aus über vier Millionen Basenpaaren und mißt gestreckt 1,3 Millimeter; die drei Milliarden Basenpaare der 23 Chromosomen in menschlichen Ei- und Samenzellen etwa einen Meter.

### Operationen des molekularen Computers

Welche Möglichkeiten kennt man, DNA-Moleküle zu manipulieren, also die gespeicherte Information zu verändern, also zu rechnen?

**Synthese.** Zur Anfertigung von Strängen speziellen Inhalts existieren verschiedene Synthesemethoden, die insbesondere auch im Bereich der kombinatorischen Chemie ihre Anwendung finden.

**Kopieren.** Ein DNA-Doppelstrang kann durch Erwärmen in zwei Einzelstränge aufgetrennt werden. Bei etwa 96°C lösen sich die verbindenden Wasserstoffbrücken, ohne die Verkettung der Nukleotide zu zerstören. Nach Temperaturabsenkung kommt es wieder zur Ausbildung von Basenpaarungen.

Einzelstränge oder Teilbereiche von ihnen können mit Hilfe des Polymerase-Kettenreaktion genannten Verfahrens wieder zu einem Doppelstrang ergänzt werden. Dabei werden zunächst kurze Primer-Sequenzen, die komplementär zu Anfang und Ende des Bereichs sind, synthetisiert. Nachdem diese Sequenzen sich an den Grenzen des Bereichs angelagert haben, wird die dazwischenliegende Sequenz unter Zuhilfenahme des Enzyms DNA-Polymerase aus den vier Nukleotidarten wieder zu einem Doppelstrang ergänzt.

Zum Beispiel heftet die DNA-Polymerase in *E. coli* pro Minute etwa 11.000 Nukleotide aneinander. Damit steht die Operation „Kopieren“ zur Verfügung.

**Rekombination.** Die wesentlichen Werkzeuge zur Manipulation der DNA sind die Restriktionsenzyme (Endonukleasen), die den Doppelstrang an gewissen Stellen, die durch die lokale Basensequenz festgelegt sind, zerschneiden können. Beispielsweise zerschneiden die Enzyme *TaqI*, *SciNI* bzw. *HhaI* einen Doppelstrang jeweils bei den Sequenzen

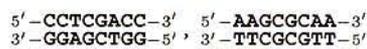
TCGA GCGC bzw. GCGC  
AGCT' CGCG bzw. CGCG'

JUSTUS-LIEBIG-  
UNIVERSITÄT  
GIESSEN

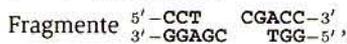
Dr. Martin Kutrib

Institut für Informatik  
Arndtstraße 2  
35392 Gießen  
Telefon (0641) 99-3 21 44  
Telefax (0641) 99-3 21 49  
e-mail martin.kutrib@informatik.uni-giessen.de

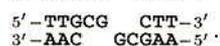
Eine Anwendung der drei Enzyme auf die Moleküle



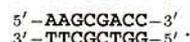
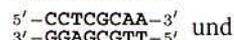
und  $5' - \text{TTGCGCTT} - 3'$   
 $3' - \text{AACCGCAA} - 5'$  liefert die



$5' - \text{AAG} & \text{CGCAA} - 3'$   
 $3' - \text{TTCGC} & \text{GTT} - 5'$  und



Das Enzym DNA-Ligase verbindet „passende“ DNA-Fragmente, die mit Restriktionsenzymen zerschnitten wurden. Hieraus ergibt sich nun die Möglichkeit, neben den Ursprungsmolekülen neue, rekombinierte Moleküle (oder manipulierte Zeichenreihen) zu erzeugen. Im obigen Beispiel sind dies:



Heute sind Tausende verschiedener Restriktionsenzyme bekannt, die an Hunderten von verschiedenen Basensequenzen als spezifisch schneidende Skalpelle ihre Arbeit verrichten können.

