Mathematisches Forschungsinstitut Oberwolfach

Report No. 41/2004

Mini-Workshop: Probability Theory on Trees and Analysis of Algorithms

Organised by
Gerold Alsmeyer (Münster)
Luc Devroye (Montreal)
Uwe Rösler (Kiel)

August 15th–August 21st, 2004

Introduction by the Organisers

The analysis of algorithms and data structures, started by D. Knuth, is a rapidly growing area at the interface of Mathematics and Theoretical Computer Science. Probability theory enters the subject in a natural way when studying an algorithm as to its performance over randomized inputs and/or if an algorithm itself takes randomization steps. The latter holds true for so-called divide and conquer algorithms, which were a central topic of this workshop. The sorting algorithm Quicksort is presumably the most prominent example. While early work was mostly based on the generating function approach, the last two decades have seen an increasing number of contributions based on probabilistic methods involving martingales, random trees and other stochastic processes. The first article of this type was written in the eighties by L. Devroye, a Humboldt award winner of 2004 and also one of the organizers. Later another pioneering contribution came by U. Rösler who, by introducing weighted branching processes and the contraction method, determined the asymptotic distributional behavior of Quicksort. One can say that much of the work presented at this workshop is to some extent spawn by these articles.

The aim and scope of the mini-workshop was to bring together leading junior and senior experts in the field with a strong probabilistic background and a focus on divide an conquer algorithms and related data structures. There were 16 one hour talks presented by 12 of the 14 participants from 7 countries. Main topics were

branching processes, the contraction method, the asymptotic analysis of random trees, stochastic fixed point equations and randomized algorithms.

Mini-Workshop: Probability Theory on Trees and Analysis of Algorithms

Table of Contents

Abstracts

Random records and cuttings in trees

Svante Janson

We consider random cutting down of rooted trees, defined as follows [6]. If T is a rooted tree with number of vertices $|T| \geq 2$, we make a random cut by choosing one edge at random. Delete this edge so that the tree separates into two parts, and keep only the part containing the root. Continue recursively until only the root is left. We let $X(T)$ denote the (random) number of cuts that are performed until the tree is gone.

The same random variable appears when we consider records in a tree. Let each edge e have a random value λ_e attached to it, and assume that these values are i.i.d. with a continuous distribution. Say that a value λ_e is a *record* if it is the largest value in the path from the root to e. Then the number of records is again given by $X(T)$, as is easily seen.

There are also vertex versions of cuttings and records. For cuttings, choose a vertex at random and destroy it together with all its descendants. Continue until the root is chosen and thus the whole tree is destroyed. For records, we assign i.i.d. values λ_v to the vertices, and define a record as above. Again, there is an equivalence between cuttings and records. The edge and vertex versions are closely related, and the results are essentially the same.

These random variables can be studied both for deterministic trees and for random trees.

If the tree is a path, we have the classical record problem studied by Rényi, [9].

Our main results are for the case when the tree T_n itself is random, more precisely a random conditioned Galton–Watson tree (also known as simply generated tree) with n vertices. (It is well-known that examples include random labelled trees and random binary trees.) Since both the records (or cutting) and the tree now are random, $X(T_n)$ can be regarded in (at least) two ways.

First, we can regard $X(T_n)$ as a random variable, obtained by picking a random tree T_n and then a random cutting of it. This point of view has been taken by Meir and Moon [6] (mean and variance for Cayley trees), Chassaing and Marchand [3] (asymptotic distribution for Cayley trees), Panholzer [7, 8] (asymptotic distribution for some special families of simply generated trees, and for non-crossing trees). We extend these results to all conditioned Galton–Watson trees. (All unspecified limits are as $n \to \infty$.)

Theorem 1. *Let T_n be a conditioned Galton–Watson tree of size n, defined by an offspring distribution ξ with mean $\mathbb{E}\,\xi = 1$ and finite variance $\sigma^2 > 0$. Then,*

$$\frac{X(T_n)}{\sigma n^{1/2}} \xrightarrow{\mathrm{d}} Z,$$

where Z has a Rayleigh distribution with density $xe^{-x^2/2}$, $x>0$. Moreover, if $\mathbb{E}\,\xi^m<\infty$ for every $m>0$, then all moments converge and thus, for every $r>0$,

$$\mathbb{E}\,X(T_n)^r \sim \sigma^r n^{r/2}\,\mathbb{E}\,Z^r = 2^{r/2}\sigma^r\Gamma\bigl(\tfrac{r}{2}+1\bigr)n^{r/2}.$$

The other point of view is to study $X(T_n)$ as a random variable conditioned on T_n. In other words, we consider the random procedure in two steps: First we choose a random tree $T=T_n$. Then we keep this tree fixed and consider random cuttings of it; this gives a random variable $X(T)$ with a distribution that depends on T. Normalizing as in the theorem above, we consider the distribution of $\sigma^{-1}n^{-1/2}X(T_n)$ given T_n; this is thus a random probability distribution. We then can show that this random probability distribution converges in distribution to a random probability distribution (that does not depend on ξ); this random distribution has moments that can be expressed as functionals of a Brownian excursion.

The proofs are based on Aldous' theory of the continuum random tree [1, 2].

Finally, we study the case when the tree is a (deterministic) complete binary tree of size n. In this case, both the methods and results are different. There is now a periodicity in the result. This is not surprising for complete binary trees, but it is a bit surprising that the periodicity is in the fractional part $\{\lg n-\lg\lg n\}$.

