Elektron. Informationsverarb. u. Kybernet. 21 (1985) 1/2, 3–7
(Journal of Information Processing and Cybernetics – EIK)

Einige Bemerkungen zu einer Arbeit von *W. Knödel* über das mittlere Verhalten von on-line-Packungsalgorithmen

Von *Helmut Prodinger*

H. Prodinger: Some remarks on a paper by *W. Knödel* on the mean behaviour of online bin packing algorithms.

Abstract: For the bin packing algorithms best fit and first fit *W. Knödel* has considered the average density of packing assuming a special stochastic input sequence. Essentially it is described asymptotically by $8c\, e^{-1/6} \sqrt{n/\pi}/9$ (c a numerical constant). In the present note the value $c = e^{1/6} \sqrt{3}/2$ is determined. By means of the trinomial coefficients the average can be computed exactly.

In einer jüngst erschienenen Arbeit [4] hat *W. Knödel* eine asymptotische Formel betreffend das mittlere Verhalten von Packungsalgorithmen unter speziellen statistischen Annahmen hergeleitet. Aus Gründen der Kürze wollen wir hier die Problemstellung und einige Rechnungen aus [4] nicht wiederholen und setzen beim Leser die Kenntnis von [4] voraus.

Knödel zeigte ([4]; S. 429, (4); Notation leicht geändert):

$$E_n \sim \frac{8c}{9\, e^{1/6}} \sqrt{\frac{n}{\pi}}. \tag{1}$$

Die Bestimmung der Konstanten c gelingt nur numerisch: $c = 1{,}02\ldots$. In dieser Note wird folgender Satz gezeigt:

Satz.

$$E_n = 3^{-n-1} \left((2n+1) \binom{n,\,3}{n} + (2n+1) \binom{n-1,\,3}{n-1} + \binom{n-1,\,3}{n-2} - 3^n \right), \tag{2}$$

$$E_n = \frac{4}{3\sqrt{3}} \sqrt{\frac{n}{\pi}} - \frac{1}{3} + O\left(\frac{1}{\sqrt{n}}\right), \tag{3}$$

$$c = \frac{\sqrt{3}}{2}\, e^{1/6} = 1{,}0230881\ldots. \tag{4}$$

Hier bezeichnet $\binom{n,\,3}{k}$ den Trinomialkoeffizienten $[x^k]\,(1 + x + x^2)^n$ (vgl. [3; S. 77]). ($[x^k]\, f(x)$ soll stets den Koeffizienten von x^k in $f(x)$ bedeuten.)

Beweis. Formel (3) aus [4] besagt im wesentlichen folgendes:

$$E_n = 3^{-n-1}[x^n]\left(T_{-1}(x) + 2\sum_{k\geq 1} kT_k(x)\right); \tag{5}$$

die Erläuterung der auftretenden erzeugenden Funktionen gelingt in anschaulicher Weise etwa entsprechend Abb. 1. In diesem Übergangsdiagramm denke man sich jede Kante mit dem Symbol x markiert. Dann sei $T_{-1}(x)$ die erzeugende Funktion der Wege von Z_0 nach Z_1, $T_0(x)$ die der Wege von Z_0 nach Z_0 und $T_k(x)$ für $k \geq 1$ die der Wege von Z_0 nach Z_{2k}. Die Herleitung von expliziten Formeln gelingt leicht, wenn man mit der Automatentheorie vertraut ist.

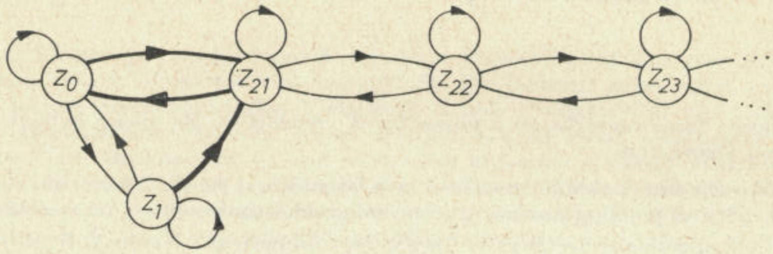

Abb. 1

Als Hilfsüberlegung denke man sich zunächst Z_0 und Z_1 eliminiert; $M(x)$ bezeichne die erzeugende Funktion der Wege, die in Z_{21} starten. Eine kanonische Faktorisierung des Weges bezüglich der Kanten von Z_{21} nach Z_{22} bzw. Z_{22} nach Z_{21} liefert