**Filtern.** Um einen Einzelstrang mit einer speziellen Teilsequenz aus einem Gemenge herauszufiltern, können zur Teilsequenz (teil-)komplementäre Stränge synthetisiert und an magnetischen Kügelchen befestigt werden. Wenn das Gemenge über die Kügelchen gegossen wird, verbinden sich komplementäre Teilstränge und der gewünschte Effekt wird erzielt.

**Separation und Nachweis von Molekülen bestimmter Länge.** Neben der Möglichkeit, DNA-Moleküle anhand einer Teilsequenz zu filtern, lassen sie sich entsprechend ihrer Länge durch Elektrophorese trennen. Dabei werden die Moleküle, die auf ihrer Oberfläche eine Ladung tragen, auf ein Gel aufgebracht, an das ein elektrisches Feld angelegt ist. Je nach Ladung wandern die einzelnen Moleküle mit unterschiedlicher Geschwindigkeit durch die Trägermasse. Aus dieser Geschwindigkeit lassen sich Rückschlüsse auf die Länge ziehen. Um festzustellen, wie weit ein Molekül gewandert ist, bedient man sich spezifischer Färbe- und Markierungsverfahren.

Spätestens jetzt, nachdem die prinzipielle Arbeitsweise eines molekula-

*Das Rundreiseproblem lässt sich auf eine Vielzahl praktischer Situationen übertragen:*

*So stellt sich einem Briefträger die Frage, in welcher Reihenfolge er am besten die Briefkästen in seinem Bezirk leert.*

*Foto: Ute Voigt*

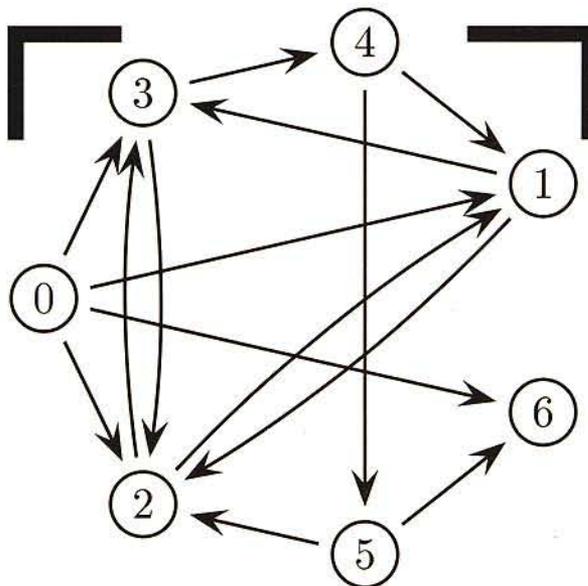


ren Computers geklärt ist, rücken mindestens zwei Fragen in den Vordergrund: Reichen die Möglichkeiten überhaupt aus, sinnvolle, das heißt nicht-triviale, Rechnungen auszuführen? Und wenn ja, ist die praktische Durchführbarkeit vielleicht doch Science Fiction?

**Es funktioniert**

Während die Idee, submikroskopische Rechner zu konstruieren, auf einen Vortrag von Richard Feynman 1959 zurückgeht, gelang es Leonard M. Adleman 1994 als erstem, im Labor eine Rechnung durchzuführen. Er hat gezeigt, wie eine Instanz des sogenannten Hamiltonschen Pfad Problems (HPP) gelöst werden kann. Dieses Problem ist von seiner Komplexität her dem Finden einer kürzesten Rundreise ebenbürtig.

Bei dem HPP handelt es sich um das Problem zu entscheiden, ob es in einem vorgegeben Graphen ausgehend von einem Startknoten einen

