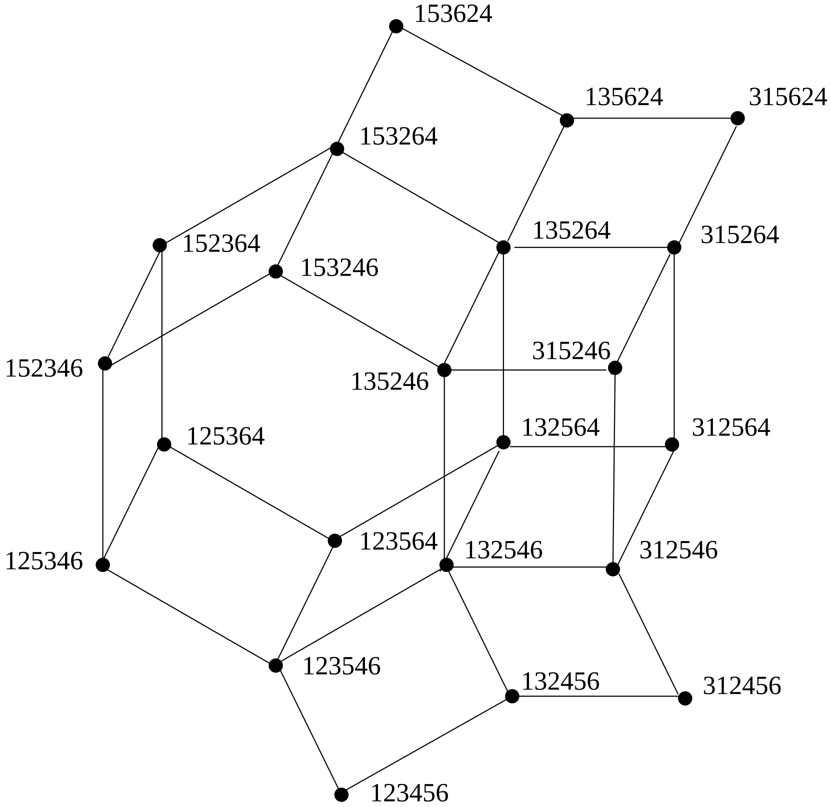
Vorlesung
Zufall und Algorithmen

Stefan Felsner

felsner@inf.fu-berlin.de

16. Februar 1999



Ausgearbeitet für die Fortsetzung des Schülerseminars 1998 in Altwriezen, Brandenburg, Februar 1999.

Inhalt

1 Zufall und Wahrscheinlichkeit

1.1 Einleitung

Die auf Erdős zurückgehende "Probabilistische Methode" in Kombinatorik und Graphentheorie sowie "Randomisierte Algorithmen" in der Theoretischen Informatik haben sich in den vergangenen zwei Jahrzehnten als äußerst wirkungsvolle Grundtechniken erwiesen, die oft zu eleganten und überraschenden Ergebnissen führen. In dieser Vorlesung möchte ich Beispiele für solche Anwendungen des Zufalls geben.

Zunächst müßen wir allerdings einige Begriffe klären. Da wir es fast ausschließlich mit diskreten Strukturen zu tun haben werden, begnügen wir uns mit einem eingeschränkten Wahrscheinlichkeitsbegriff.

1.2 Experimente, Wahrscheinlichkeit, Zufall

Was bedeuten die folgenden Sätze?

Morgen liegt die Niederschlagswahrscheinlichkeit bei 70%.

Wer beim Roulette auf rot setzt hat eine 18:19 Chance zu gewinnen.

Die Wahrscheinlichkeitstheorie bietet einen formalen Rahmen, in dem zumindest die zweite Aussage sinnvoll und richtig ist. Dabei gehen wir von einem Experiment aus. Ein *Ereignis* ist ein Ausgang des Experiments. Der *Ereignisraum* ist die Menge aller möglichen Ausgänge des Experiments (bei uns höchstens abzählbar viele).

Beispiel 1. Typische Experimente: Münze oder Würfel werfen, eine Kugel aus einer Urne ziehen, einen Ball mit verbundenen Augen in eines von n Kästchen werfen.

Angenommen wir führen ein Experiment immer wieder, also N mal, durch und bezeichnen einen Ausgang als Erfolg, wenn ein Ereignis ω auftritt. Es ist eine Erfahrungstatsache(?!), daß sich mit wachsendem N das Verhältnis der erfolgreichen Experimente zur Gesamtzahl immer mehr einer Zahl p zwischen 0 und 1 angleicht (gegen die Erfolgswahrscheinlichkeit p konvergiert).

In der Mathematik wird der Spieß umgekehrt. Wir definieren Wahrscheinlichkeit als eine Abbildung von den Ereignissen in die reellen Zahlen, die gewisse Eigenschaften erfüllt. Die angesprochene 'Erfahrungstatsache' findet man in der Theorie als "Gesetz der großen Zahlen" wieder.

Formal beginnen wir mit einem *Wahrscheinlichkeitsraum*, er besteht aus einem Ereignisraum Ω und einer *Wahrscheinlichkeitsverteilung* $\Pr : \Omega \to [0, 1]$, so daß gilt:

$$\sum_{\omega \in \Omega} \Pr(\omega) = 1.$$

Wir wollen nun auch Teilmengen $A \subseteq \Omega$ *Ereignisse* nennen. Durch $\Pr(A) := \sum_{\omega \in A} \Pr(\omega)$ ist die Wahrscheinlichkeit des Ereignisses festgelegt. Die Elemente $\omega \in \Omega$ heißen zur Unterscheidung auch *Elementarereignisse*.

