

Rechnen mit DNS

von Eberhard Bertsch & Peter Imming

Studium Integrale Journal
17. Jahrgang / Heft 1 - Mai 2010

Seite 16 - 20

Zusammenfassung: DNS-Moleküle enthalten in kodierter Form den Bauplan für Proteine, die in Lebewesen als Baustoffe oder Biokatalysatoren fungieren. Der Bauplancharakter und die – verglichen mit anderen Biopolymeren – relativ leichte Herstellbarkeit von DNS, **a to, že lze DNA poměrně snadno vyrobit** wird gentechnisch inzwischen vielfältig genutzt. Die Möglichkeiten, der menschlichen Nutzung von DNS beschränken sich jedoch nicht darauf **Možnosti lidského využití DANN tím nejsou vyčerpány**. Seit etwa zwanzig Jahren wird theoretisch und experimentell daran gearbeitet, DNS-Sequenzen auch als Träger von Rechenoperationen **početních operací** nutzbar zu machen, also „Computer“ zu bauen, deren Kernbestandteile Nukleotidketten **nukleotid** und Eiweiße sind. Von besonderer Bedeutung war dabei ein Experiment von Adleman (1994) **Adelmannův**, das die prinzipielle **Lösbarkeit** einer, **která je jak známo obtížná** bekannt schwierigen mathematischen Aufgabe **demonstrier**. Es konnte gezeigt werden **Podářilo se mu prokázat**, dass auf DNS-Molekülen Programme für beliebige Zwecke darstellbar sind. Der Sachverhalt **Tato skutečnost tedy umožňuje srovnat molekoly DANN s technickým hardwarem, na němž je..legt** somit* den Vergleich mit einer technischen Hardware für darauf programmierte Software **nahe**. Nach Auffassung der Autoren ist er ein weiteres Indiz dafür, dass DNS als ein intelligentes Konstrukt angesehen werden sollte und nicht als Produkt **nepromyšlených náhodilostí a nutností** undurchdachter Zufälle und Notwendigkeiten.

* **so|mit** <Adv.>: *wie daraus zu schließen, zu folgern ist; folglich, also, mithin:*

In diesem Artikel möchten wir ein Thema vorstellen, **na němž pracují odborníci přibližně dvě desetiletí** das seit etwa zwei Jahrzehnten von Fachleuten bearbeitet wird und in dieser Zeit zu vielen theoretischen und praktischen Ergebnissen geführt hat, aber wegen der Komplexität der **Fragestellung komplexnosti otázek/problematiky, které jsou s ní spjaty**, bisher noch nicht in der breiten Öffentlichkeit bekannt ist. Es geht um das Rechnen – genauer: die Informationsverarbeitung – mit Desoxyribonukleinsäure, abgekürzt DNS, dem Trägermaterial der genetischen Eigenschaften aller Lebewesen **živočichů**. Es geht nicht um die Struktur von DNS als solcher, um Gentechnik oder Gentherapie, sondern wir **zeigen** einen erstaunlichen weiteren Aspekt der ohnehin genial konstruierten DNS **auf**. **Rádi bychom upozornili na další překvapivý aspekt DNA**

oh|ne|hin <Adv.>: **unabhängig davon; auf jeden Fall; sowieso**: das hätte uns o. nichts genützt.

cukr-fosfátové kostra

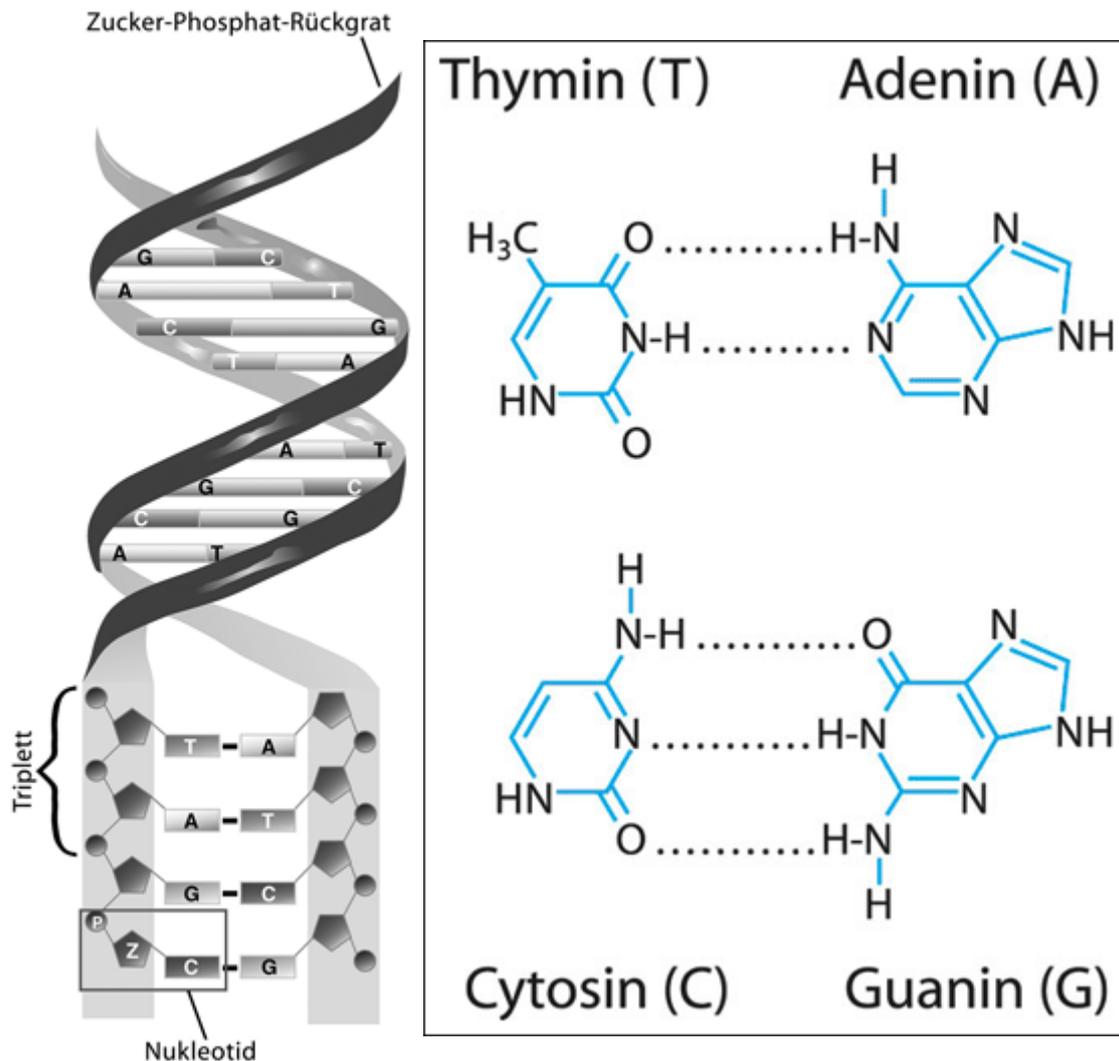


Abb. 1: Aufbau der DNS. Links Doppelhelix, Z = Zucker Desoxyribose, P Phosphatrest. Rechts die vier Grundbausteine (Basen) der DNS. Je zwei Basen passen genau zueinander und sind daher „komplementär“. Die Basenpaarung erfolgt über Wasserstoffbrücken (gepunktete Linien). Mittels der DNS-Sequenz (Abfolge) werden die Reihenfolgen der Aminosäuren in den Proteinen programmiert. Ein Triplett steht für eine Aminosäure. (Aus Junker & Scherer 2006)

Struktura DNA:

Nalevo: **Dvoušroubovice**, **cukr**, deoxyribózea, [deoxyribonukleové kyseliny](#)

Vždy dvě báze se do sebe přesně zapadají a jsou tedy komplementární, Báze se párují prostřednictvím vodíkových můstků (vytečkované linky). Prostřednictvím sekvence (posloupnosti) DNA jsou naprogramována pořadí aminokyselin v bílkovinách. Jeden triplet kóduje jednu aminokyselinu.