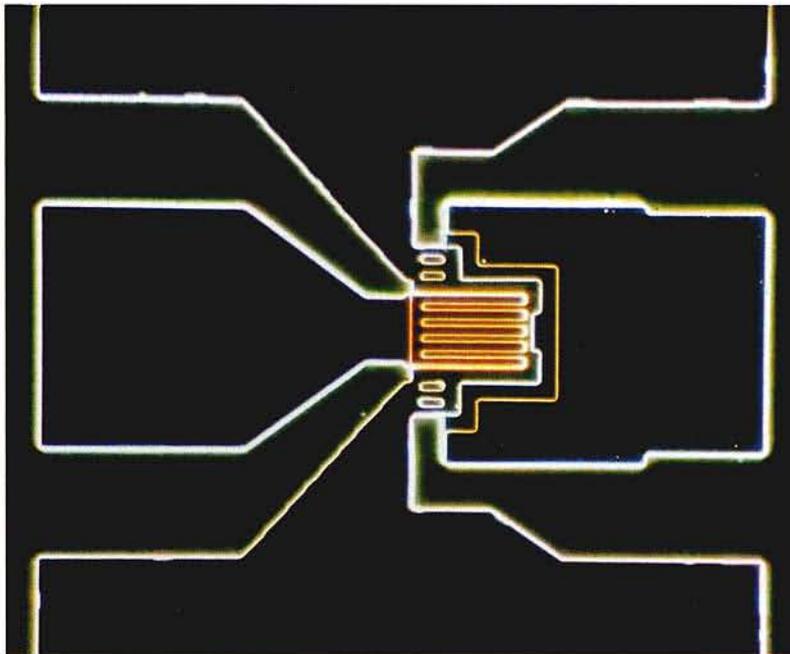


*In diesem Graphen gibt es genau einen Pfad von 0 nach 6, bei dem jeder Knoten nur einmal besucht wird, nämlich den Weg entlang der Pfeile 0 → 1, 1 → 2, 2 → 3, 3 → 4, 4 → 5, 5 → 6. Leonard Adleman demonstrierte an ihm exemplarisch, wie die Lösung eines hartnäckigen Problems mittels eines molekularen Computers gefunden werden kann.*

Weg entlang der Pfeile zu einem Zielknoten gibt, so daß jeder Knoten genau einmal besucht wird. Der in der Abbildung auf der vorhergehenden Seite dargestellte Graph besitzt für den Startknoten 0 und den Zielknoten 6 genau einen Hamiltonschen Pfad, nämlich den Weg entlang der Pfeile  $0 \Rightarrow 1, 1 \Rightarrow 2, 2 \Rightarrow 3, 3 \Rightarrow 4, 4 \Rightarrow 5, 5 \Rightarrow 6$ . Würde man den Knoten 3 als Startknoten und den Knoten 5 als Zielknoten vorgeben, so gäbe es in diesem Graphen keinen Hamiltonschen Pfad von 3 nach 5 (da es beispielsweise keinen Pfeil gibt, der in

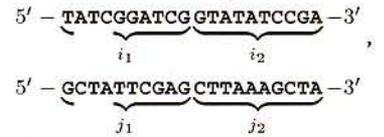
den Knoten 0 hineinführt). Das von Adleman benutzte Lösungsverfahren basiert auf einer Brute-Force-Methode: alle vermeintlichen Lösungskandidaten werden in einer Lösung (in chemischem Sinne) in DNA kodiert erzeugt und dann in einer Folge von Laborschritten (parallel) auf ihre Richtigkeit hin überprüft. Exemplarisch konnte er damit für den in Abbildung 1 angegebenen Graphen den Nachweis der Existenz eines Hamiltonschen Pfades von 0 nach 6 führen.

Zunächst mußte er eine geeignete



Ein Transistor, von denen sich in einem handelsüblichen Prozessor einige Millionen auf der Fläche einer Briefmarke befinden. Nach der Voraussage des sogenannten Mooreschen Gesetz verdoppelt sich alle 18 Monate die Rechenleistung dieser Prozessoren. Und trotzdem gibt es Rechenprobleme in der Mathematik, die auch dann nicht in vernünftigen Zeiten zu lösen sein werden, wenn das Mooresche „Gesetz“ in Zukunft seine Gültigkeit behalten sollte. Bieten „molekulare Rechner“ einen Ausweg?  
Foto: Mercedes Benz