Theorem 2. *Suppose that $n\to\infty$ such that $\{\lg n-\lg\lg n\}\to\gamma\in[0,1]$. Then*

$$\Bigl(X(T_n)-\frac{n}{\lg n}-\frac{n\lg\lg n}{\lg^2 n}\Bigr)\Big/\frac{n}{\lg^2 n}\xrightarrow{\mathrm{d}}-W_\gamma$$

where W_γ has an infinitely divisible distribution with characteristic function

$$\mathbb{E}\,e^{\mathrm{i}tW_\gamma}=\exp\Bigl(\mathrm{i}f(\gamma)t+\int_0^\infty\bigl(e^{\mathrm{i}tx}-1-\mathrm{i}tx\mathbf{1}[x<1]\bigr)\,d\nu_\gamma(x)\Bigr),$$

where $f(\gamma):=2^\gamma-1-\gamma$ and the Lévy measure ν_γ is supported on $(0,\infty)$ and has density

$$\frac{d\nu_\gamma}{dx}=2^{\{\lg x+\gamma\}}x^{-2}.$$

The strategy of the proof is to approximate $X(T_n)$ by a sum of independent random variables derived from $\{\lambda_e\}$; it turns out that only exceptionally small values at level $\approx\lg\lg n$ have a significant influence on $X(T_n)$. We will then apply a classical limit theorem for triangular arrays.

For details, see [4, 5].

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[2] D. Aldous, The continuum random tree III. *Ann. Probab.* **21**, no. 1, 248–289.

[3] P. Chassaing & R. Marchand. In preparation.

[4] S. Janson, Random cutting and records in deterministic and random trees. Preprint, 2003. Available from `http://www.math.uu.se/~svante/papers`

[5] S. Janson, Random records and cuttings in complete binary trees. In *Mathematics and Computer Science III, Algorithms, Trees, Combinatorics and Probabilities (Vienna 2004)*, Eds. M. Drmota, P. Flajolet, D. Gardy, B. Gittenberger, Birkhäuser, Basel, 2004.
[6] A. Meir & J.W. Moon, Cutting down random trees. *J. Australian Math. Soc.* **11** (1970), 313–324.
[7] A. Panholzer, Cutting down very simple trees. Preprint, 2003.
[8] A. Panholzer, Non-crossing trees revisited: cutting down and spanning subtrees. *Proceedings, Discrete Random Walks 2003*, Cyril Banderier and Christian Krattenthaler, Eds., *Discr. Math. Theor. Comput. Sci.* **AC** (2003), 265–276.
[9] A. Rényi, (1962). On the extreme elements of observations. *MTA III, Oszt. Közl.* **12** (1962) 105–121. Reprinted in *Collected Works*, Vol III, pp. 50-66, Akadémiai Kiadó, Budapest, 1976.

Moment conditions for weighted branching processes

Dirk Kuhlbusch

The talk presents moment conditions determining the asymptotic behaviour of *weighted branching processes (WBP)* and *weighted branching processes in random environments (WBPRE).*
In the case of a WBPRE $(Z_n)_{n\geq 0}$, the underlying reproduction mechanism is driven by a stationary ergodic sequence $(U_n)_{n\geq 0}$ of (random) probability measures which can be viewed as environmental changes over time. In the first part of the talk, we obtain a nonnegative martingale by normalizing the underlying process with its conditional means (given the environmental sequence) and present necessary and sufficient conditions guaranteeing $\mathfrak{L}_1$-convergence of this martingale. These conditions are substantially simplified in the case of i.i.d. environmental sequences. The proofs of these results are based on a change of measure on the set of weighted family trees associated with the underlying process, thus adapting the techniques used in [5] and [4] for *Galton-Watson processes* and *branching random walks.*
The second part of the talk focusses on ordinary WBP (deterministic and nonvarying environments). Given any $p \in (1, \infty)$, we give a complete characterization of $\mathfrak{L}_p$-convergence of the corresponding martingale. Moreover, we analyse when the martingale limit has finite moments with respect to a certain class of regularly varying functions. The techniques used for these results exploit the inherent *double martingale structure*, generalizing the methods used in [1] for Galton-Watson processes.

References

[1] Alsmeyer, G. and Rösler, U. On the existence of ϕ-moments of the limit of a normalized supercritical Galton-Watson process. To appear in *J. Theor. Probab.* (2004).
[2] Kuhlbusch, D. On weighted branching processes in random environment. *Stoch. Proc. Appl.* **109**, 113-144 (2004).
[3] Kuhlbusch, D. *Moment conditions for weighted branching processes.* Dissertation, University of Münster (2004).

[4] Lyons, R. A simple path to Biggins' martingale convergence for branching random walk. In K.B. Athreya and P. Jagers (eds), *Classical and Modern Branching Processes.* IMA Vol. Math. Appl. **84**, 217-221. Springer, New York (1997).

[5] Lyons, R., Pemantle, R. and Peres, Y. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **25**, 1125-1138 (1995).

Two-way chaining with reassignment

Luc Devroye

(joint work with K. Dalal and E. Malalla)

In classical uniform hashing with chaining, a set of s keys are inserted into a hash table with n separate chains (or linked lists) via a uniform hash function. The insertion time is constant, and the average search time is proportional to the load factor of the hash table $\alpha := s/n$. However, even for constant load factor, the worst-case search time (the length of the longest chain) is asymptotic to $\log n / \log \log n$, in probability [17, 24].