Heute übliche Rechner (Computer) funktionieren auf der Grundlage eng vernetzter („integrierter“) elektronischer Schaltungen **elektronických přípojů**. Deren elementare Bausteine sind Feldeffekt-Transistoren **tranzistory řízené elektrickým polem**, zumeist aus Silizium, von denen viele Millionen auf einer Fläche von der Größe **velikosti** eines Fingernagels untergebracht sind. Vor der Entdeckung und technischen Nutzung von Transistoren um die Mitte des vorigen Jahrhunderts gab es mechanische und elektromechanische Rechner, deren Funktion auf anderen physikalischen Grundlagen beruhte.

Auch mit optischen, sogar mit hydraulischen Bauteilen könnten Rechner konstruiert werden.

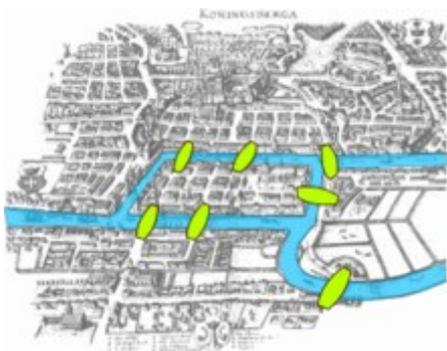
Hier ist nun die Rede vom Rechnen mit DNS, also der Nutzung einer in Lebewesen vorkommenden Substanz als Träger der Rechenschritte **nositel početních úkonů** (s. Glossar, „Informationsverarbeitung“). Die Rechen- und Speicherungsschritte **Jednotlivé početní kroky jakož i úkony ukládání informace**, die in heutigen Computern mit elektromagnetischen

Mitteln durchgeführt werden, erfolgen **se děje** mittels DNS-Ketten. Die allgemeine Struktur von DNS-Sequenzen **ukazuje obr.** wird in Abb. 1 gezeigt.

Voraussetzung dafür ist das tiefere Verständnis **hlubší porozumění molekulárním procesům** der molekularen Prozesse in lebenden Zellen, das seit etwa Mitte der sechziger Jahre gewonnen wurde. Zwar handelt es sich für die praktische Anwendung um Zukunftsmusik **Praktické využití je sice hudbou budoucnosti, avšak existuje množství důvodů, pro něž jsou..než klasické.** Aber es gibt durchaus Gründe, warum Rechnen mit DNS Vorteile gegenüber klassischem Rechnen haben könnte. Hauptsächlicher Vorteil ist die gleichzeitige Ausführbarkeit einzelner Schritte an sehr vielen Stellen innerhalb eines kleinen Volumens – sozusagen ein weiteres nanotechnologisches Verfahren. Die kompakte Speicherung in DNS-Ketten ermöglicht es, **že díky paralelnosti probíhá mnoho procesů na celkově malém prostoru.** viele Prozesse zwecks Parallelität in einem insgesamt kleinen Raum auszuführen. Parallelität kennen wir auch von der Informationsverarbeitung im Gehirn und im Auge. Die Bildverarbeitung kann in verschiedenen Bereichen des Auges parallel ablaufen. Die 1994 von Leonard Adleman veröffentlichte Beschreibung eines Experiments, bei dem mit DNS-Molekülen gerechnet wurde, führte zu großem Aufsehen in der Fachwelt **vzbudil v odborném světě velký rozruch.** Das Experiment benutzt DNS zur Ausführung eines Algorithmus, der mathematische Fragen aus der Graphentheorie löst.

Die im Experiment **zu beantwortende** Frage ist, **zněla, zda existuje taková posloupnost hran, aby spojila* všechny vrcholy spojí přesně jednou** ob es in einem vorgegebenen Graphen eine Kantenfolge gibt, die alle Knoten genau einmal erreicht. Interessanterweise wurde dieses Problem – natürlich ohne die technische Nutzung von Computern oder DNS – schon 1736 von dem berühmten Mathematiker Leonhard Euler behandelt, und zwar anhand der Brücken der Stadt Königsberg (heute Kaliningrad). Die Frage war, ob es einen Rundweg gibt, bei dem man alle sieben Brücken der Stadt über den Fluss Pregel genau einmal überquert. Euler bewies mit einer einfachen Argumentation, dass es keinen solchen Rundweg gab.

1}Graph «, **Graf, der; -en, -en** [zu griech. gráphein= schreiben] (bes. Math., Naturw.): *grafische Darstellung (z. j. B. von Relationen) in Form von [markierten] Knoten[punkten] u. verbindenden Linien (Kanten).*



Die Lösung solcher Aufgaben erfordert bei großer Knotenanzahl viel Rechenzeit **dlouhou výpočetní dobu.** Wenn man statt sieben Brücken (Kanten) eine größere Anzahl hat, steigt der Zeitaufwand **roste doba výpočtu nesmírně rychle** rasend schnell. Durch hohen Parallelitätsgrad – wenn also sehr viel gleichzeitig **úkonů zároveň** ausgerechnet wird – lässt sich der Zeitbedarf stark senken. Im Falle des genannten Adleman-Experimentes wurde ein Graph mit sieben Knoten untersucht. Die Berechnung mit DNS dauerte mehrere Tage. Das ist natürlich im Vergleich zum Zeitbedarf eines Computers **to ovšem není porovnáno časem, za**

něž tuto operaci zvládne počítač nikterak ohromující alles andere als beeindruckend. Das Wichtige und Aufsehen erregende war aber, dass eine Aufgabe dieser Art überhaupt mit DNS berechnet werden konnte.

Einführung

In diesem Artikel möchten wir ein Thema vorstellen, das seit etwa zwei Jahrzehnten von Fachleuten bearbeitet wird und in dieser Zeit zu vielen theoretischen und praktischen Ergebnissen geführt hat, aber wegen der **Komplexität** der Fragestellung bisher noch nicht in der breiten Öffentlichkeit bekannt ist. Es geht um das Rechnen – genauer: die Informationsverarbeitung – mit Desoxyribonukleinsäure, abgekürzt DNS, dem Trägermaterial der genetischen Eigenschaften aller Lebewesen. Es geht nicht um die Struktur von DNS als solcher, um Gentechnik oder Genthherapie, sondern wir zeigen einen erstaunlichen weiteren Aspekt der ohnehin genial konstruierten DNS auf.