Beispiel 2. Würfelergebnisse mit zwei gleichen Würfeln
$\Omega = \{\{i,j\} : 1 \leq i \leq j \leq 6\}$ und

$$\Pr(\{a,b\}) = \begin{cases} 1/36 & : \text{ falls } a = b \\ 1/18 & : \text{ sonst} \end{cases}$$

Oft wird in Beispielen und Aufgaben ein Experiment beschrieben, aber keine Wahrscheinlichkeitsverteilung auf den Ausgängen angegeben, Beispiel: 'wähle i zufällig aus $1, .., n$'. In solchen Fällen ist die Gleichverteilung gemeint. Die *Gleichverteilung* auf einer endlichen Menge Ω ist die Wahrscheinlichkeitsverteilung mit $\Pr(\omega) = 1/|\Omega|$ für jedes $\omega \in \Omega$.

Eine *Zufallsvariable* ist eine Funktion $X : \Omega \to \mathbb{R}$, die einem Elementarereignis eine Zahl zuordnet. Der *Erwartungswert* einer Zufallsvariablen ist definiert als

$$E(X) = \sum_{\omega \in \Omega} X(\omega)\Pr(\omega).$$

Ist Pr die Gleichverteilung, dann ist der Erwartungswert nichts anderes als der Mittelwert der Zahlen $X(\omega)$.

Ereignisse, die über Zufallsvariable definiert sind, werden oft abgekürzt notiert. Für $R \subseteq \mathbb{R}$ und $z \in \mathbb{R}$ bezeichnen $(X \in R)$ und $(X < z)$ die Mengen $\{\omega : X(\omega) \in R \subseteq \mathbb{R}\}$ und $\{\omega : X(\omega) < z\}$. In endlichen Ereignisräumen kann man den Erwartungswert damit auch so bestimmen:

$$E(X) = \sum_{x \in \mathbb{R}} x \cdot \Pr(X = x)$$

Beispiel 3. Sei $(\Omega, \Pr)$ wie in Beispiel 2. Die Zufallsvariable S sei die Summe der Augenzahlen. Es gilt $\Pr(S = 4) = \Pr(\{1,3\}) + \Pr(\{2,2\}) = 1/12$. Zeige: $E(S) = 7$.

Beispiel 4. Gegeben ein Experiment mit Erfolgswahrscheinlichkeit p. Wie oft muß das Experiment im Mittel wiederholt werden, um zu einem erfolgreichen Ausgang zu kommen? Sei X die Zufallsvariable, deren Erwartungswert wir bestimmen wollen, d.h. $X = k$, wenn Erfolg im k-ten Experiment eintritt. Wir setzen $q = 1 - p$.

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k) = \sum_{k=1}^{\infty} kq^{k-1}p = p\frac{d}{dq}(\sum_{k=0}^{\infty} q^k) = p\frac{d}{dq}(\frac{1}{1-q}) = \frac{p}{(1-q)^2} = \frac{1}{p}$$

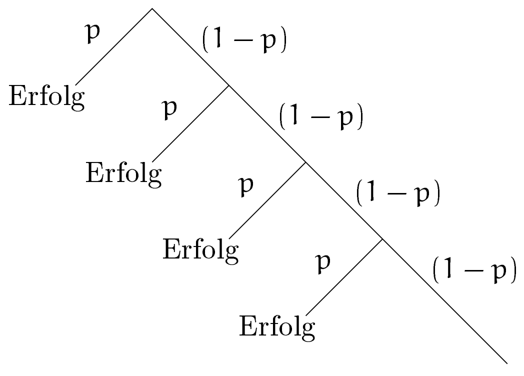


Abbildung 1. Der Baum zum Erfolg.

Der Erwartungswert hat die folgende einfache, aber wichtige Eigenschaft:

Linearität des Erwartungswertes.
Seien X_1, X_2 Zufallsvariablen auf Ω. Die Zufallsvariable $X_1 + X_2$ ist definiert durch $(X_1 + X_2)(\omega) = X_1(\omega) + X_2(\omega)$. Es gilt

$$E(X_1 + X_2) = E(X_1) + E(X_2).$$

Übung 1. Beweise die Linearität des Erwartungswertes.

Neben der Linearität des Erwartungswertes benötigt man häufig auch den
Erwartungswert von 0-1 Variablen.
Sei X eine Zufallsvariable die nur die Werte 0 und 1 annimmt, dann ist $E(X) = 0 \cdot \Pr(X = 0) + 1 \cdot \Pr(X = 1) = \Pr(X = 1)$.

Beispiel 5. Es werden n Bälle zufällig in n Kästchen geworfen. Wie viele leere Kästchen sind zu erwarten?

- Sei X_i eine Zufallsvariable mit $X_i = 0$ wenn Kiste i einen Ball enthält und $X_i = 1$ wenn Kiste i leer.
- Die erwartete Anzahl leerer Kästchen ist demnach $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$.
- $E(X_i) = \Pr(X_i = 1) = \left(\frac{n-1}{n}\right)^n \underset{n\to\infty}{\longrightarrow} \frac{1}{e}$.
- Die erwartete Anzahl leerer Kästchen ist also n/e oder annähernd ein Drittel.