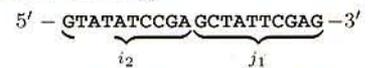
Beschreibung der Lösungskandidaten in Form von DNA finden, das heißt, geeignete Wörter in der genetischen Sprache. Für jeden Knoten wurde eine eindeutige Darstellung als DNA-Molekül bestehend aus 20 Basen gewählt, die sich für die weiteren Schritte hinreichend voneinander zu unterscheiden hatten. Beispielsweise können drei Knoten  $i, j$  und  $k$  durch die Stränge



und



dargestellt werden. Die Darstellung eines Pfeils mittels DNA basiert nun – der entscheidende Trick – auf der Darstellung der beiden Knoten, die der Pfeil miteinander verbindet. Sie ist das DNA-Molekül, das sich durch Verbindung der zehn 3'-Basen des Ausgangsknotens mit den zehn 5'-Basen des Zielknotens ergibt. Der Pfeil  $i \Rightarrow j$  würde damit entsprechend die Darstellung



und der Pfeil  $j \Rightarrow k$  die Darstellung



besitzen.

Um alle möglichen Lösungskandidaten zu erzeugen, wurden hinreichende Mengen von solchen DNA-Molekülen zusammengegeben, die einerseits Pfeile und andererseits die komplementären Stränge von Knoten darstellten (mit Ausnahme der Knoten 0 und 6). In einer Ligase-Reaktion wurden diese Moleküle dann bausteinartig zu DNA-Doppelsträngen zusammengelagert, die denkbaren Pfaden in dem Graphen entsprachen.

Der Pfad  $i \Rightarrow j, j \Rightarrow k$  würde beispielsweise unter Benutzung des Komplements



des Knotens  $j$  durch folgendes Fragment dargestellt:





Martin Kutrib hat Informatik und Raumflugtechnik an der Universität Braunschweig studiert. Seit 1989 ist er in der Arbeitsgruppe Informatik und am Institut für Informatik der Universität Gießen tätig, wo er 1993 promoviert wurde. Zwischenzeitlich war er ein Jahr an der Universität-GH Siegen beschäftigt. Sein besonderes wissenschaftliches Interesse gilt neben der Theorie der Polyautomaten der Komplexitätstheorie und der Parallelität in Algorithmen und Programmiersprachen.

Komplexität	Eingabegröße $n$					
	10	20	30	40	50	60
$n$	0.00001 Sekunden	0.00002 Sekunden	0.00003 Sekunden	0.00004 Sekunden	0.00005 Sekunden	0.00006 Sekunden
$n^2$	0.0001 Sekunden	0.0004 Sekunden	0.0009 Sekunden	0.0016 Sekunden	0.0025 Sekunden	0.0036 Sekunden
$n^3$	0.001 Sekunden	0.008 Sekunden	0.027 Sekunden	0.064 Sekunden	0.125 Sekunden	0.216 Sekunden
$n^5$	0.01 Sekunden	3.2 Sekunden	24.3 Sekunden	1.7 Minuten	5.2 Minuten	13.0 Minuten
$2^n$	0.001 Sekunden	1.0 Sekunden	17.9 Minuten	12.7 Tage	35.7 Jahre	366 Jahrhunderte
$3^n$	0.059 Sekunden	58 Minuten	6.5 Jahre	3855 Jahrhunderte	$2 \cdot 10^8$ Jahrhunderte	$1.3 \cdot 10^{13}$ Jahrhunderte

Die zu erwartende Rechenzeit in Abhängigkeit von der Zeitkomplexität und der Instanzgröße eines Problems. Den Zeiten liegt die Annahme zugrunde, der ausführende Rechner benötige eine Mikrosekunde für jede Elementaroperation (das entspricht  $10^6$  Operationen pro Sekunde). Für  $n$  ist jeweils die Instanzgröße einzusetzen, etwa die Anzahl der Städte zur Bestimmung einer Rundreise, die Anzahl der Knoten eines Graphen, etc.