Azar et al. [2] suggested a novel approach called the *greedy two-way chaining* paradigm. It uses two independent uniform hash functions to insert the keys where each key is inserted *on-line* into the shorter chain, with ties broken randomly. The insertion time is still constant, while the average search time cannot be more than twice the average search time of classical uniform hashing. However, the expected maximum search time is only $2 \log_2 \log n + 2\alpha + O(1)$ [2, 3, 21]. The two-way chaining paradigm has been effectively used to derive many efficient algorithms [4, 5, 6]. A further variant of on-line two-way chaining [25] improves the maximum search time by a constant factor.

On the other hand, one can show that the *off-line* version of two-way chaining, where all the hashing values of the keys are known in advance, yields better worst-case performance [2, 7, 22]. Czumaj and Stemann [7] proved that if $s \leq 1.67545943... \times n$, one can find an assignment for the keys such that the maximum chain length is at most 2, w.h.p. (with high probability, i.e., with probability tending to one as $n \to \infty$). There is a large gap between the worst-case performances of the on-line and off-line versions of two-way chaining. Subsequently, one naturally wonders if it is possible to design an efficient on-line two-way chaining algorithm whose worst-case search time is close enough to its off-line one, while preserving constant insertion time and $O(\alpha)$ average search time. Our goal here is to obtain constant expected maximum search and deterministic $O(1)$ insertion times when the load factor of the hash table is constant.

Many hashing schemes that achieve constant worst-case search time have been developed [10, 11, 12, 15]. However, these schemes use a large number of hash functions, sometimes employ rehashing techniques, and have insertion times that are constant only in expected amortized sense. The closest to our work is a new hashing scheme called cuckoo hashing [23, 9] which utilizes the two-choice paradigm to improve the worst-case performance, but it relies also on the idea of

reallocation of the inserted keys. It inserts n keys into a hash table that is partitioned into two parts, each of size $\lceil (1+\epsilon)n \rceil$, for some constant $\epsilon > 0$. It uses two independent hash functions. Each key is hashed initially by the first function to a cell in the first sub-table. If the cell is full, then the new key is inserted there anyway, and the old key is kicked out to the second sub-table to be hashed by the second function. The same rule is applied in the second sub-table. Keys are moved back and forth until a key moves to an empty location or a limit of $O(\log n)$ moves is reached. When the limit is reached, new independent hash functions are chosen, and the whole table is rehashed. The worst-case search time is at most two, but the insertion time is constant only in an amortized expected sense. An off-line and static version of this algorithm previously appeared in [22].

In our presentation at oberwolfach, we presented a two-way chaining algorithm that is close to cuckoo hashing while achieving constant worst-case insertion time, deterministically, and constant worst-case search time asymptotically almost surely, when the load factor is constant. The space consumption is also linear. The idea is based on the structure of a random multi-graph, a key reallocation technique, and a deamortization method. The algorithm is divided into stages where at each stage the hash table is modelled by a random graph with n vertices representing the chains and m edges denoting the the pairs of hash values for the keys inserted during the stage. Inserting keys into chains corresponds to orienting edges towards vertices. Our goal then is to minimize the maximum out-degree. This model has been used earlier to analyze the off-line version of two-way chaining [7]. When the graph is a forest, it is easy to orient the edges such that the maximum out-degree is one. In order to keep the maximum out-degree as low as possible, some edges need to be reoriented when two trees are joined during the hashing process, and this means that the corresponding keys also need to be reallocated. Furthermore, cycles could occur in the random graph. Since the hashing process is on-line, we use a queue to control the orientation process, thereby ensuring that every insertion operation takes only a constant time of work. This leads us to the elegant deamortization method introduced by Gajewska and Tarjan [16].

A key is immediately assigned a bin according to its first hash value. Finding out in which tree of the forest a value lies, we follow arrows up to the root of the tree stored so far. Unfinished work (root finding and reorientations) are left in the queue, as we permit ourselves to carry out a constant number of steps per insertion. Thus, if the queue is long, all the keys involved in the queue are quite arbitrarily placed. Fortunately, if $m < (1-\epsilon)n/2$, then the (random) graph consists basically of trees and a few unicyclic components. As edges that cause a cycle are ignored (and thus, the corresponding keys are also arbitrarily placed), we only store trees. The maximal tree is known to be $O(\log n)$ in size, and $O(1)$ in size on average. This implies that the queue will never be long, and thus, all keys involved in it are very likely different and even from different components in the forest. Hence, the (temporary) arbitrary placement of keys increases the maximal chain length by $O(1)$ with high probability. The placement of cycle-causing keys also accounts for no more than $O(1)$ w.h.p. Furthermore, the tree structure orientation assures

that those keys that have been included in the trees are all placed in different bins. If m exceeds the bound given above, then we junk the forest, and start anew as soon as the limit $(1-\epsilon)n/2$ is reached, assigning keys in the same way using newly built forests. This can be repeated about $2m/((1-\epsilon)n$ times. The maximal chain length is thus not more than a constant times the number of such stages.

We can thus show that if $m = \lfloor \beta n \rfloor \leq s = O(n \log n)$ for some constant $\beta < 1/2$, and if we pick the constant amount of work we do per insertion as an appropriate function of β, then w.h.p. the maximum search time is at most $2 \lceil s/m \rceil + 6$ and the maximum chain length is at most $\lceil s/m \rceil + 3$. The maximum is taken over the entire history of all chains, not just at the end. The theorem confirms that if the load factor of the hash table $s/n = O(1)$, then asymptotically almost surely the maximum search time is constant. Since there is a trivial lower bound of $2s/n$, we see that we are roughly within $1/\beta$ of the best possible, recalling that β can be picked arbitrarily close to $1/2$.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, 2nd ed., John Wiley, New York, 2000.