Werden nun m Bälle in n Kästchen geworfen, dann ist die Wahrscheinlichkeit, daß Kästchen i leer bleibt $\left(\frac{n-1}{n}\right)^m = \left(\frac{n-1}{n}\right)^{n\cdot\frac{m}{n}} \underset{n\to\infty}{\longrightarrow} \frac{1}{e}^{\frac{m}{n}}$. Eine kleine Rechnung zeigt, daß die erwartete Anzahl leerer Kästchen für $m > n\ln(n)$ kleiner als eins wird. Damit haben wir eine Lösung für das wichtige

Coupon Collectors Problem: Es gibt n verschiedene Sammelbildchen, beim Kauf eines Bildchens bekommt man eines der n Motive gleichverteilt zugewiesen. Wie viele Bildchen muß ein Sammler kaufen der eine vollständige Sammlung haben will?

Hier eine zweite Lösung für das Problem: Bezeichne den j-ten Bildkauf als Erfolg, wenn das gekaufte Bildchen neu für den Sammler ist. Nun zerlegen wir die Folge der Zukäufe in Phasen, wobei die k-te Phase die Bildkäufe nach dem $(k-1)$-ten Erfolg und bis zum k-ten Erfolg enthält. Die Zufallsvariable Y_k stehe für die Anzahl der Bilder, die in der k-ten Phase gekauft werden. Jeder Bildkauf in der k-ten Phase ist mit Wahrscheinlichkeit $\frac{n-(k-1)}{n}$ ein Erfolg. Aus den Überlegungen in Beispiel 4 wissen wir also, daß $E(Y_i) = n/(n-k+1)$. Die erwartete Zahl von Bildkäufen ist daher

$$E(\sum_{k=1}^{n} Y_k) = \sum_{k=1}^{n} E(Y_k) = \sum_{k=1}^{n} \frac{n}{n-k+1} = n\sum_{k=1}^{n} \frac{1}{k}$$

Die Summe $H_n = \sum_{k=1}^{n} \frac{1}{k}$ ist als n-te *Harmonische Zahl* bekannt. Für diese Zahl hat man eine recht genaue Abschätzung: $\ln(n) < H_n < \ln(n) + 1$. Damit ist gezeigt, daß unser Sammler das Geld für $n\log(n)$ Bildchen bereithalten sollte.

Übung 2. Ein Experiment mit Erfolgswahrscheinlichkeit p wird n mal wiederholt. Sei X die Anzahl der Erfolge, berechne $E(X)$

Übung 3. Zufällig wird eine Permutation π von $\{1,..,n\}$ gezogen, wieviele Fixpunkte sind zu erwarten? Ein Fixpunkt von π ist ein i mit $\pi(i) = i$.

Übung 4. Auf der Knotenmenge $V = \{1,...,n\}$ wird ein zufälliger Graph erzeugt indem die zuerst leere Kantenmenge E wie folgt vergrössert wird:

- Wähle i und j zufällig (gleichverteilt) aus V. Wenn $i \neq j$ und $\{i,j\} \notin E$ dann ist $E := E \cup \{i,j\}$

Sei T die Anzahl der Iterationen bis der Graph zusammenhängend ist. Zeige, daß der Erwartungswert von T zwischen $\frac{1}{2}n\ln(n)$ und $n\ln(n)$ liegt.

Oft ist es von Interesse, wie weit eine Zufallsvariable von ihrem Erwartungswert abweicht. Eine einfache Antwort auf diese Art Frage liefert die

Markov Ungleichung.
Für eine Zufallsvariable X, die nur Werte ≥ 0 annimmt, und $\alpha > 0$ gilt

$$\Pr(X > \alpha\, E(X)) \leq 1/\alpha.$$

Abbildung 2. Weshalb die Markov Ungleichung gilt.

Folgerung: Die Wahrscheinlichkeit, daß bei n Würfen auf n Kästchen mehr als die Hälfte der Kästchen leer bleiben, ist kleiner als 3/4: Wir wählen in der Markov Ungleichung $\alpha = e/2$ und berechnen $1/\alpha = 2/e = 0.7357588$.

Beispiel 6. Ein Turnier ist eine Orientierung des vollständigen Graphen K_n. Wir wissen schon (VL Diskrete Mathe; Übung 9), daß es in jedem Turnier einen gerichteten Hamilton-Pfad gibt. Mit der Probabilistischen Methode werden wir nun zeigen, daß manche Turniere $\frac{n!}{2^{n-1}}$, also sehr viele, Hamilton-Pfade enthalten.