Die Probleme mit einer polynomiellen Zeitkomplexität in den Zeilen 2 bis 4 stellen auch für größere  $n$  kaum Schwierigkeiten dar, wohingegen Probleme mit exponentieller Zeitkomplexität nur für kleine Größen praktisch berechenbar sind. Insbesondere würde bei letzteren selbst eine Beschleunigung der Ver-

arbeitungszeit um den Faktor 10.000 nur unwesentliche Vorteile bringen. Da für NP-schwierige Probleme bisher nur Lösungsverfahren bekannt sind, die eine exponentielle Zeitkomplexität aufweisen, könnten molekulare Computer hier Lösungen für Instanzgrößen liefern, die heute noch unvorstellbar sind.

Die Mengen an benutzter DNA stellten dabei sicher, daß mindestens ein DNA-Doppelstrang entstand, der einen Hamiltonschen Pfad darstellte, sofern der betrachtete Graph überhaupt über einen solchen verfügte.

Nachdem alle Lösungskandidaten erzeugt waren, wurden mittels Polymerase genau diejenigen verviel-

fältigt, die im Knoten 0 starteten und im Knoten 6 endeten, und für den nächsten Schritt extrahiert. Dieser bestand in einer Elektrophorese, die dazu diente, die Stränge anhand ihrer Länge zu selektieren und nur solche weiterzuverwenden, die Pfade durch genau sieben Knoten repräsentierten.

Nach Auftrennen in Einzelstränge

wurden aus den verbliebenen DNA-Molekülen diejenigen herausgefiltert, die als Teilsequenz jede Knotenkodierung der Knoten 1 bis 5 (mindestens) einmal enthielten. Hierfür waren insgesamt 5 aufeinanderfolgende Filterungen nötig. Die nun noch in der Lösung vorhandenen DNA-Moleküle konnten nur noch Pfade von 0 nach 6 darstellen,



„Ich hab ihn mit dem Rundreiseproblem gefüttert. Er sagt, wir sollten das Reisen aufgeben und in die Finanzbranche wechseln.“

Karikatur: Sidney Harris



Thomas Buchholz studierte in Gießen Mathematik mit Nebenfach Informatik. Seit ein- und einhalb Jahren ist er Mitarbeiter am Institut für Informatik. Dort gilt sein wissenschaftliches Interesse vornehmlich massiv parallelen Maschinenmodellen insbesondere im Hinblick auf komplexitätstheoretische Fragestellungen. Molekulare Computer haben von daher wegen ihrer inhärenten Parallelität sowie hinsichtlich formal-sprachlicher Aspekte seine Aufmerksamkeit erlangt.

die jeden Knoten des Graphen genau einmal besuchten, d.h. Hamiltonsche Pfade von 0 nach 6 waren.

Nach erneuter Vervielfältigung und Separation konnten in der Lösung tatsächlich entsprechende DNA-Moleküle nachgewiesen und das HPP damit positiv beantwortet werden.

### NP-schwierige Probleme

Was aber macht die Existenz eines Hamiltonschen Pfades in einem Graphen so interessant? Das HPP gehört zur Klasse der sogenannten NP-schwierigen Probleme. Sie zeichnet sich dadurch aus, daß es wohl effiziente Algorithmen gibt, um eine vermeintliche Lösung eines dieser Probleme zu verifizieren, man jedoch keine effizienten Verfahren kennt, um eine derartige Lösung zu ermitteln. Die besten bekannten Verfahren weisen eine exponentielle Laufzeit auf, so daß sie in der Praxis nur für kleine Instanzgrößen sinnvoll einsetzbar sind. Die Aussicht auf deren Lösbarkeit, zumindest für mitt-

lere Instanzgrößen, birgt die Motivation, sich mit molekularen Rechnern zu beschäftigen. Es hat sich nämlich gezeigt, daß viele Probleme des täglichen Lebens gerade in diese Klasse fallen.