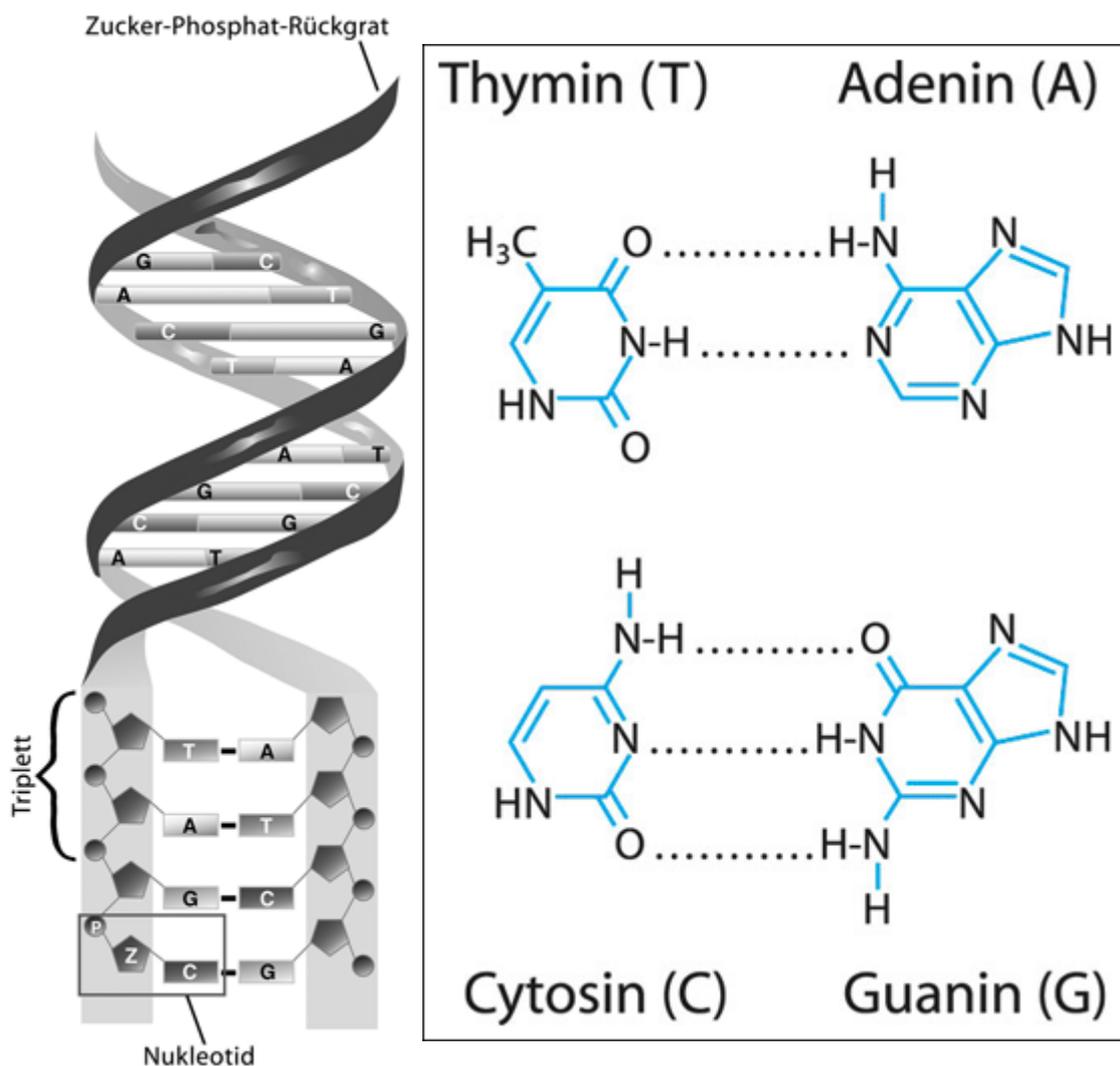


Abb. 1: Aufbau der DNS. Links Doppelhelix, Z = Zucker Desoxyribose, P Phosphatrest. Rechts die vier Grundbausteine (Basen) der DNS. Je zwei Basen passen genau zueinander und sind daher „komplementär“. Die Basenpaarung erfolgt über Wasserstoffbrücken (gepunktete Linien). Mittels der DNS-Sequenz (Abfolge) werden die Reihenfolgen der Aminosäuren in den Proteinen programmiert. Ein Triplett steht

für eine Aminosäure. (Aus Junker & Scherer 2006)

Heute übliche Rechner (Computer) funktionieren auf der Grundlage eng vernetzter („integrierter“) elektronischer Schaltungen. Deren elementare Bausteine sind Feldeffekt-Transistoren, zumeist aus Silizium, von denen viele Millionen auf einer Fläche von der Größe eines Fingernagels untergebracht sind. Vor der Entdeckung und technischen Nutzung von Transistoren um die Mitte des vorigen Jahrhunderts gab es mechanische und elektromechanische Rechner, deren Funktion auf anderen physikalischen Grundlagen beruhte.

Auch mit optischen, sogar mit hydraulischen Bauteilen könnten Rechner konstruiert werden.

Hier ist nun die Rede vom Rechnen mit DNS, also der Nutzung einer in Lebewesen vorkommenden Substanz als Träger der Rechenschritte (s. Glossar, „Informationsverarbeitung“). Die Rechen- und Speicherungsschritte, die in heutigen Computern mit elektromagnetischen Mitteln durchgeführt werden, erfolgen mittels DNS-Ketten. Die allgemeine Struktur von DNS-Sequenzen wird in Abb. 1 gezeigt.

Voraussetzung dafür ist das tiefere Verständnis der molekularen Prozesse in lebenden Zellen, das seit etwa Mitte der sechziger Jahre gewonnen wurde. Zwar handelt es sich für die praktische Anwendung um Zukunftsmusik. Aber es gibt durchaus Gründe, warum Rechnen mit DNS Vorteile gegenüber klassischem Rechnen haben könnte. Hauptsächlicher Vorteil ist die gleichzeitige Ausführbarkeit einzelner Schritte an sehr vielen Stellen innerhalb eines kleinen Volumens – sozusagen ein weiteres nanotechnologisches Verfahren. Die kompakte Speicherung in DNS-Ketten ermöglicht es, viele Prozesse zwecks Parallelität in einem insgesamt kleinen Raum auszuführen. Parallelität kennen wir auch von der Informationsverarbeitung im Gehirn und im Auge. Die Bildverarbeitung kann in verschiedenen Bereichen des Auges parallel ablaufen. Die 1994 von Leonard Adleman veröffentlichte Beschreibung eines Experiments, bei dem mit DNS-Molekülen gerechnet wurde, führte **zu großem Aufsehen in der Fachwelt**. Das Experiment benutzt DNS zur Ausführung eines Algorithmus, der mathematische Fragen aus der Graphentheorie löst.