Sei T ein Turnier, das entsteht, indem die Orientierung für jede Kante des K_n durch einen Münzwurf festgelegt wird. Zu jeder Permutation σ der Knoten des K_n definieren wir eine Zufallsvariable

$$X_\sigma = \begin{cases} 1 & \text{wenn die Knotenfolge } \sigma \text{ ein HP von } T \text{ ist;} \\ 0 & \text{sonst.} \end{cases}$$

Die erwartete Anzahl von Hamilton-Pfaden in T ist

$$E(\sum_\sigma X_\sigma) = \sum_\sigma E(X_\sigma) = \sum_\sigma \Pr(X_\sigma = 1).$$

Die Wahrscheinlichkeit, daß eine Permutation σ einen Hamilton-Pfad von T beschreibt, ist $\left(\frac{1}{2}\right)^{n-1}$, da jede der $n-1$ Kanten auf dem Pfad richtig orientiert sein muß. Da es $n!$ Permutationen gibt, ist die erwartete Anzahl von Hamilton-Pfaden in einem zufälligen Turnier $\frac{n!}{2^{n-1}}$. Also gibt es Turniere, die zumindest soviele Hamilton-Pfade enthalten.

2 Permutationen und Sortieren

Eines der grundlegenden Probleme der Informatik ist das Sortierproblem. Man hat eine Folge $a_1, a_2, \ldots, a_n$ von Zahlen gegeben, mit denen in Anwendungen meist noch eine zusätzliche Information verknüpft ist – z.B. Löhne mit weiteren Personendaten. Die Aufgabe ist nun eine Umordnung $a_1', a_2', \ldots, a_n'$ der Daten zu bestimmen, die $a_1' \leq a_2' \leq \ldots \leq a_n'$ erfüllt. Der Aufwand für den Sortiervorgang wird meist über die Anzahl der benötigten Vergleiche von je zwei Elementen abgeschätzt. Es läßt sich zeigen, daß jeder vergleichsbasierte Sortieralgorithmus im schlechtesten Fall zumindest etwa $n \log(n)$ Vergleiche benötigt. Andererseits gibt es Algorithmen, die jede Eingabe mit etwa sovielen Vergleichen sortieren. Der folgende rekursive Algorithmus Quicksort ist nicht so gut. Wir bezeichnen die zu sortierende Menge mit A und nehmen zur Vereinfachung an, daß alle a_i verschieden sind.

QSort(A)

(1) Wähle ein Split-Element $a \in A$

(2) Vergleiche a mit jedem Element aus A und bestimme so die Mengen $A_< = \{x \in A : x < a\}$ und $A_> = \{x \in A : a < x\}$

(3) Rufe QSort($A_<$) und QSort($A_>$) auf, die Rückgaben seien $A_<'$ und $A_>'$

(4) Gib $(A_<', a, A_>')$ zurück

Mit einem Lauf von QSort kann man einen (gerichteten, gewurzelten, binären) Baum assoziieren.

Beispiel 7. Eingabe $(6, 4, 9, 5, 2, 7, 8, 3, 1)$

Nach der Zerlegung mit dem Split-Element 6 sind die Eingaben für den rekursiven Aufruf die Folgen $(4, 5, 2, 3, 1)$ und $(9, 7, 8)$. Wählt man weiterhin jeweils das erste Element als Split-Element, dann entsteht der in der Abbildung dargestellte Baum.

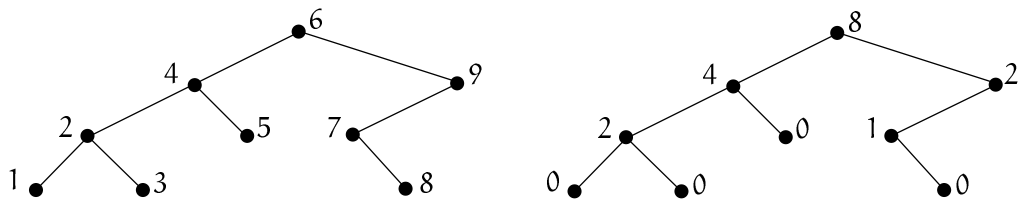


Abbildung 3. Ein Sortierbaum; links steht bei jedem Knoten das Split-Element, rechts die bei diesem Knoten durchgeführte Anzahl Vergleiche.

Übung 5. Zeige, daß QSort im schlimmsten Fall $\binom{n}{2}$ Vergleiche benötigt.

Die Übung deutet darauf hin, daß Quicksort kein so guter Algorithmus ist. Trotzdem ist Quicksort ein bei Praktikern sehr beliebtes und oft verwendetes Sortierverfahren. Um das zu verstehen, analysieren wir als nächstes eine randomisierte Variante.

2.1 Randomisiertes Quicksort

Die randomisierte Variante von Quicksort (RQSORT) unterscheidet sich von QSORT nur in der Wahl des Split-Element.

(1*) Wähle ein zufälliges $a \in A$ als Split-Element

Nun ist also jedes $x \in A$ mit Wahrscheinlichkeit $1/|A|$ das Split-Element. Die Hoffnung ist, daß oft mit einem Element aus der Mitte gesplittet wird und daher viele Vergleiche gespart werden.

Sei Z die Zufallsvariable, die die Anzahl der von RQSORT durchgeführten Vergleiche zählt. Wir interessieren uns für den Erwartungswert von Z. Um den zu bestimmen, stellen wir Z wieder einmal als Summe von einfacheren Zufallsvariablen dar. Mit

$$X_{ij} = \begin{cases} 1 & \text{wenn } i \text{ und } j \text{ verglichen werden,} \\ 0 & \text{sonst} \end{cases}$$

gilt $Z = \sum X_{ij}$ und also $E(Z) = \sum_{1 \le i < j \le n} \Pr(X_{ij} = 1)$. Nun müssen wir für gegebenes $i < j$ die Wahrscheinlichkeit, daß i und j verglichen werden, bestimmen.