[2] Y. Azar, A. Z. Broder, A. R. Karlin and E. Upfal, "Balanced allocations," *SIAM Journal on Computing*, vol. 29 (1), pp. 180–200, 2000. A preliminary version of this paper appeared in *Proceedings of the 26th ACM Symposium on Theory of Computing (STOC)*, pp. 593–602, 1994.

[3] P. Berenbrink, A. Czumaj, A. Steger, and B. Vöcking, "Balanced allocations: the heavily loaded case," in: *Proceedings of the 32nd ACM Symposium on Theory of Computing (STOC)*, pp. 745–754, 2000.

[4] A. Broder and M. Mitzenmacher, "Using multiple hash functions to improve IP lookups," in: *Proceedings of the IEEE INFOCOM 2001 Conference*, Anchorage, Alaska USA , April 2001. Full version available as Technical Report TR–03–00, Department of Computer Science, Harvard University, Cambridge, MA, 2000.

[5] J. Byers, J. Considine, and M. Mitzenmacher, "Simple load balancing for distributed hash tables," in: *Proceedings of the 2nd International Workshop on Peer-to-Peer Systems*, pp. 80–87, 2003.

[6] A. Czumaj, F. Meyer auf der Heide, and V. Stemann, "Contention resolution in hashing based shared memory simulations," *SIAM Journal on Computing*, vol. 29, No. 5, pp. 1703–1739, 2000.

[7] A. Czumaj and V. Stemann, "Randomized allocation processes," *Random Structures and Algorithms*, vol. 18, Issue 4, pp. 297–331, June 2001.

[8] L. Devroye, "Branching processes and their applications in the analysis of tree structures and tree algorithms," in: *Probabilistic Methods for Algorithmic Discrete Mathematics*, ed. M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed, pp. 249–314, 1998.

[9] L. Devroye and P. Morin, "Cuckoo hashing: further analysis," *Information Processing Letters*, vol. 86, pp. 215–219, 2004.

[10] M. Dietzfelbinger, A. Karlin, K. Mehlhorn, F. Meyer auf der Heide, H. Rohnert, and R. Tarjan, "Dynamic perfect hashing: upper and lower bounds," *SIAM Journal on Computing*, vol. 23 (4), pp. 738–761, 1994. A preliminary version appeared in: *Proceedings of the 29th IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 524–531, 1988.

[11] M. Dietzfelbinger and F. Meyer auf der Heide, "A new universal class of hash functions and dynamic hashing in real time," in: *Proceedings of the 17th International Colloquium on Automata Languages and Programming*, LNCS 443, Springer-Verlag, pp. 6–19, 1990.

[12] M. Dietzfelbinger and F. Meyer auf der Heide, "High performance universal hashing, with applications to shared memory simulations," in: *Data Structures and Efficient Algorithms*, LNCS 594, Springer-Verlag, pp. 250–269, 1992.
[13] P. Erdös, "Some remarks on the theory of graphs," *Bulletin of the American Mathematical Society*, vol. 53, pp. 292–294, 1947.
[14] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publ. Math. Ins. Hunger. Acad. Sci.*, Vol. 5, PP. 17-61, 1960.
[15] M. Fredman, J. Komlós, E. Szemerédi, "Storing a sparse table with $O(1)$ worst case access time," *Journal of the ACM*, vol. 31, pp. 538–544, 1984.
[16] H. Gajewska and R. E. Tarjan, "Deques with heap order," *Information Processing Letters*, vol. 22(4), pp. 197–200, 1986.
[17] G. H. Gonnet, "Expected length of the longest probe sequence in hash code searching," *Journal of the ACM*, vol. 28, pp. 289–304, 1981.
[18] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, John Wiley & Sons, New York, 2000.
[19] R. M. Karp, "The transitive closure of a random digraph," *Random Structures and Algorithms*, vol. 1, PP. 73-93, 1990.
[20] E. Malalla, *Two-way Hashing with Separate Chaining and Linear Probing*, Ph.D. thesis, School of Computer Science, McGill University, 2004.
[21] M. Mitzenmacher, A. Richa, and R. Sitaraman, "The power of two random choices: A survey of the techniques and results," in: *Handbook of Randomized Computing*, (P. Pardalos, S. Rajasekaran, and J. Rolim, eds.), pp. 255–305, 2000.
[22] R. Pagh, "On the cell probe complexity of membership and perfect hashing," in: *Proceedings of 33rd ACM Symposium on Theory of Computing (STOC)*, pp. 425–432, 2001.
[23] R. Pagh and F. F. Rodler, "Cuckoo hashing," in: *Proceedings of the European Symposium on Algorithms*, LNCS 2161, Springer-Verlag, pp. 121–133, 2001. A previous version is available as BRICS Report Series RS–01–32, Department of Computer Science, University of Aarhus, 2001.
[24] M. Raab and A. Steger, ""Balls into bins"—a simple and tight analysis," in: *Proceedings of the 2nd Workshop on Randomization and Approximation Techniques in Computer Science*, vol. 1518, Lecture Notes in Computer Science, Springer-Verlag, pp. 159–170, 1998.
[25] B. Vöcking, "How asymmetry helps load balancing," in: *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 131–141, 1999.