Die im Experiment zu **beantwortende** Frage ist, ob es in einem vorgegebenen Graphen eine Kantenfolge gibt, die alle Knoten genau einmal erreicht. Interessanterweise wurde dieses Problem – natürlich ohne die technische Nutzung von Computern oder DNS – schon 1736 von dem berühmten Mathematiker Leonhard Euler behandelt, und zwar anhand der Brücken der Stadt Königsberg (heute Kaliningrad). Die Frage war, ob es einen Rundweg gibt, bei dem man alle sieben Brücken der Stadt über den Fluss Pregel genau einmal überquert. Euler bewies mit einer einfachen Argumentation, dass es keinen solchen Rundweg gab.

Die Lösung solcher Aufgaben erfordert bei großer Knotenanzahl viel Rechenzeit. Wenn man statt sieben Brücken (Kanten) eine größere Anzahl hat, steigt der Zeitaufwand rasend schnell. Durch hohen Parallelitätsgrad – wenn also sehr viel gleichzeitig ausgerechnet wird – lässt sich der Zeitbedarf stark senken. Im Falle des genannten Adleman-Experimentes wurde ein Graph mit sieben Knoten untersucht. Die Berechnung mit DNS dauerte mehrere Tage. Das ist natürlich im Vergleich zum Zeitbedarf eines Computers alles andere als beeindruckend. Das Wichtige und Aufsehen erregende war aber, dass eine Aufgabe dieser Art überhaupt mit DNS berechnet werden konnte.

Wir beschreiben den Aufbau und den Ablauf des Adleman'schen Verfahrens hier in groben Zügen.

Einzelheiten sind Adleman (1994) oder auch dem Lehrbuch von Hinze & Sturm (2004) zu entnehmen.

Jeder Knoten und jede Kante wird durch ein einzelnes DNS-Molekül dargestellt, und zwar durch einen ganz

bestimmten **DNS-Einzelstrang** *jednotlivými vlákny* |DNS mit einer bestimmten Sequenz aus 20 Basen (20mer, lineare Abfolge aus 20 Basen). Die Zuordnung zwischen den Objekten im Graphen und den DNS-Einzelsträngen geschieht willkürlich, aber verbindlich. Ein bestimmtes **20mer** *20-mer* repräsentiert in allen zusammengesetzten Strängen immer denselben Knoten oder dieselbe Kante. Im Experiment werden 7 Knoten und 14 Kanten verwendet. Die vorderen 10 Basen jeder Kantendarstellung sind komplementär zu den hinteren 10 einer Knotendarstellung, die hinteren 10 der Kantendarstellung sind komplementär zu den vorderen 10 einer anderen Knotendarstellung. Auf diese Weise lassen sich mittels der chemischen Paarungsgesetzmäßigkeiten (A immer mit T, C immer mit G) lange **Doppelstränge** *dvouvlákna* bilden, in denen der eine **Strang** aus einer Abfolge von kodierten Kanten und der andere Strang aus einer Abfolge von kodierten Knoten besteht. Die Kantenfolge verbindet Knoten auf einem Pfad innerhalb des dargestellten Graphen. Meistens gibt es in einem einzelnen Graphen viele unterschiedliche Pfade.

<http://gene-abc.ch/de/lexikon/deutsches-lexikon/d-f.html>

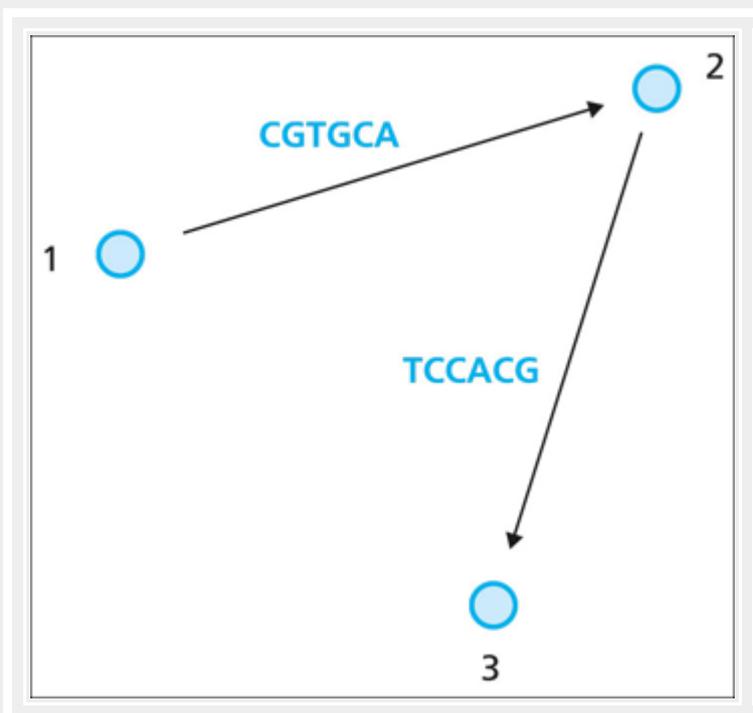
Die Abbildung im **Kastentext** *obr. V rámečku ukazuje* zeigt einen aus zwei Kanten und drei Knoten bestehenden **daného** Graphen. Der einfacheren Darstellung **Pro zjednodušení není vlákno tvořeno dvaceti bázemi, nýbrž poze šesti halber** sind die Stränge nicht 20, sondern nur 6 Basen lang.

Grafik zu „Rechnen mit DNS“

Wir wollen den „Weg“ durch den aus drei Knoten 1, 2 und 3 und zwei Kanten (1, 2) und (2, 3) bestehenden Graphen (Abb. 2) in Form einer Nukleotidsequenz ausdrücken.

Zu diesem **Zweck sei vorab vereinbart**, dass die drei Knoten 1, 2 und 3 durch die Nukleotidketten ATCGCA, CGTAGG und TGCAGA und die beiden Kanten (1, 2) und (2, 3) durch CGTGCA und TCCACG dargestellt werden müssen, also in der Weise, dass die Darstellung jeder Kante gerade die Komplementärkette zur hinteren Hälfte ihres ersten und zur vorderen Hälfte ihres zweiten Knotens ist.

Dann drückt die folgende Verkettung der drei Knoten 1, 2, 3 im oberen Strang und der beiden Kanten (1, 2) und (2, 3) im dazu komplementären unteren Strang gerade aus, dass es einen Weg gibt, der von 1 über 2 nach 3 verläuft.



Knotenfolge

A T C G C A - C G T A G G - T G C A G A
 C G T G C A - T C C A C G

Kantenfolge

Die Bindestriche **Pomlčky stojí na těch místech** sind hier an solchen Stellen angegeben, wo **Ausgangsketten** auf der einen oder anderen Seite des **Doppelstrangs** zusammengefügt wurden.

Im Adleman-Experiment wurden zunächst Billionen kurzer Ketten für die Knoten und Kanten hergestellt; und es wurde diesen ermöglicht, sich in jeder beliebigen Weise aneinanderzufügen.

Die entstandenen Doppelstränge wurden danach mit molekularbiologischen Methoden daraufhin überprüft **U ... bylo přezkoumáno, zda** ob sie mit einem bestimmten Teilstück **anfingen**, mit einem bestimmten anderen aufhörten und dazwischen jedes einen Knoten darstellende Teilstück genau einmal vorkam.