Das Problem der kürzesten Rundreise hat wichtige wirtschaftliche und industrielle Anwendungen. Bei der Herstellung von gelochten Leiterplatten z. B. muß ein Bohrer alle Lochpositionen nacheinander anfahren. Zur Verkürzung der gesamten Bohrzeit wäre ein möglichst minimaler Gesamtweg wünschenswert. Ein minimaler Weg kann auch Kosten senken, wenn Roboter in automatischen Lagern verschiedene Stationen anfahren oder Briefkästen in einem Bezirk geleert werden müssen.

Ein ähnliches Problem mit offensichtlicher Kostenrelevanz ist das Verlegen von Telefon- oder Datenleitungen zwischen Ortschaften. Weitere NP-schwierige Probleme von praktischem Interesse sind Zuschneide- und dual zu ihnen Packalgorith-

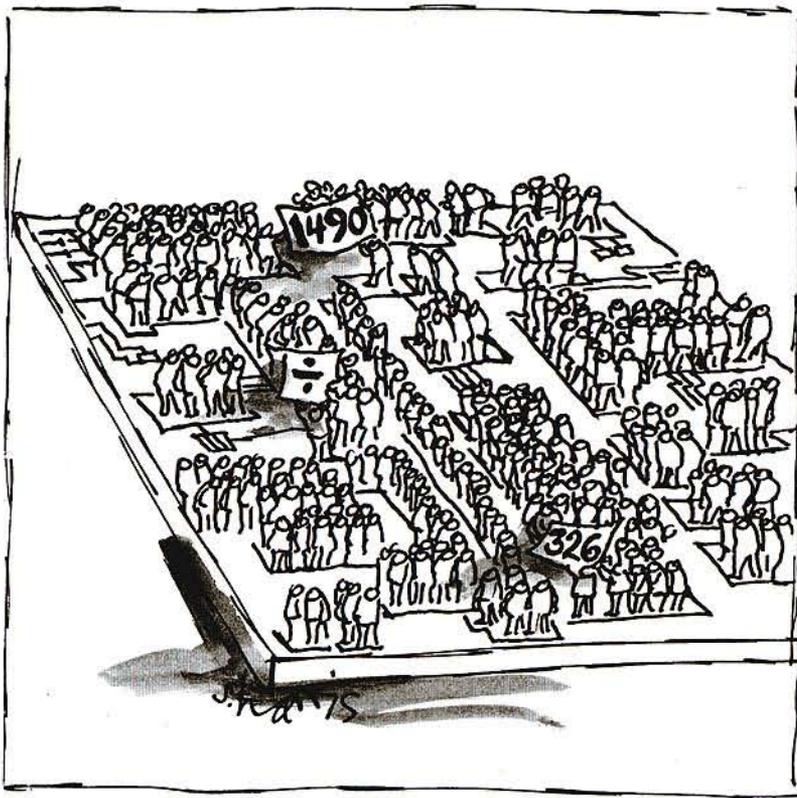
men. Beim Ausstanzen verschiedener Formen beispielsweise aus einem Blech soll der Rest möglichst minimal sein. Das Ausnutzen von zur Verfügung stehender Packfläche auf Paletten und in Containern hat Einfluß auf die anfallenden Logistikkosten.

Notgedrungen arbeitet man in der Praxis mit Heuristiken und Näherungslösungen, um zumindest einen Teil des Einsparungspotentials auszuschöpfen.

### Splicing von Zeichenreihen

Wie Adlemans Versuch gezeigt hat, sind sowohl die Art der Kodierung als auch die anschließende Vorgehensweise problemspezifisch. Obwohl theoretisch möglich, ist man weit davon entfernt, einen in heutigem Sinne universellen molekularen Rechner zu haben. Gerade darum ist es von Interesse, die grundlegenden (Rechen-)Möglichkeiten, die die Rekombination von Molekülen in sich birgt, zu untersuchen.

Die Theoretische Informatik stellt



*Menschlicher Silizium-Chip, fähig zu sechs Berechnungen pro Minute.*

*Karikatur: Sidney Harris*

esse ausschließlich einer möglichen Realisierung in Computern, die mit Prozessoren aus Silizium arbeiten, werden jetzt auch Wege zur Implementierung als molekulare Computer gesucht, nicht zuletzt in der Hoffnung, bisherige Erkenntnisse über die untersuchten Modelle zu übertragen.