$$M(x) = \sum_{i\geq 0}\left(\frac{x}{1-x}\cdot xM(x)\right)^i \cdot \frac{1}{1-x} = \frac{1}{1-x-x^2M(x)},$$

und daher

$$M(x) = \frac{1-x-\sqrt{1-2x-3x^2}}{2x^2}.$$

Um T_0 zu bestimmen, faktorisiert man die Wege nach den stark gezeichneten Kanten. Die Wege, die erstmalig Z_{21} erreichen, werden durch $x/(1-2x)$ beschrieben, die, von Z_{21} startend, nach Z_0 führen, durch $xM(x)$. Schließlich kann man noch „am Ende" von Z_0 nach Z_0 fahren, und nur Z_1 als Zwischenknoten verwenden. Daher erhält man:

$$T_0(x) = \sum_{i\geq 0}\left(\frac{x}{1-2x}\cdot xM(x)\right)^i \cdot \sum_{j\geq 0}\left(\frac{x^2}{(1-x)^2}\right)^j \cdot \frac{1}{1-x} = \frac{1-x}{1-2x-x^2M(x)}.$$

$T_{-1}(x)$ ist nun aus $T_0(x)$ durch Multiplikation mit $x/(1-x)$ zu gewinnen:

$$T_{-1}(x) = \frac{x}{1-2x-x^2M(x)}.$$

Um $T_i(x)$ für $i \geq 1$ zu bestimmen, faktorisiert man die Wege: Man kommt schließlich nach Z_{21} und benützt dann Z_0 und Z_1 nicht mehr; der Beitrag ist

$$\left(T_0(x) + T_1(x)\right)x.$$

Sodann faktorisiert man, nachdem man Z_{2j} erreicht hat und nie mehr Z_{2s} mit $s < j$ benützt. Der Beitrag ist

$$x^{i-1}\left(M(x)\right)^i.$$

Insgesamt ergibt sich

$$T_i(x) = \frac{x^i(M(x))^i}{1 - 2x - x^2M(x)}.$$

Eine leichte Rechnung ergibt als Kontrolle

$$\sum_{i \geqq -1} T_i(x) = \frac{1}{1 - 3x},$$

wie auch zu erwarten ist. Weiters gilt

$$\sum_{k \geqq 1} kT_k(x) = \frac{\partial}{\partial y} \sum_{k \geqq 1} T_k(x)\, y^k|_{y=1} = \frac{1}{1 - 2x - x^2M(x)} \frac{xM(x)}{(1 - xM(x))^2}.$$

Um die Koeffizienten in (5) tatsächlich ablesen zu können, bedienen wir uns eines Kunstgriffs und der Cauchyschen Integralformel; wir verweisen auf [2, 5].

Sei $x = v/(1 + v + v^2)$, dann ist $M(x) = 1 + v + v^2$ und $T_{-1}(x) = v/(1 - v)$. Daher ist

$$\begin{aligned}[x^n]\, T_{-1}(x) &= \frac{1}{2\pi i} \int^{(0_+)} \frac{dx}{x^{n+1}}\, T_{-1}(x) = \frac{1}{2\pi i} \int^{(0_+)} \frac{(1 - v^2)\, dv}{v^{n+1}}\, (1 + v + v^2)^{n-1} \cdot \frac{v}{1 - v} \\ &= [v^{n-1}]\, (1 + v)\, (1 + v + v^2)^{n-1} = \binom{n-1,\, 3}{n-1} + \binom{n-1,\, 3}{n-2}. \qquad (6)\end{aligned}$$