All diejenigen, auf die das zutrifft, sind Lösungen des zuvor beschriebenen mathematischen Problems.
Všechny tyto části

Zunächst wird durch biochemische Synthese dafür gesorgt, dass alle Stränge billionenfach vervielfältigt vorkommen. Dann werden sie in ein gemeinsames Reagenzgefäß umgefüllt. Es wird davon ausgegangen, dass sich alle DNS-Stücke zusammenlagern, die komplementäre Überlappungen aufweisen.

Wahrscheinlichkeitsaussagen oder Untersuchungen darüber finden sich allerdings in der Originalarbeit nicht. Damit alle möglichen Stränge mit komplementären Überlappungen entstehen, muss die Menge der in der biochemischen Synthese verwendeten DNS-Stränge in exponentieller Größenordnung in der Menge der Knoten und Kanten des Problems sein. Als nächstes werden unter den nun vorhandenen Ketten nur die weiter verwendet, die mit dem gewünschten Knoten beginnen und mit dem anderen gewünschten Knoten enden.

Labortechnisch geschieht dies durch DNS-Primerstränge* mit bestimmten kurzen Basenfolgen, die eine starke Vervielfältigung für alle mit dieser Basenfolge am Anfang übereinstimmenden Stränge bewirken. Hierbei kommt es außerdem auf die präzise Einstellung der Reaktionstemperatur an.

Dann sind die Stränge mit der richtigen Länge auszuwählen. In der hier beschriebenen Anordnung sind es diejenigen mit einer Länge von 140 Basenpaaren. Dies geschieht auf mechanischem Wege mit Hilfe einer Gelelektrophorese*, die die vorhandenen Stränge nach deren „Länge“ „sortiert“. Stränge mit einem Längenunterschied von etwa 10 Basenpaaren lassen sich so voneinander trennen. Das ist ein Grund für die Wahl der Länge 20 bei den einzelnen Knoten und Kanten.

Nachdem nun diejenigen Stränge ausgewählt sind, die richtig anfangen und richtig aufhören und außerdem die richtige Länge haben, ist noch zu prüfen, ob jeder Knoten in den Strängen vorkommt. Dazu sind so viele weitere Separationen nötig, wie der Graph Knoten enthält. Jeder dieser Schritte folgt einem komplizierten Protokoll, in dem eine **Anlagerung** von Strängen an den zu einer Knotendarstellung komplementären Strang vorgenommen wird. Dabei wird der vorgegebene Strang **biotinyliert***, was die Möglichkeit der räumlichen Fixierung und anschließenden Trennung der ankoppelnden Stränge durch Zentrifugieren zur Folge hat. Sobald nur diejenigen Stränge übrig sein können, in denen jede Kantendarstellung vorkommt, ist zu fragen, ob tatsächlich solche Stränge übrig sind oder nicht. Dazu dient wieder eine Elektrophorese.

Auf mathematischem Wege konnte in den vergangenen **Jahren bewiesen** werden, dass jede überhaupt mögliche Berechnung (Turing-Berechenbarkeit [Asteroth & Baier 2002]) zumindest prinzipiell – wenngleich in vielen Fällen nur mit unverträglich hohem Aufwand an Material und Zeit – von DNS-Molekülen durchgeführt werden kann.

Glossar

Biotinylierung: Biotinylierung ist eine biochemische Methode, bei der z.B. eine makromolekulare Nukleinsäure wie DNS mit einem kleinen Molekül namens Biotin verbunden wird. Die Biotin-Gruppierung bindet sehr stark an bestimmte Proteine, die Avidine. Diese starke Bindung kann benutzt werden, Moleküle mit Biotin-Einheiten aus einem Gemisch „herauszufischen“. **Gelelektrophorese:** Die Gelelektrophorese ist ein biochemisches Trennverfahren, bei dem die Wanderung von geladenen Teilchen (Molekülen) in einem elektrischen Feld zu ihrer Trennung benutzt wird. Man löst das zu trennende Stoffgemisch und lässt die Lösung durch eine elektrisch neutrale, feste Gelmatrix aus Agarose oder Polyacrylamid wandern. Die Wanderungsgeschwindigkeit hängt von der Stärke des Feldes, der Nettoladung, Form und Größe der Makromoleküle sowie von physikochemischen Eigenschaften der Lösung ab. **Graphentheorie:** Die Graphentheorie beschäftigt sich mit Strukturen, die aus einzelnen Punkten (genannt Knoten, englisch: nodes) und Verbindungen zwischen diesen (genannt Kanten, englisch: edges) bestehen. Durch direkte Verbindungen zwischen einzelnen Knoten sind dann auch indirekte Verbindungen gegeben, die mehrere Kanten nacheinander durchlaufen, sowie Zyklen, die über mehrere Kanten zum Ausgangsknoten zurückführen. Ein Beispiel für den praktischen Nutzen der Graphentheorie ist die Herstellung und dynamische Änderung von Verbindungen im Telefonnetz oder Internet. Auch das „Problem des Handelsreisenden“ (Travelling Salesman Problem) ist Gegenstand der algorithmischen Graphentheorie. Es geht dabei um Verfahren, die auf einer Landkarte die insgesamt kürzeste Rundreise über alle ausgewählten Städte bestimmen. Mit geringen Änderungen tritt diese Frage auch in der industriellen Fertigung (Robotik) auf. Man kann beweisen, dass optimale Lösungen für große Graphen ebenso schwer zu finden sind wie bei dem von Adleman behandelten Graphenproblem. **Informationsverarbeitung:** Das Rechnen mit Zahlen, das wir schon in der Grundschule lernen, ist ein spezieller Fall der „Informationsverarbeitung“. Informationsverarbeitung sind alle schrittweisen, nach bestimmten Regeln erfolgenden Veränderungen von Zeichenfolgen. Zeichenfolgen können beispielsweise aus Ziffern oder Buchstaben bestehen. Etwa im Falle der chinesischen Schriftsprache besteht der Zeichenvorrat, mit dem Folgen gebildet werden können, aus Tausenden unterschiedlicher Zeichen. Regeln können lauten: Ersetze Zeichen „1 x 1“ durch das Zeichen „2“, ersetze Zeichen „2 x 2“ durch das Zeichen „4“, und so weiter, wenn wir als Beispiel das Multiplizieren nehmen. Eine andere einfache Möglichkeit ist: Ersetze alle „a“ durch „b“, alle „b“ durch „c“, alle „c“ durch „d“ und so weiter. Bei diesem Verfahren handelt sich um eine schon von Gaius Julius Caesar verwendete Verschlüsselungstechnik. Aus informationsverarbeitenden Schritten aufgebaute Verfahren werden Algorithmen genannt. **Primer:** Primer sind kurze DNS-Sequenzen, die chemisch-synthetisch hergestellt und einem Reagenzgemisch zugesetzt werden, das DNS vervielfältigt. Sie dienen als Startsequenz für DNS-vervielfältigende Enzyme, die DNS-Polymerasen, die im Gemisch enthalten sind. DNS-Polymerasen können nur die Verlängerung eines schon existierenden kurzen DNS-Strangs katalysieren. Die DNS-Vervielfältigung verläuft wie bei der Zellteilung lebender Zellen.