Branching random walks on trees and the Brownian snake

Svante Janson
(joint work with Jean-François Marckert)

Consider a rooted ordered finite tree T where each edge is assigned a real number called *value*. We then let, for every vertex v, S_v be the sum of the values of the edges along the paths from the root to v. We will assume that the values of the edges are independent random variables with a common distribution. We let Y denote one of these values.

We study the case when the tree itself is random, more precisely a random conditioned Galton–Watson tree (or simply generated tree) with n vertices.

Each S_v is a sum of i.i.d. variables, and the number of terms is the depth of v, which typically is of the order $n^{1/2}$. Hence, by the central limit theorem, S_v is typically of order $n^{1/4}$ if $\mathbb{E}\,Y = 0$ and $\operatorname{Var} Y < \infty$, but S_v is typically of order $n^{1/2}$ if $\mathbb{E}\,Y \neq 0$.

To study the case $\mathbb{E}\,Y = 0$ in more detail, we take the values $n^{-1/4}S_v$ in the order given by the depth first walk on the tree, extend this by linear interpolation to a continuous function, and rescale to obtain a function $r_n(s)$ on $[0,1]$.

Before proceeding, recall that the Brownian snake [4, 7] is a random function that can be described as follows: Let $\zeta(s)$ be a random non-negative function on a given interval I; in our case, ζ is a standard Brownian excursion on $I = [0,1]$. (Another common version is with ζ reflected Brownian motion on $[0,\infty)$.) Then the corresponding Brownian snake $W(s,t)$ is a random function of two variables (or stochastic field), $s \in I$ and $t \geq 0$, such that conditioned on ζ, $W(s,t)$ is a Gaussian process with mean $\mathbb{E}\,W(s,t) = 0$ and covariance function, if $s_1 \leq s_2$,

$$\operatorname{Cov}\bigl(W(s_1,t_1), W(s_2,t_2)\bigr) = \min\Bigl(t_1,\, t_2,\, \inf_{u\in[s_1,s_2]} \zeta(u)\Bigr).$$

In particular, for fixed s, $t \mapsto W(s,t)$ is a Brownian motion stopped at $t = \zeta(s)$. Two such Brownian motions for s_1 and s_2 are identical for $t \leq \inf_{u\in[s_1,s_2]} \zeta(u)$, and then evolve independently.

Let $r(s) := W(s,\zeta(s))$ (known as the head of the snake).

Assume $\mathbb{E}\,Y = 0$ and $\operatorname{Var} Y = 1$, and let $n \to \infty$. It has been shown by Marckert and Mokkadem [8] and Chassaing and Schaeffer [3] in special cases, and by Gittenberger [5] in general, assuming $\mathbb{E}\,|Y|^{8+\varepsilon} < \infty$, that then $r_n \xrightarrow{\mathrm{d}} r$ in $C[0,1]$ (i.e. in the uniform topology).

We want to weaken the moment condition on Y as far as possible. First, it is easy to see that $r_n \to r$ in the sense of finite-dimensional distributions without further assumptions. Weak convergence in $C[0,1]$ is equivalent to the convergence of the finite-dimensional distributions together with tightness. Often, the tightness is a technical nuisance that can be verified with more or less work. Here, that is not the case and we need a stronger condition on Y in order to obtain convergence.

Theorem 1. *Assume $\mathbb{E}\,Y = 0$. Then $r_n \xrightarrow{\mathrm{d}} r$ in $C[0,1]$ if and only if $\mathbb{P}(|Y| \geq y) = o(y^{-4})$.*

In particular, this holds if $\mathbb{E}\,Y^4 < +\infty$, and no weaker moment condition suffices.

When this condition fails, we do not have convergence because the extreme values of Y will cause thin spikes in r_n. These spikes are at random positions, and are therefore not seen by the finite-dimensional distributions. We also have convergence $r_n \xrightarrow{\mathrm{d}} r$ in, for example, $L^2[0,1]$.

If Y have tails that are exactly of the order y^{-4}, then r_n converges in distribution to a "hairy snake", i.e. a Brownian snake with hairs added. The hairs are vertical line segments going up or down from the snake; their positions and lengths are given by a Poisson process, so the number of them is infinite, but there is only a finite number of them with length larger than a given number.

Note that this limiting object, the hairy snake, is *not* a function, and therefore the convergence does not take place in $C[0,1]$. Instead we identify continuous functions on $[0,1]$ with their graphs, and obtain convergence in the space of compact subsets of $\mathbb{R}^2$. This seems to be a novel type of convergence in this context.

If the tails of Y are even larger, the spikes dominate and we may after suitable rescaling obtain convergence to a flat (or dead) hairy snake, with hairs as above added to the line segment from $(0,0)$ to $(1,0)$, and thus without the Brownian part.

The proofs are based on Aldous' theory of the continuum random tree [1, 2]. For details, see [6].

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[2] D. Aldous, The continuum random tree III. *Ann. Probab.* **21** (1993), no. 1, 248–289.

[3] P. Chassaing & G. Schaeffer, Random planar lattices and integrated superBrownian excursion. *Prob. Theo. Rel. Fields*, to appear.

[4] T. Duquesne & J.F. Le Gall, Random trees, Lévy processes and spatial branching processes. *Astérisque* **281** (2002).

[5] B. Gittenberger, A note on "State spaces of the snake and its tour – Convergence of the discrete snake" by J.-F. Marckert and A. Mokkadem. *J. Theo. Prob.*, **16** (2003), 1063–1067.