Dazu ordnen wir die Elemente in der Reihenfolge, in der sie während des Laufs von RQSORT als Split-Elemente verwendet werden. Das liefert eine Permutation π_R.

($\star$) Ein Vergleich zwischen i und j findet genau dann statt, wenn kein Element aus dem offenen Intervall (i, j) vor i und j als Split-Element gewählt wird. Anders gesagt, i und j werden verglichen wenn eines der beiden Elemente in π_R vor allen anderen Elementen aus dem Intervall $[i, j]$ steht.

Die Wahrscheinlichkeit, daß das erste Split-Element, das aus dem Intervall $[i, j]$ gezogen wird, i oder j ist, ist genau $\frac{2}{j-i+1}$ (weil jedes Element mit gleicher Wahrscheinlichkeit das erste ist). Um $E(Z)$ zu bestimmen, müssen wir nur noch ein bißchen rechnen.

$$E(Z) = \sum_{1 \le i < j \le n} \Pr(X_{ij} = 1) = \sum_{1 \le i < j \le n} \frac{2}{j - i + 1} = 2 \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} \frac{1}{k+1} < 2n \sum_{k=1}^{n} \frac{1}{k}$$

Mit der beim Cupon Collector erwähnten Abschätzung der n-ten Harmonischen Zahl durch $\ln(n)$ erhält man

Satz 1. *Die randomisierte Version von Quicksort benötigt im Erwartungswert etwa* $2n \ln(n)$ *Vergleiche.*

Mit etwas stärkeren Techniken, als sie uns hier zur Verfügung stehen, kann man zeigen, daß die Verteilung "stark um den Erwartungswert konzentriert ist", d.h. es kommt sehr selten vor, daß der Algorithmus deutlich mehr als $2n \ln(n)$ Vergleiche benötigt.

3 Zufallsspaziergänge und Markov Ketten

3.1 Mischen

Das zum Sortieren inverse Problem wollen wir Mischen nennen. Hier ist eine Folge $a_1 \dots a_n$ gegeben; das Ziel ist es, die Elemente der Folge in eine zufällige Reihenfolge zu bringen. Idealerweise soll jede Permutation π der Elemente mit der gleichen Wahrscheinlichkeit

$1/n!$ auftreten. Natürlich gibt es hier eine einfache Lösung, die anhand eines Kartenstapels $A = (a_1 \ldots a_n)$ dargestellt werden soll.

PERMUTATION(A)
- (1) Reserviere n Plätze $P_1 \ldots P_n$
- (2) Wiederhole für $k = 1$ bis n
  - (2a) Wähle ein zufälliges $j \in \{1, \ldots, n-k+1\}$
  - (2b) Lege die oberste Karte von A auf den j-ten freien Platz
- (3) Bilde einen Stapel A', in dem die Karte von Platz P_i an i-ter Stelle liegt und gib A' zurück.

Wer regelmäßig Karten spielt, weiß, daß dieses Mischverfahren in der Praxis nicht viel Wert hat. Wir wollen nun ein abstraktes Modell von realistischeren Mischvorgängen beschreiben: Der Mischer nimmt den Stapel $A = A_0$ und führt eine Mischbewegung durch, nun hat er A_1 in der in der Hand. Weitere Mischbewegungen führen zu $A_2, A_3 \ldots$, bis der Mischer nach k Mischbewegungen meint, es ist genug und den Stapel A_k als gemischt zurückgibt. Wichtig für unsere Mischüberlegungen sind zwei Beobachtungen

(1) Mit einer einzelnen Mischbewegung kann der Stapel A nicht in jede andere Sortierung B überführt werden. Vielmehr handelt es sich bei einer Mischbewegung um einen Zug, der nur eine "kleine" Veränderung bewirkt.

(2) Wird so ein eingeschränkter Zug oft genug wiederholt, so führt das (hoffentlich) zu einem zufällig sortierten, also gemischten Stapel.

Wir wollen nun einen konkreten Mischvorgang betrachten. Ein Zug besteht hier aus der zufälligen Wahl einer Karte aus dem Stapel, diese Karte wird herausgenommen und oben auf den Stapel gelegt. Wir nennen die Methode RANDOM TO TOP. Die folgende Abbildung zeigt einen *Übergangsgraphen* für dieses Mischverfahren. Der Graph enthält eine Kante $\pi \to \pi'$, wenn π' in einem Zug von π aus erreicht werden kann, die Schleifen $\pi \to \pi$ sind nicht dargestellt.

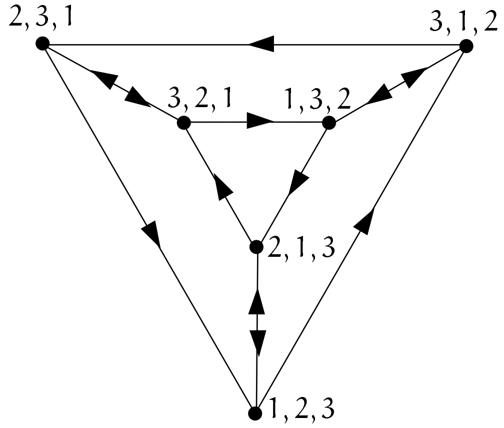


Abbildung 4. Der Übergangsgraph von RANDOM TO TOP auf 3-elementigen Permutationen.