Einführung

In diesem Artikel möchten wir ein Thema vorstellen, das seit etwa zwei Jahrzehnten von Fachleuten bearbeitet wird und in dieser Zeit zu vielen theoretischen und praktischen Ergebnissen geführt hat, aber wegen der **Komplexität** der Fragestellung bisher noch nicht in der breiten Öffentlichkeit bekannt ist. Es geht um das Rechnen – genauer: die Informationsverarbeitung – mit Desoxyribonukleinsäure, abgekürzt DNS, dem Trägermaterial der genetischen Eigenschaften aller Lebewesen. Es geht nicht um die Struktur von DNS als solcher, um Gentechnik oder Genthherapie, sondern wir zeigen einen erstaunlichen weiteren Aspekt der ohnehin genial konstruierten DNS auf.

Heute übliche Rechner (Computer) funktionieren auf der Grundlage eng vernetzter („integrierter“) elektronischer Schaltungen. Deren elementare Bausteine sind Feldeffekt-Transistoren, zumeist aus Silizium, von denen viele Millionen auf einer Fläche von der Größe eines Fingernagels untergebracht sind. Vor der Entdeckung und technischen Nutzung von Transistoren um die Mitte des vorigen Jahrhunderts gab es mechanische und elektromechanische Rechner, deren Funktion auf anderen physikalischen Grundlagen beruhte.

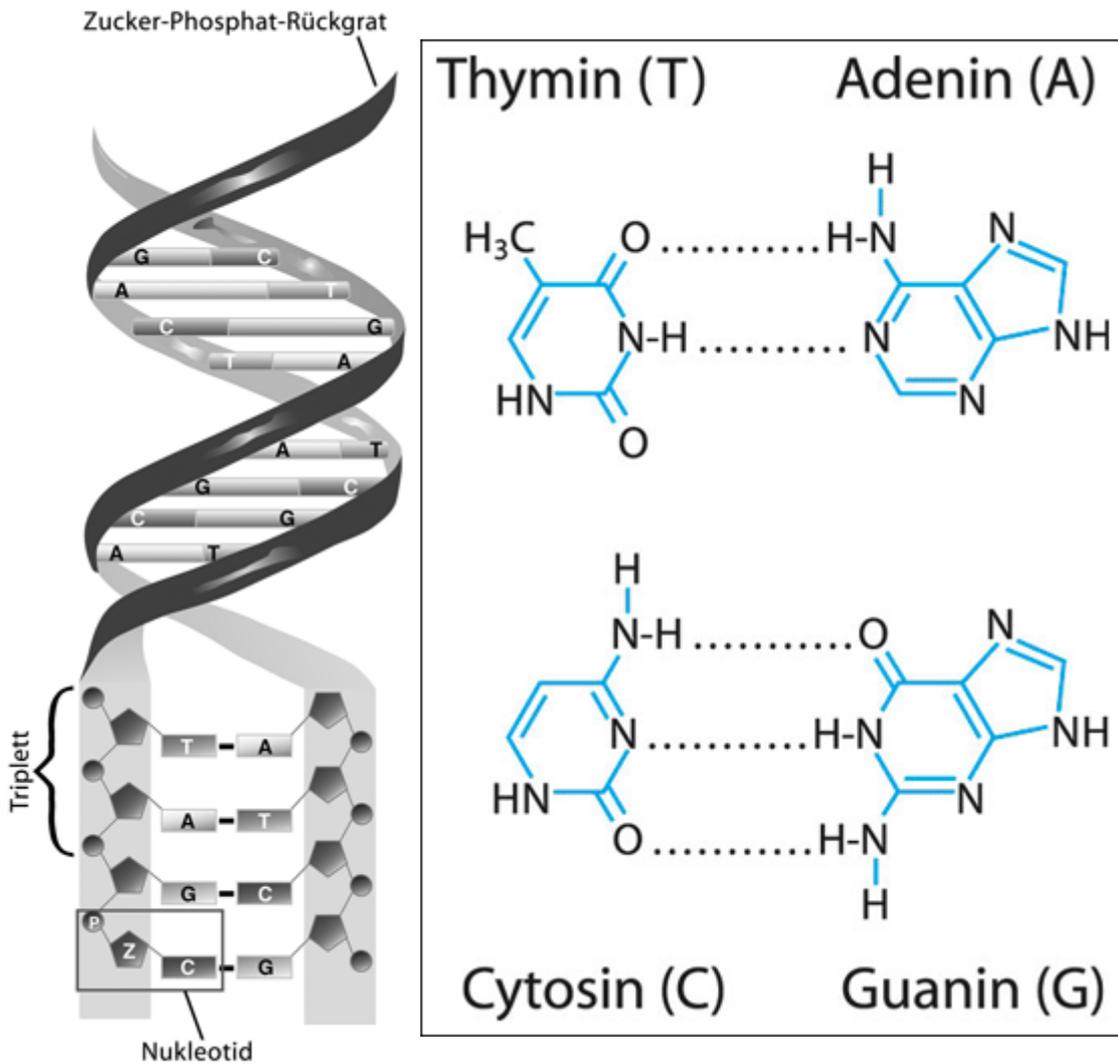


Abb. 1: Aufbau der DNS. Links Doppelhelix, Z = Zucker Desoxyribose, P Phosphatrest. Rechts die vier Grundbausteine (Basen) der DNS. Je zwei Basen passen genau zueinander und sind daher „komplementär“. Die Basenpaarung erfolgt über Wasserstoffbrücken (gepunktete Linien). Mittels der DNS-Sequenz (Abfolge) werden die Reihenfolgen der Aminosäuren in den Proteinen programmiert. Ein Triplettsymbol steht für eine Aminosäure. (Aus Junker & Scherer 2006)

Auch mit optischen, sogar mit hydraulischen Bauteilen könnten Rechner konstruiert werden.

Hier ist nun die Rede vom Rechnen mit DNS, also der Nutzung einer in Lebewesen vorkommenden Substanz als Träger der Rechenschritte (s. Glossar, „Informationsverarbeitung“). Die Rechen- und Speicherungsschritte, die in heutigen Computern mit elektromagnetischen Mitteln durchgeführt werden, erfolgen mittels DNS-Ketten. Die allgemeine Struktur von DNS-Sequenzen wird in Abb. 1 gezeigt.

Voraussetzung dafür ist das tiefere Verständnis der molekularen Prozesse in lebenden Zellen, das seit etwa Mitte der sechziger Jahre gewonnen wurde. Zwar handelt es sich für die praktische Anwendung um Zukunftsmusik. Aber es gibt durchaus Gründe, warum Rechnen mit DNS Vorteile gegenüber klassischem Rechnen haben könnte. Hauptsächlich Vorteil ist die gleichzeitige Ausführbarkeit einzelner Schritte an

sehr vielen Stellen innerhalb eines kleinen Volumens – sozusagen ein weiteres nanotechnologisches Verfahren. Die kompakte Speicherung in DNS-Ketten ermöglicht es, viele Prozesse zwecks Parallelität in einem insgesamt kleinen Raum auszuführen. Parallelität kennen wir auch von der Informationsverarbeitung im Gehirn und im Auge. Die Bildverarbeitung kann in verschiedenen Bereichen des Auges parallel ablaufen. Die 1994 von Leonard Adleman veröffentlichte Beschreibung eines Experiments, bei dem mit DNS-Molekülen gerechnet wurde, führte **zu großem Aufsehen in der Fachwelt**. Das Experiment benutzt DNS zur Ausführung eines Algorithmus, der mathematische Fragen aus der Graphentheorie löst.