[6] S. Janson & J.-F. Marckert, Convergence of discrete snakes. Preprint, 2003. Available from `http://www.math.uu.se/~svante/papers`

[7] J.F. Le Gall, *Spatial branching processes, random snakes and partial differential equations.* Lectures in Mathematics ETH Zürich, Birkhäuser, Basel, 1999.

[8] J.F. Marckert & A. Mokkadem, States spaces of the snake and of its tour – Convergence of the discrete snake. *J. Theo. Prob.*, **16** (2003), 1015–1046.

Recent developments on the contraction method

Ralph Neininger

In this talk we describe a systematic approach to limit laws for sequences of random vectors which satisfy distributional recursions as they appear under various models of randomness for parameters of trees, characteristics of divide-and-conquer algorithms, or, more generally, for quantities related to recursive structures.

While there are also strong analytic techniques to the subject we extend and systematize a more probabilistic approach, the contraction method. This method was first introduced for the analysis of Quicksort by Rösler (1991) and further on developed independently by Rösler (1992) and by Rachev and Rüschendorf (1995). The name of the method refers to the fact that the analysis makes use of an underlying map of measures, which is a contraction with respect to some probability metric.

In this context often an approach based on the minimal L_2 metric ℓ_2 is used along the lines of Rösler's original work on Quicksort. Although the ℓ_2 approach works well for many problems leading to non-normal limit distributions its main defect is, that it typically will not work for normal limit laws. To overcome this problem, we propose to use as alternative metrics the Zolotarev metrics ζ_s which

are more flexible and at the same time still manageable. The advantage of alternative metrics as the Zolotarev metric for the analysis of algorithms had already been demonstrated at some examples in the paper Rachev and Rüschendorf (1995).

The flexibility of the ζ_s metrics consists in the fact that, where for $s = 2$ we reobtain the common ℓ_2 theory, we may also use ζ_s with $s > 2$, which gives access to normal limit laws or, $s < 2$, leading to results where we can weaken the assumption of finite second moments, an assumption being usually present in the ℓ_2 approach.

In his 1999 paper "Normal convergence problem? Two moments and a recurrence may be the clues." Pittel stated as a heuristic principle that various global characteristics of large size combinatorial structures such as graphs and trees are asymptotically normal if the mean and variance are "nearly linear" in n. As technical reason he argued that the normal distribution with the same two moments "almost" satisfies the recursion. He exemplified this idea at the independence number of uniformly random trees. An essential step in the proof of our limit theorem is the introduction of an accompanying sequence which fulfills approximatively a recursion of the same form as the characteristics do and is formulated essentially in terms of the limiting distribution. This is similar to the technical idea proposed by Pittel.

In Neininger and Rüschendorf (2004) we obtain a general limit theorem for divide-and-conquer recursions where the conditions are formulated in terms of relationships of moments and a condition ensuring the asymptotic stability of the recursive structure. These conditions can quite easily be checked in a series of examples and allow to (re)derive many examples from the literature. In fact, for the special case of normal limit laws we need — according to Pittel's principle — the first and second moment to apply the method.

As applications of our general transfer result we (re)-derive various central limit laws for random recursive structures, ranging from the size of m-ary search trees or random tries, path lengths in digital search trees, tries, and Patricia tries, via top-down mergesort and the maxima in right triangles to parameters of random recursive trees and plane-oriented versions thereof.

We also show how refined asymptotics of the moments for these problems can be turned into rates of convergence in Zolotarev's metric.

References

[1] Neininger, R. and L. Rüschendorf (2004) A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab. 14*, 378–418.

[2] Pittel, B. (1999) Normal convergence problem? Two moments and a recurrence may be the clues. *Ann. Appl. Probab. 9*, 1260–1302.

[3] Rösler, U. (1991) A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl. 25*, 85–100.

[4] Rösler, U. (1992) A fixed point theorem for distributions. *Stochastic Process. Appl. 42*, 195–214.

[5] Rachev, S. T. and L. Rüschendorf (1995) Probability metrics and recursive algorithms. *Adv. in Appl. Probab. 27*, 770–799.

The Contraction Method for max-Recursive Sequences

Ludger Rüschendorf

We consider max-recursive sequences (Y_n) which satisfy a recursive stochastic equation of the form

$$Y_n \stackrel{d}{=} \bigvee_{r=1}^{k} \left(A_r(n) Y^{(r)}_{I_r^{(n)}} + b_r(n) \right) \tag{1}$$

where $A_r(n)$, $b_r(n)$ are stochastic factors, $I_r^{(n)}$ are random subgroup sizes, $0 \leq I_r^{(n)} \leq n-1$, and $\left(Y_k^{(r)}\right)_{r=1,\ldots,k}$ are iid copies of Y_k. This type of recursions arises typically in the analysis of worst cases, of the height of trees or in parallel search algorithms.

After normalization and assuming convergence of the normalized coefficients the first problem is to analyze *limit equations* of the form

$$X \stackrel{d}{=} \bigvee_{r=1}^{k} (A_r X_r + b_r) \tag{2}$$

where A_r, b_r are random. The right-hand side defines an operator T on the set of all probability measures $\mathcal{M}$ on $\mathbb{R}^1$ defined by

$$T\mu \stackrel{d}{=} \bigvee_{r=1}^{k} (A_r X_r + b_r) \text{ with } (X_r) \text{ independent}, \ X_r \stackrel{d}{=} \mu. \tag{3}$$

The following general existence and uniqueness result is due to Neininger and Rüschendorf (2004). We denote for $s > 0$ by ℓ_s the minimal L_s-metric.