Ein Mischvorgang mit RANDOM TO TOP kann als Zufallsspaziergang auf dem Übergangsgraphen aufgefasst werden. Wenn ein Zufallsspaziergänger im Knoten π angekommen ist, dann wählt er jede der ausgehenden Kanten mit der gleichen Wahrscheinlichkeit $1/n$ und gelangt so zu π'.

Denken wir uns als Graphen ein $n \times n$ Gitter und einen Zufallsspaziergänger, der in $(0,0)$ startet und in jedem Schritt jede der von seinem Knoten ausgehenden Kanten mit Wahrscheinlichkeit $1/n$ benutzt. Wieviele Schritte muß er gehen, bis er sich in einem zufälligen Gitterpunkt befindet? Offenbar zumindest $2n$, denn vorher kann er nicht in (n,n) ankommen. Vorher ist sein Aufenthaltsort noch von seinem Standpunkt abhängig und also nicht zufällig. Wir wollen nun untersuchen, nach wievielen RANDOM TO TOP Zügen wir erwarten können, daß die Abhängigkeit vom Startstapel verloren ist.

Dazu verändern wir die Regel ein bißchen: Wir wählen nun nicht die Position sondern die Nummer der Karte zufällig. Damit hat sich aber eigentlich nichts geändert! Nun folgendes Gedankenexperiment:

Wir legen neben den Stapel A noch einen beliebigen weiteren Stapel B. Dann führen wir jeden Zug mit derselben zufälligen Kartennummer an beiden Stapeln aus. Wenn nach k Schritten jede Kantennummer einmal gezogen wurde, gilt $A_k = B_k$. Da B beliebig war, ist gezeigt, daß die Abhängigkeit vom Startstapel verloren ist. Wie groß ist k? Da wir es hier mit einem Cupon Collection Problem zu tun haben, kennen wir die Antwort aus Beispiel 5:

Satz 2. *Um einen Stapel aus n Karten mit* RANDOM TO TOP *zu mischen, benötigt man im Erwartungswert etwa $n \ln(n)$ Züge.*

Übung 6. Sei A' der Stapel, der entstanden ist, nachdem RANDOM TO TOP jede Karte einmal umgelegt hat. Zeige, jede Permutation von A tritt mit gleicher Wahrscheinlichkeit als A' auf.

Die Methode mit der wir RANDOM TO TOP analysiert haben ist unter dem Namen *Coupling* bekannt. Eine Mischmethode die von Mischanfängern tatsächlich angewandt wird ist die Folgende: In jedem Zug wird die oberste Karte des Stapels genommen und an eine zufällige Stelle in den Stapel zurückgesteckt.

Klarerweise* zeigt diese Methode TOP TO RANDOM das gleiche Mischverhalten wie RANDOM TO TOP. In der folgenden Übung soll ein unabhängiger Beweis dafür gegeben werden:

Übung 7. Wir wenden TOP TO RANDOM auf einen Stapel A an. Sei n die unterste Karte des Stapels. Zeige:

(1) Zu jeder Zeit befinden sich die Karten unter n in einer zufälligen Reihenfolge
(2) TOP TO RANDOM erzeugt mit erwartet etwa $n \ln(n)$ Zügen eine zufällige Permutation. Hinweis: Zerlege den Prozeß in Phasen.

In der Literatur werden noch zahlreiche andere Mischverfahren untersucht. vom RIFFLE SHUFFEL bei dem A in zwei etwa gleichgroße Teile zerlegt wird, die dann zufällig ineinander geschoben werden, ist gezeigt, daß $\frac{3}{2}\log(n)$ Züge zum Mischen ausreichen. Für $n = 52$ sind also 8 bis 9 Züge nötig.

*oder auch nicht?

3.2 Theoretischer Hintergrund: Markov Ketten

Eine Markov Kette besteht aus einer Menge $V = \{v_1 \dots v_N\}$ von Zuständen und einer Übergangswahrscheinlichkeit p_{ij} für jedes Paar $1 \leq i, j \leq N$, die angibt, mit welcher Wahrscheinlichkeit man vom Zustand v_i in einem Schritt in den Zustand v_j kommt. Für jedes i muß gelten $\sum_{j=1}^{N} p_{ij} = 1$. Man faßt die Wahrscheinlichkeiten p_{ij} in einer Matrix $M = (p_{ij})_{1 \leq i,j \leq N}$ zusammen.

Zu einer Markov Kette kann man den Übergangsgraphen G_π betrachten, indem eine gerichtete Kante $v_i \to v_j$ für die i, j mit $p_{ij} > 0$ enthalten ist. Ist $q : V \to [0, 1]$ mit $\sum_{i=1}^{N} q_i = 1$ ein Vektor von Aufenthaltswahrscheinlichkeiten. Dann ist $q' = qM$ die Aufenthaltswahrscheinlichkeit nach einem Schritt des Markov Prozesses. Aus einer Startverteilung $q^{(0)}$ ergibt sich mit $q^{i+1} = q^{(i)}M$ eine Folge von Verteilungen. Ist $q^{(0)}$ von der Gestalt $(0 \dots 0, 1, 0 \dots 0)$ mit der 1 an der i-ten Stelle, dann beschreibt $q^{(k)}$ die Aufenthaltswahrscheinlichkeiten eines Zufallsspaziergängers, der in G_M im Knoten v_i gestartet ist nach k Schritten. Der folgende Schritt wird als der Hauptsatz über Markov Ketten bezeichnet.