In ähnlicher Weise erhält man

$$\begin{aligned}&[x^n] \frac{1}{1 - 2x - x^2M(x)} \cdot \frac{xM(x)}{(1 - xM(x))^2} = \frac{1}{2\pi i} \int^{(0_+)} \frac{dv(1 + v)\, (1 + v + v^2)^n}{v^n(1 - v)^2} \\ &= [v^{n-1}]\, (1 + v)\, (1 + v + v^2)^n \sum_{i \geqq 0} (i + 1)\, v^i \\ &= \sum_{i \geqq 0} (i + 1) \left(\binom{n,\, 3}{n - 1 - i} + \binom{n,\, 3}{n - 2 - i} \right) \\ &= \sum_{k=0}^{n-1} (n - k) \binom{n,\, 3}{k} + \sum_{k=0}^{n-1} (n - 1 - k) \binom{n,\, 3}{k} \\ &= (2n - 1) \sum_{k=0}^{n-1} \binom{n,\, 3}{k} - 2 \sum_{k=0}^{n-1} k \binom{n,\, 3}{n}. \qquad (7)\end{aligned}$$

Die erste Summe ist aus Symmetriegründen leicht zu bestimmen:

$$\sum_{k=0}^{n-1} \binom{n,\, 3}{k} = \frac{1}{2} \left(3^n - \binom{n,\, 3}{n} \right).$$

Die zweite Summe ist auch nicht viel schwieriger zu bearbeiten:

$$\begin{aligned}\sum_{k=0}^{n-1} k \binom{n,\, 3}{k} &= -n \binom{n,\, 3}{n} + [x^n] \frac{1}{1 - x} \sum_{k \geqq 0} k \binom{n,\, 3}{k} x^k \\ &= -n \binom{n,\, 3}{n} + [x^n] \frac{x}{1 - x}\, n(1 + x + x^2)^{n-1}\, (1 + 2x)\end{aligned}$$

$$= -n\binom{n,3}{n} + n\sum_{i\geq 0}\left(\binom{n-1,3}{n-1-i} + 2\binom{n-1,3}{n-2-i}\right)$$
$$= -n\binom{n,3}{n} - 2n\binom{n-1,3}{n-1} + 3n\sum_{i=0}^{n-1}\binom{n-1,3}{i}$$
$$= -n\binom{n,3}{n} - 2n\binom{n-1,3}{n-1} + 3n\cdot\frac{1}{2}\left(3^{n-1} + \binom{n-1,3}{n-1}\right)$$
$$= \frac{n}{2}3^n - \frac{n}{2}\binom{n-1,3}{n-1} - n\binom{n,3}{n}.$$

Damit erhält man für (7)

$$(2n-1)\frac{1}{2}\left(3^n - \binom{n,3}{n}\right) - n3^n + n\binom{n-1,3}{n-1} + 2n\binom{n,3}{n}$$
$$= -\frac{1}{2}3^n + \left(n+\frac{1}{2}\right)\binom{n,3}{n} + n\binom{n-1,3}{n-1}.$$

Schließlich erhält man aus (5)

$$3^{n+1}E_n = \binom{n-1,3}{n-1} + \binom{n-1,3}{n-2} - 3^n + (2n+1)\binom{n,3}{n} + 2n\binom{n-1,3}{n-1},$$

und das ist Formel (2).

Nun gilt [5] für $n \to \infty$ und $k = \mathcal{O}(1)$:

$$3^{-n}\binom{n,3}{n-k} = \sqrt{\frac{3}{4\pi n}}\cdot\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

woraus sich (3) durch Einsetzen ergibt. (4) ist trivial.

Bemerkung. Wäre man nicht an der expliziten Formel (2), sondern nur an der asymptotischen Formel (3) interessiert, könnte man mit Hilfe der Darbouxschen Methode (vgl. [1]) die Rechnung viel kürzer gestalten. Allerdings scheint gerade die Existenz einer expliziten Formel einigermaßen verblüffend zu sein, so daß darauf nicht verzichtet wurde. Die asymptotische Formel könnte man mit leichter Mühe genauer gestalten.