Die im Experiment zu **beantwortende** Frage ist, ob es in einem vorgegebenen Graphen eine Kantenfolge gibt, die alle Knoten genau einmal erreicht. Interessanterweise wurde dieses Problem – natürlich ohne die technische Nutzung von Computern oder DNS – schon 1736 von dem berühmten Mathematiker Leonhard Euler behandelt, und zwar anhand der Brücken der Stadt Königsberg (heute Kaliningrad). Die Frage war, ob es einen Rundweg gibt, bei dem man alle sieben Brücken der Stadt über den Fluss Pregel genau einmal überquert. Euler bewies mit einer einfachen Argumentation, dass es keinen solchen Rundweg gab.

Die Lösung solcher Aufgaben erfordert bei großer Knotenanzahl viel Rechenzeit. Wenn man statt sieben Brücken (Kanten) eine größere Anzahl hat, steigt der Zeitaufwand rasend schnell. Durch hohen Parallelitätsgrad – wenn also sehr viel gleichzeitig ausgerechnet wird – lässt sich der Zeitbedarf stark senken. Im Falle des genannten Adleman-Experimentes wurde ein Graph mit sieben Knoten untersucht. Die Berechnung mit DNS dauerte mehrere Tage. Das ist natürlich im Vergleich zum Zeitbedarf eines Computers alles andere als beeindruckend. Das Wichtige und Aufsehen erregende war aber, dass eine Aufgabe dieser Art überhaupt mit DNS berechnet werden konnte.

Wir beschreiben den Aufbau und den Ablauf des Adleman'schen Verfahrens hier in groben Zügen.

Einzelheiten sind Adleman (1994) oder auch dem Lehrbuch von Hinze & Sturm (2004) zu entnehmen.

Jeder Knoten und jede Kante wird durch ein einzelnes DNS-Molekül dargestellt, und zwar durch einen ganz bestimmten DNS-Einzelstrang mit einer bestimmten Sequenz aus 20 Basen (20mer, lineare Abfolge aus 20 Basen). Die Zuordnung zwischen den Objekten im Graphen und den DNS-Einzelsträngen geschieht willkürlich, aber verbindlich. Ein bestimmtes 20mer repräsentiert in allen zusammengesetzten Strängen immer denselben Knoten oder dieselbe Kante. Im Experiment werden 7 Knoten und 14 Kanten verwendet. Die vorderen 10 Basen jeder Kantendarstellung sind komplementär zu den hinteren 10 einer Knotendarstellung, die hinteren 10 der Kantendarstellung sind komplementär zu den vorderen 10 einer anderen Knotendarstellung. Auf diese Weise lassen sich mittels der chemischen Paarungsgesetzmäßigkeiten (A immer mit T, C immer mit G) lange Doppelstränge bilden, in denen der eine Strang aus einer Abfolge von kodierten Kanten und der andere Strang aus einer Abfolge von kodierten Knoten besteht. Die Kantenfolge verbindet Knoten auf einem Pfad innerhalb des dargestellten Graphen. Meistens gibt es in einem einzelnen Graphen viele unterschiedliche Pfade.

Die Abbildung im Kastentext zeigt einen aus zwei Kanten und drei Knoten bestehenden Graphen. Der

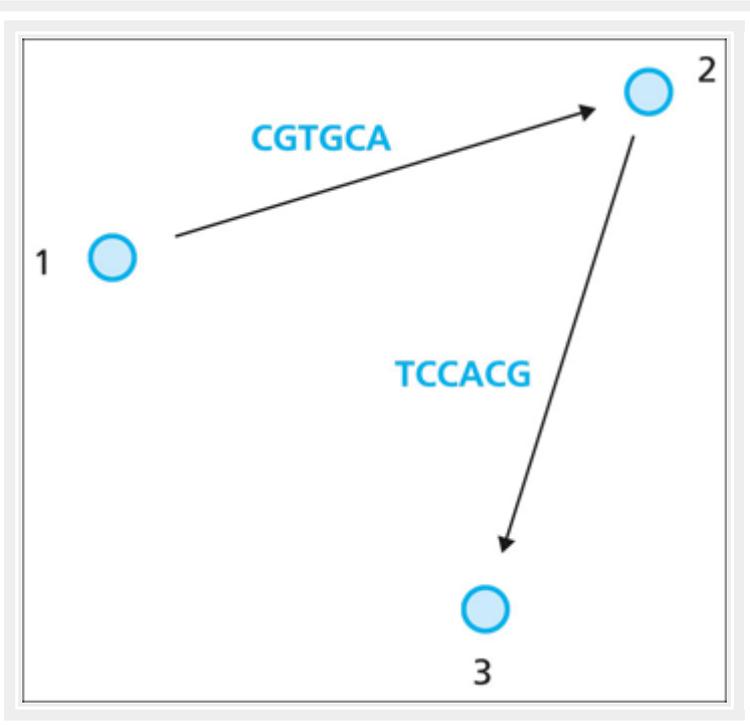
Grafik zu „Rechnen mit DNS“

einfacheren Darstellung halber sind die Stränge nicht 20, sondern nur 6 Basen lang.

Wir wollen den „Weg“ durch den aus drei Knoten 1, 2 und 3 und zwei Kanten (1, 2) und (2, 3) bestehenden Graphen (Abb. 2) in Form einer Nukleotidsequenz ausdrücken.

Zu diesem Zweck sei vorab vereinbart, dass die drei Knoten 1, 2 und 3 durch die Nukleotidketten ATCGCA, CGTAGG und TGCAGA und die beiden Kanten (1, 2) und (2, 3) durch CGTGCA und TCCACG dargestellt werden müssen, also in der Weise, dass die Darstellung jeder Kante gerade die Komplementärkette zur hinteren Hälfte ihres ersten und zur vorderen Hälfte ihres zweiten Knotens ist.

Dann drückt die folgende Verkettung der drei Knoten 1, 2, 3 im oberen Strang und der beiden Kanten (1, 2) und (2, 3) im dazu komplementären unteren Strang gerade aus, dass es einen Weg gibt, der von 1 über 2 nach 3 verläuft.



Knotenfolge

A T C G C A - C G T A G G - T G C A G A
 C G T G C A - T C C A C G

Kantenfolge

Die Bindestriche sind hier an solchen Stellen angegeben, wo Ausgangsketten auf der einen oder anderen Seite des Doppelstrangs zusammengefügt wurden.

Im Adleman-Experiment wurden zunächst Billionen kurzer Ketten für die Knoten und Kanten hergestellt; und es wurde diesen ermöglicht, sich in jeder beliebigen Weise aneinanderzufügen.

Die entstandenen Doppelstränge wurden danach mit molekularbiologischen Methoden daraufhin überprüft, ob sie mit einem bestimmten Teilstück anfangen, mit einem bestimmten anderen aufhörten und dazwischen jedes einen Knoten darstellende Teilstück genau einmal vorkam.