Satz 3. *Ist in G_M jeder Knoten von jedem anderen aus über einen gerichteten Pfad erreichbar und gibt es ein i mit $p_{ii} > 0$, dann gibt es eine eindeutige Verteilung ρ mit $\rho_i = \sum_j \rho_j p_{i,j}$ für jedes i, oder kurz $\rho = \rho M$. Diese Verteilung ρ heißt die* stationäre Verteilung. *Ausserdem gilt für jede Anfangsverteilung $q^{(0)}$, daß $q^{(k)}$ sich mit der Zeit der stationären Verteilung annähert, $\rho_i = \lim_{k \to \infty} q_i^{(k)}$ für jedes i. In unseren Anwendungen geht es meistens darum, die Mischgewschwindigkeit zu untersuchen, d.h. zu analysieren, wie groß man k wählen muß, um nahe an die Verteilung ρ zu kommen.*

Übung 8. Sei $G = (V, E)$ ein ungerichteter zusammenhängender und nicht bipartiter Graph. Betrachte die Markov Kette mit

$$p_{ij} = \begin{cases} 1/d(i) & \text{wenn } (i,j) \in E \\ 0 & \text{sonst.} \end{cases}$$

wobei $d(i)$ der Grad von v_i ist. Zeige: Die eindeutige stationäre Verteilung (der Hauptsatz gilt in diesem Falle obwohl $p_{ii} = 0$ für alle i) ist gegeben durch $\rho_i = \frac{d(i)}{2|E|}$.

Übung 9. Zeige, daß der Hauptsatz nicht gilt, wenn man die Voraussetzung "nicht bipartit" fallenläßt.

Übung 10. Eine Maschine arbeitet den ganzen Tag und geht mit Wahrscheinlichkeit p in der Nacht kaputt. In diesem Fall wird sie in 2 Tagen repariert. Die Reparatur ist mit Wahrscheinlichkeit p erfolgreich, andernfalls ist ein neuer Reparaturversuch (weitere 2 Tage) erforderlich. In welchem Prozentsatz der Tage arbeitet die Maschine?

3.3 Anwendung: Das Problem des Handlungsreisenden

Für Definitionen bezüglich des Problems siehe (VL Diskrete Mathematik), dort haben wir gesehen, daß es sich beim TSP um ein notorisch schwieriges Optimierungsproblem handelt. Ein Lösungsansatz, der in vielen Fällen zu guten Ergebnissen, fast optimalen Touren führt, soll im folgenden beschrieben werden.

Für jede Tour T definiert man eine Menge von benachbarten Touren. Eine typische solche Definition ist die 2-Opt Nachbarschaft, bei der zwei Touren benachbart sind, wenn sie durch den Austausch von zwei Kanten auseinander hervorgehen; siehe Abbildung. Nun wertet man die Lösung der zu T benachbarten Touren aus, für das weitere Vorgehen hat man verschiedene Möglichkeiten

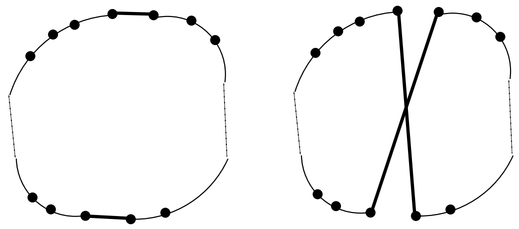


Abbildung 5. Zwei Touren die 2-opt benachbart sind.

(1) Wähle die kürzeste der benachbarten Touren als nächsten Zustand oder stoppe wenn die Nachbartouren alle länger sind als T.

(2) Wähle eine der benachbarten Touren zufällig nach einer Übergangsverteilung, die abhängig von den Kosten ist.

Das 1. Verfahren ist eine *lokale Optimierung* und interessiert hier nicht weiter. Das zweite Verfahren hat viele subtile Varianten, z.B. *simuliertes Abkühlen.* Bemerkenswert ist in unserem Zusammenhang, daß es sich um einen Markov Prozess handelt, den man anwenden kann, ohne den gesamten Übergangsgraphen bzw. die Übergangsmatrix zu kennen. Diese Objekte wären nämlich viel zu groß.

3.4 Anwendung: Domino Tilings (Domino Parkettierungen)

IN der statistischen Physik untersucht man gewisse Eigenschaften von sogenannten *Diner Systemen.* Dabei handelt es sich um Domino Tilings einer Grundfläche, z.B. eines $n \times n$ Brettes. Aus gewissen Gründen möchte man solche Tilings zufällig, jedes mit der gleichen Wahrscheinlichkeit, erzeugen. Die Methode der Wahl ist auch hier ein Markov Prozess, man bezeichnet zwei Tilings als benachbart, wenn sie durch eine elementare Rotation, siehe Abbildung, auseinander hervorgehen. Ist ein Tiling D gegeben, dann wählt man ein zufälliges 2×2 Subquadrat, enthält dieses Subquadrat genau zwei parallele Dominos, dann notiert man diese, andernfalls bleibt man bei D. Aus dem Hauptsatz folgt, daß dieser Prozess gegen die Gleichverteilung strebt.