Theorem 1: Let A_r, $b_r \in L^s$, $\mu_0 \in \mathcal{M}$ and $\zeta_s := E \sum_{r=1}^{k} |A_r|^s < 1$ and let $\ell_s(\mu_0, T\mu_0) < \infty$. Then the stochastic equation (2) has a unique solution in $\mathcal{M}_s(\mu_0) = \{\mu \in \mathcal{M} : \ell_s(\mu, \mu_0) < \infty\}$.

The proof of Theorem 1 is based on the following contraction property: For any $s > 0$ and any $P, Q \in \mathcal{M}$ holds

$$\ell_s(TP, TQ) \leq \left(E \sum_{r=1}^{l} |A_r|^s \right)^{1/s \wedge 1} \ell_s(P, Q). \tag{4}$$

(4) extends an earlier related contraction result in Rachev and Rüschendorf (1992) which was applied there to limit theorems for maxima. Based on this theorem the following limit theorem for max-recursive sequences was established in Neininger and Rüschendorf (2004).

Theorem 2: Consider the max-recursive equation in L^s given in (1). We assume

A1) Stabilization:

$$\left(A_1^{(n)}, \ldots, A_k^{(n)}, b_1(n), \ldots, b_k(n) \right) \xrightarrow{L^s} (A_1^*, \ldots, A_k^*, b_1^*, \ldots, b_k^*)$$

A2) Contraction condition:

$$E\sum_{r=1}^{k}|A_r^*|^s < 1$$

A3) Nondegeneracy: For all $\ell \in \mathbb{N}$ and $r = 1, \ldots, k$ holds

$$E\left[1_{\{I_r^{(n)} \le \ell\} \cup \{I_r^{(n)} = n\}} |A_r^{(n)}|^s\right] \longrightarrow 0.$$

Then $\ell_s(X_n, X^*) \to 0$ where X^* is the unique solution of the limit equation

$$X^* = \bigvee_{r=1}^{k} (A_r^* X_r^* + b_r^*)$$

in the class of distributions with finite sth moments.

One direct application of Theorem 2 is the limit theorem for the worst case M_n of the FIND algorithm

$$M_n \xrightarrow{\mathcal{D}} M^* \tag{5}$$

where M^* is the unique solution in $\mathcal{M}_s$, $s > 1$ of

$$M \stackrel{d}{=} 1 + UM \vee (1-U)\bar{M}. \tag{6}$$

This result was first given in Grübel and Rösler (1996) based on process theory for the FIND process.

References

[1] R. Grübel und U. Rösler (1996). Asymptotic distribution theory for Hoare's selection algorithm. *Advances in Applied Probability 28*, 252–269.

[2] R. Neininger and L. Rüschendorf (2004). General limit theorem for recursive algorithms and combinatorial structures. *Annals of Applied Probability 14*, 378–418.

[3] S. T. Rachev and L. Rüschendorf (1992). Rate of convergence for sums and maxima and doubly ideal metrics. *Theory Prob. Appl. 37*, 276–289.

Stochastic fixed points of max-type

Uwe Roesler

Our aim is to give an overview on solutions of the fixed point equation (FPE)

$$X \stackrel{\mathcal{D}}{=} \sup_{i \in I\!N} (T_i X_i + C_i) \tag{1}$$

of max-type. The random variables $(T_i, C_i)_i, X_j, j \in I\!N$ are independent and X_i has the same distribution as X. We consider here only positive factors T_i and exclude trivial cases.

The oldest literature on max-type equations seems to be Rachev and Rueschendorf in 92 [3] using ideal metrics. An overview of examples is in Aldous-Bandyopadhyay 04 [1], however under a different point of view.

Stochastic FPE appear naturally in tree structures, the weighted branching process [6]. Let $V = I\!N^*$ be the rooted branching tree of finite sequences. Every node $v \in V$ carries a random cost function $C(v)$ and every edge (v, vi) carries a random factor $T_i(v)$. Different families behave independently, $(C(v), T_1(v), T_2(v), \ldots), v \in V$ are iid rvs, but within the families we allow arbitrary dependence,

$$C(v), T_1(v), T_2(v), \ldots$$

dependent for fixed v. The weight or layer $L(v)$ of a node v is the product of all factors on the path back to the root.

Common objects are the sum $Z_n := \sum_{|v|=n} L(v)C(v)$ or the largest value $L_n := \sup_{|v|=n} L(v)C(c)$. They satisfy the backward recursive equations $Z_{n+1} = \sum_i T_i Z_n(i)$ and $L_{n+1} = \sup_i T_i L_n(i)$. Under suitable conditions converge Z_n or L_n to some limit X, which satisfy a FPE of sum-type respectively of max-type

$$X \stackrel{\mathcal{D}}{=} \sum_i T_i X_i \qquad X \stackrel{\mathcal{D}}{=} \sup_i T_i X_i.$$

FPEs with cost function appear by summing over all nodes up to the n-th generation. The equation (1) comes up as the limit of $X_n := \sum_{|v|=n} \sum_w L(w)C(wv_{||w|+1})$ summing over all strict ancestors w of v.

The contraction method [6] was exploited by Neininger-Rueschendorf 03 [4]. The minimal l_p-metric is well suited for additive cost functions C_i, the weighted Kolmogorov distance for no costs. However, for zero cost the problem arises of choosing a good starting distribution.