Literatur

[1] *Bender, E. A.*, Asymptotic Methods in Enumeration. SIAM Rev. **16** (1974), 485–515.
[2] *de Bruijn, N. G., D. E. Knuth, S. O. Rice*, The average height of planted plane trees. In: Graph Theory and Computing; Ed.: *R. C. Read*; Academic Press, 1972; pp. 15–22.
[3] *Comtet, L.*, Advanced Combinatorics. Reidel Publ. Comp., Dordrecht—Boston 1974.
[4] *Knödel, W.*, Über das mittlere Verhalten von on-line-Packungsalgorithmen. Elektron. Informationsverarb. u. Kybernet. EIK **19** (1983) 9, 427–433.
[5] *Prodinger, H.*, The average height of a stack where three operations are allowed and some related problems. J. Combin. Inform. System Sci. **5** (1980), 287–304.

Kurzfassung

W. Knödel hat für die Packungsalgorithmen best fit und first fit unter Annahme einer speziellen stochastischen Eingabefolge den Mittelwert der Packungsdichte betrachtet. Diese wird im wesentlichen durch $8c\,e^{-1/6}\sqrt{n/\pi}/9$ asymptotisch beschrieben (c eine numerische Konstante). In dieser Note wird der Wert $c = e^{1/6}\sqrt{3}/2$ bestimmt. Mit Hilfe der Trinomialkoeffizienten kann der Mittelwert sogar explizit angegeben werden.

Резюме

В. Кнёдель изучал среднее значение плотности упаковки для best fit- и first fit-алгоритмов в предположении специальной случайной входной последовательности. Эта плотность асимптотически равна $8c\,e^{-1/6}\sqrt{n/\pi}/9$ (c — некоторая константа). В настоящей заметке найдено значение $c = e^{1/6}\sqrt{3}/2$. Среднее значение плотности может быть найдено в явном виде с помощью тригономинальных коэффициентов.

(Eingegangen am 28. 12. 1983)

Anschrift des Verfassers:
H. Prodinger
Institut für Algebra und Diskrete Mathematik,
Abteilung für Theoretische Informatik
Technische Universität Wien
Gußhausstraße 27—29
1040 Wien
Österreich

Elektron. Informationsverarb. u. Kybernet. 21 (1985) 1/2, p. 8
(Journal of Information Processing and Cybernetics — EIK)

Book Review

Leben und Werk von John von Neumann. Ein zusammenfassender Überblick. Herausgegeben von T. Legendi und T. Szentivanyi. (Aus dem Ungarischen übersetzt von R. Nienhaus). Bibliographisches Institut, Mannheim-Wien-Zürich 1983. 151 S.

Im Buch sind Beiträge verschiedener Autoren zur Persönlichkeit und wissenschaftlichen Bedeutung von *John v. Neumann* zusammengefaßt. Die Beiträge geben einen guten Überblick seiner höchst beeindruckenden Vielseitigkeit und Produktivität auf vielen unterschiedlichen Wissenschaftsgebieten. Dabei werden sowohl seine Leistungen auf dem engeren Gebiet der Mathematik — und auch dort auf einer Vielzahl unterschiedlicher Teilgebiete — als auch davon ausstrahlend, auf weitere Wissenschaftsdisziplinen, insbesondere Physik, Ökonomie und Kybernetik, dargestellt. Er hat dabei fundamentale Problemstellungen dieser Wissenschaftsdisziplinen erschlossen und geklärt. Seine Ergebnisse haben die Gesamtentwicklung dieser Disziplinen maßgeblich mit beeinflußt und sind — trotz des hohen Entwicklungstempos dieser Gebiete — immer noch Bezugspunkt höchst aktueller Forschungsprobleme.

Das Erscheinen des Buches ist sehr zu begrüßen, da damit, zumindest im deutschsprachigen Raum, eine bestehende Lücke zur Würdigung einer herausragenden Wissenschaftspersönlichkeit dieses Jahrhunderts geschlossen wird. Das Buch ist sowohl als Lektüre für Fachwissenschaftler als auch für den interessierten Laien geeignet.

Die umfangreiche Bibliographie ermöglicht dem Leser die selbständige weitere Beschäftigung mit dieser Persönlichkeit. Der Stil der Beiträge ist nicht immer überzeugend, und die Gliederung des Druckbildes mit einer sehr ausgeprägten absatzweisen Gliederung steht nicht immer in Einklang mit der dargestellten Gedankenfolge. Diese kleineren Mängel mindern aber nicht die gute Qualität und den Wert des Buches. G. Wolf