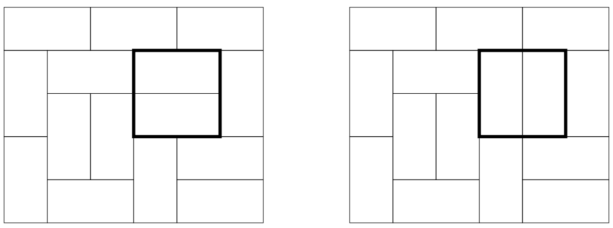


Abbildung 6. Eine Domino-Rotation

Physiker haben die Methode auch seit langem benutzt. Erst 1995 jedoch konnte gezeigt werden, daß man nach $64n^4 \log_2(1/\epsilon)$ Schritten ϵ-nahe an der angestrebten Verteilung ist.

Zusätzliche Übungsaufgaben

Übung 11. A und B werfen je n mal eine Münze. Zeige, daß die Wahrscheinlichkeit dafür, daß beide gleich oft Kopf geworfen haben, genau $\binom{2n}{n}\frac{1}{2^{2n}}$ ist. (Der schönste Beweis kommt ohne Rechnung aus.)

Übung 12. A hat eine Münze, die mit Wahrscheinlichkeit p Kopf und mit $(1-p)$ Zahl zeigt. B bittet A um Zufallsbits und fordert, daß 0 und 1 je mit Wahrscheinlichkeit $1/2$ auftreten. Wie kann A die Münze nutzen, um die Bitte zu erfüllen? Hinweis: Betrachte Paare von Würfen.

Übung 13. (Spielers Pleite) Ein Mensch möchte einen Jaguar für N Mark kaufen, er hat schon k Mark und versucht den fehlenden Rest durch ein Münzspiel zu gewinnen bei dem er in jeder Runde je mit Wahrscheinlichkeit $1/2$ eine Mark gewinnt oder verliert. Er spielt bis er mit 0 Mark pleite ist oder den Jaguar kaufen kann. Wie groß ist die Wahrscheinlichkeit für die Pleite?

Übung 14. Dieser Spieler spielt das gleiche Spiel wie der vorherige, allerdings hat er einen reichen Onkel der ihn jedesmal wenn er auf 0 ist mit einer Mark weiterhilft. Am Ende wird dieser Spieler den Jaguar kaufen, wie lange muss er dafür im Mittel spielen?

Übung 15. Ein Zufallsspaziergänger auf einem Kreis mit Knoten $\{0, 1, ..., n\}$ startet in 0 und geht bis er jeden Knoten besucht hat, im letzten neu besuchten Knoten bleibt er stehen. Wo beendet es seinen Weg mit welcher Wahrscheinlichkeit?

Übung 16. Ähnlich wie bei den Spieler-Aufgaben betrachten wir einen Zufallsspaziergang auf dem Intervall $\{1, ..., n\}$. Das heißt für $i \neq 1, n$ geht man je mit Wahrscheinlichkeit $1/2$ zu $i-1$ oder $i+1$, in den beiden Randknoten bleibt man je mit Wahrscheinlichkeit $1/2$ stehen.

(1) Zeige: Die stationäre Verteilung ist die Gleichverteilung.

(2) Verwende ein Coupling-Argument um abzuschätzen wie viele Schritte ein Zufallsspaziergänger im Mittel benötigt bis sein Aufenthaltsort zufällig ist.

Erläuterung zum Titelbild

Sei G ein gerichteter Graph auf n Knoten, der keinen gerichteten Kreis enthält. Eine topologische Sortierung von G ist eine Permutation, π der Knoten von G so daß für jede Kante $i \to j$ von G gilt, i kommt in π vor j, d.h. alle Kanten zeigen nach rechts. Ein Markov Prozess auf den topologischen Sortierungen läuft wie folgt:

Wähle eine Zahl i zufällig aus $\{1, \ldots, n-1\}$. Nun betrachte die Knoten $\pi(i), \pi(i+1)$ in der aktuellen topologischen Sortierung. Wenn $\pi(i) \to \pi(i+1)$ eine Kante von G ist, dann bleibe bei π, andernfalls gehe durch Vertauschen von $\pi(i)$ und $\pi(i+1)$ zu π' über.

Das Titelbild zeigt den Übergangsgraphen dieser Markov Kette für den gerichteten Graphen G aus der Abbildung.

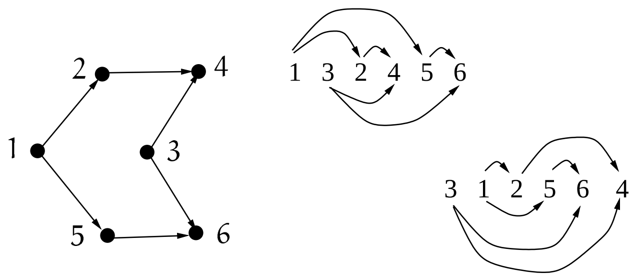


Abbildung 7. Ein Graph mit zwei topologischen Sortierungen.