All diejenigen, auf die das zutrifft, sind Lösungen des zuvor beschriebenen mathematischen Problems.

Zunächst wird durch biochemische Synthese dafür gesorgt, dass alle Stränge billionenfach vervielfältigt vorkommen. Dann werden sie in ein gemeinsames Reagenzgefäß umgefüllt. Es wird davon ausgegangen, dass sich alle DNS-Stücke zusammenlagern, die komplementäre Überlappungen aufweisen.

Wahrscheinlichkeitsaussagen oder Untersuchungen darüber finden sich allerdings in der Originalarbeit nicht. Damit alle möglichen Stränge mit komplementären Überlappungen entstehen, muss die Menge der in der biochemischen Synthese verwendeten DNS-Stränge in exponentieller Größenordnung in der Menge der Knoten und Kanten des Problems sein. Als nächstes werden unter den nun vorhandenen Ketten nur die weiter verwendet, die mit dem gewünschten Knoten beginnen und mit dem anderen gewünschten Knoten enden.

Labortechnisch geschieht dies durch **DNS-Primerstränge*** mit bestimmten kurzen Basenfolgen, die eine

starke Vervielfältigung für alle mit dieser Basenfolge am Anfang übereinstimmenden Stränge bewirken. Hierbei kommt es außerdem auf die präzise Einstellung der Reaktionstemperatur an.

Dann sind die Stränge mit der richtigen Länge auszuwählen. In der hier beschriebenen Anordnung sind es diejenigen mit einer Länge von 140 Basenpaaren. Dies geschieht auf mechanischem Wege mit Hilfe einer **Gelelektrophorese***, die die vorhandenen **Stränge** nach deren „Länge“ „sortiert“^{tridi}. Stränge mit einem Längenunterschied von etwa 10 Basenpaaren lassen sich so voneinander trennen. Das ist ein Grund für die Wahl der Länge 20 bei den einzelnen Knoten und Kanten.

Nachdem nun diejenigen Stränge ausgewählt sind, die richtig anfangen und richtig aufhören und außerdem die richtige Länge haben, ist noch zu prüfen, ob jeder Knoten in den Strängen vorkommt. Dazu sind so viele weitere Separationen nötig, wie der Graph Knoten enthält. Jeder dieser Schritte folgt einem komplizierten Protokoll, in dem eine **Anlagerung** von Strängen an den zu einer Knotendarstellung komplementären Strang **vorgenommen wird**. Dabei wird der vorgegebene Strang **biotinyliert***, was die Möglichkeit der räumlichen Fixierung und anschließenden Trennung der ankoppelnden Stränge durch **Zentrifugieren** zur Folge hat. Sobald nur diejenigen Stränge übrig sein können, in denen jede Kantendarstellung vorkommt, ist zu fragen, ob tatsächlich solche Stränge übrig sind oder nicht. Dazu dient wieder eine Elektrophorese.

Auf mathematischem Wege konnte in den vergangenen Jahren bewiesen werden, dass jede überhaupt mögliche Berechnung (Turing-Berechenbarkeit [Asteroth & Baier 2002]) zumindest prinzipiell – wenngleich in vielen Fällen nur mit unvermeidbar hohem Aufwand an Material und Zeit – von DNS-Molekülen durchgeführt werden kann.

Glossar

Biotinylierung: Biotinylierung ist eine biochemische Methode, bei der z.B. eine makromolekulare Nukleinsäure wie DNS mit einem kleinen Molekül namens Biotin verbunden wird. Die Biotin-Gruppierung bindet sehr stark an bestimmte Proteine, die Avidine. Diese starke Bindung kann benutzt werden, Moleküle mit Biotin-Einheiten aus einem Gemisch „herauszufischen“. **Gelelektrophorese:** Die Gelelektrophorese ist ein biochemisches Trennverfahren, bei dem die Wanderung von geladenen Teilchen (Molekülen) in einem elektrischen Feld zu ihrer Trennung benutzt wird. Man löst das zu trennende Stoffgemisch und lässt die Lösung durch eine elektrisch neutrale, feste Gelmatrix aus Agarose oder Polyacrylamid wandern. Die Wanderungsgeschwindigkeit hängt von der Stärke des Feldes, der Nettoladung, Form und Größe der Makromoleküle sowie von physikochemischen Eigenschaften der Lösung ab. **Graphentheorie:** Die Graphentheorie beschäftigt sich mit Strukturen, die aus einzelnen Punkten (genannt Knoten, englisch: nodes) und Verbindungen zwischen diesen (genannt Kanten, englisch: edges) bestehen. Durch direkte Verbindungen zwischen einzelnen Knoten sind dann auch indirekte Verbindungen gegeben, die mehrere Kanten nacheinander durchlaufen, sowie Zyklen, die über mehrere Kanten zum Ausgangsknoten zurückführen. Ein Beispiel für den praktischen Nutzen der Graphentheorie ist die Herstellung und dynamische Änderung von Verbindungen im Telefonnetz oder Internet. Auch das „Problem des Handelsreisenden“ (Travelling Salesman Problem) ist Gegenstand der algorithmischen Graphentheorie. Es geht dabei um Verfahren, die auf einer Landkarte die insgesamt kürzeste Rundreise über alle ausgewählten Städte bestimmen. Mit geringen Änderungen tritt diese Frage auch in der industriellen Fertigung (Robotik) auf. Man kann beweisen, dass optimale Lösungen für große Graphen ebenso schwer zu finden sind wie bei dem von Adleman behandelten Graphenproblem. **Informationsverarbeitung:** Das Rechnen mit Zahlen, das wir schon in der Grundschule lernen, ist ein spezieller Fall der „Informationsverarbeitung“. Informationsverarbeitung sind alle schrittweisen, nach bestimmten Regeln erfolgenden Veränderungen von Zeichenfolgen. Zeichenfolgen können beispielsweise

aus Ziffern oder Buchstaben bestehen. Etwa im Falle der chinesischen Schriftsprache besteht der Zeichenvorrat, mit dem Folgen gebildet werden können, aus Tausenden unterschiedlicher Zeichen. Regeln können lauten: Ersetze Zeichen „1 x 1“ durch das Zeichen „2“, ersetze Zeichen „2 x 2“ durch das Zeichen „4“, und so weiter, wenn wir als Beispiel das Multiplizieren nehmen. Eine andere einfache Möglichkeit ist: Ersetze alle „a“ durch „b“, alle „b“ durch „c“, alle „c“ durch „d“ und so weiter. Bei diesem Verfahren handelt sich um eine schon von Gaius Julius Caesar verwendete **Verschlüsselungstechnik**. Aus informationsverarbeitenden Schritten aufgebaute Verfahren werden Algorithmen genannt. **Primer**: Primer sind kurze DNS-Sequenzen, die chemisch-synthetisch hergestellt und einem Reagenzgemisch zugesetzt werden, das DNS vervielfältigt. Sie dienen als Startsequenz für DNS-vervielfältigende Enzyme, die **DNS-Polymerasen**, die im Gemisch enthalten sind. DNS-Polymerasen können nur die Verlängerung eines schon existierenden kurzen DNS-Strangs **katalysieren**. Die DNS-Vervielfältigung verläuft wie bei der Zellteilung lebender Zellen.