#### Grenzen der DNA-Computer

Wohin die Zukunft der molekularen Computer führt, vermag heute wohl kaum jemand zu sagen. Der Stand der Entwicklung läßt sich vergleichen mit dem Stand der Entwicklung elektronischer Rechner zu Zeiten, als der Transistor erfunden wurde. Auch damals haben nur die wenigsten geahnt, in welchem Ausmaß Computer fast alle Bereiche beeinflussen werden. Werden eines Tages die molekularen die elektronischen Rechner ersetzen? Da jede primitive Operation im Reagenzglas einige Stunden dauern kann, werden auch zukünftig elektronische Rechner für einfache Aufgaben die bessere (oder einzige) Wahl sein. Der Vorteil der massiven Parallelität sollte aber Probleme mit sehr hoher Zeitkomplexität für Instanzgrößen in vernünftiger Zeit lösbar machen, für die herkömmliche Computer sehr viele Jahre bräuchten.

Es bleibt die Frage nach der Lösbarkeit schwieriger Probleme mit hoher Instanzgröße. Wird es molekularen Computern gelingen, z.B. sehr große Zahlen in relativ kurzer Zeit zu faktorisieren? Viele Verschlüsselungssysteme basieren auf dem Unvermögen von herkömmlichen Rechnern, genau dieses zu leisten. Die Konsequenzen wären enorm, um das mindeste zu sagen. Mal abgesehen von vielen technischen Schwierigkeiten, die hier nicht im Detail erwähnt werden konnten, würde das Verfahren von Adleman angewandt auf einen Graphen mit 200 Knoten eine DNA-Menge erfordern, die das Gewicht der Erde überträte. Also: Auch in Zukunft erfordern schwere Probleme schwere Computer. •

dabei entsprechende Methoden bereit. Ein informationstragendes Makromolekül wird repräsentiert als Wort einer formalen Sprache über einem zunächst beliebigen Alphabet. Die Wirkung von Restriktionsenzymen und Ligase wird durch sogenannte Splicing-Regeln modelliert. An Stellen, die durch lokale Teilwörter präzisiert werden, werden zwei Wörter zertrennt und die Fragmente anschließend vertauscht wieder zusammgefügt. Eine Menge von Splicing-Regeln kann selbst wieder als formale Sprache aufgefaßt werden.

Formale Sprachen lassen sich zum Beispiel anhand ihrer strukturellen Komplexität zu Familien zusammenfassen. Eine Fragestellung ist jetzt, aus welchen Familien die Sprachen der Ausgangswörter und der Splicing-Regeln zu wählen sind, damit das Ergebnis der wiederholten Anwendung der Regeln, wiederum eine formale Sprache, zur allgemeinsten Familie gehört? Oder anders ausgedrückt, wie kompliziert müssen Regeln und Ausgangsmaterial gewählt werden, um ein System zu erhalten, das genauso leistungsfähig ist wie eine Turingmaschine? Einerseits konnte nachgewiesen werden, daß kompliziertere Regeln einfaches Ausgangsmaterial ermöglichen und umgekehrt. Andererseits existieren sogar unendliche Hierarchien von Sprachfamilien als Ergebnis des wiederholten Splicing in Abhängigkeit von der Kompliziertheit der Regelmengen.

Die Informatik stellt verschiedene Modelle für existierende und (noch) nicht existierende Rechnertypen zur Verfügung. Galt das bisherige Inter-

JUSTUS-LIEBIG-  
UNIVERSITÄT  
GIESSEN

Dipl.-Math. Thomas Buchholz

Institut für Informatik  
Arndtstraße 2  
35392 Gießen  
Telefon (0641) 99-32143  
Telefax (0641) 99-32149  
e-mail thomas.buchholz@informatik.uni-giessen.de