We concentrate now on no costs, $C_i \equiv 0$. Notice equation (1) is equivalent to each of the following

$$\ln X \stackrel{\mathcal{D}}{=} \sup_i (\ln X_i + \ln T_i) \qquad \frac{1}{X} \stackrel{\mathcal{D}}{=} \inf_i \frac{1}{T_i}\frac{1}{X_i} \qquad X^\alpha \stackrel{\mathcal{D}}{=} \sup_i T_i^\alpha X_i^\alpha$$

$\alpha > 0$. In particular the max-type and min-type solutions are the same. Using $P(\frac{1}{X} \geq t) = \overline{F}(t)$ FPE (1) is equivalent to survival function solutions

$$\overline{F}(t) \stackrel{\mathcal{D}}{=} E \prod_i \overline{F}(tT_i). \tag{2}$$

There are two classes of solutions of (1), X unbounded and X bounded. The first is connected to FPEs of sum-type and the second to a monotone operator. These approaches are supplementary to each other.

Theorem 1 (Jagers-Roesler)**.** *If $m(\alpha) := E\sum_i T_i^\alpha = 1$ and $m'(\alpha) \leq 0$ for some $\alpha > 0$ then (2) has the solution $\overline{F}(t) = Ee^{-t^\alpha W}$ where W is a non trivial solution of $W \stackrel{\mathcal{D}}{=} \sum_{i \in I\!N} T_i^\alpha W(i)$.*

The existence of a non trivial solution W for the factors T_i^α follows in this generality by a result of Lyons. The Laplace transform $\varphi(t) = Ee^{-tW}$ solves $\varphi(t) = E\prod_i \varphi(tT_i^\alpha)$. $\overline{F}(t) = \varphi(t^\alpha)$ does the job.

In the case $\sum_i T_i^\alpha \equiv 1$ the solutions W are constant c and $\overline{F}(t) = e^{-t^\alpha c}$. These solutions appear as additional solutions to the Quicksort FPE,

$$X \stackrel{\mathcal{D}}{=} UX_1 + (1-U)X_2 + C. \tag{3}$$

If Q is the Quicksort distribution satisfying (3) then Q plus a constant solves also (3). These are all solutions of (3).

Theorem 2 (Caliebe-Roesler)**.** *Assume the max-type FPE has a solution such that $0 < \overline{F}(t) < 1$ for all t. If the factors are bounded between some constants $0 < c_1 \le c_2 < \infty$ then the max-type solution is unique up to multiplicative constants.*

The monotone operator approach was introduced by Jagers-Roesler 04 [5]. Consider the operator $K(\mu) = \sup_i(T_i X_i + C_i)$, where $X_i \sim \mu$. Fact: If $\mu \preceq K(\mu)$ in stochastic order, then $K^n(\mu)$ increases to a fixed point of K. The FP can have mass at infinity. Under certain conditions we obtain a measure on the positive reals. For simplicity we present only Peter's water delay example.

Peter's water delay: The edges of a binary rooted tree are open or closed by iid Bernoulli(p) rvs. Water poured into the root crosses an open edge with no time delay and otherwise has one unit time delay. When does the first water arrive at n-th generation and converges τ_n to a limit τ? Taking the view of Galton-Watson processes, open=offspring, the event $\tau = 0$ of instantaneous water at infinity has the probability 1 minus the extinction probability of the Galton-Watson process. Strict positivity is equivalent to $2p > 1$. Notice, for every path the delay time of water will be infinity. The rv $X = 2^{-\tau}$ solves (2) with iid factors T_1 and T_2 and distribution $P(T_1 = 1) = p = 1 - P(T_1 = 1/2)$.

All solutions of (2) are known for deterministic factors.

Theorem 3 (Alsmeyer-Roesler)**.** *If the factors are real numbers then (2) has non-trivial solutions if and only if $m(\alpha) = 1$ for some $\alpha > 0$. If the closed group generated by the factors T_i*

–is $\mathbb{R}_>$ then the only solutions are $\overline{F}(t) = e^{-ct^\alpha}$

–is $\{r^{\mathbb{Z}}\}$ then the only solutions are $\overline{F}(t) = \exp(-t^\alpha h(t))$, where h is r-multiplicative periodic and $\overline{F}$ decreasing.

The proof runs by Choquet-Deny Theorem and cannot be applied to random factors.

References

[1] David Aldous and Bandyopadhyay, *A Survey of Max-Type Recursive Distributional Equations.* Preprint 2004
http://front.math.ucdavis.edu/math.PR/0401388

[2] Gerold Alsmeyer and Uwe Roesler, *Max-type fixed points for deterministic factors.*
http://www-computerlabor.math.uni-kiel.de/stochastik/roesler/ Preprint 2004.

[3] Svetlozar Rachev and Ludger Rueschendorf, *Rate of convergence for sums and maxima and doubly ideal metrics.* Theory Prob. Appl. **37**, 276-289, 1992.

[4] Ralph Neininger and Ludger Rueschendorf, *Analysis of algorithms by the contraction method: additive and max-recursive sequences.*
http://www.math.uni-frankfurt.de/∼neiningr/ Preprint 2003.

[5] Peter Jagers and Uwe Roesler, *Stochastic fixed points for the maximum.* Mathematics and Computerscience III, ed: Drmota, Flajolet, Gardy, Gittenberger; Birkhäuser 2004.
[6] Uwe Roesler and Ludger Rueschendorf, *The contraction method for recursive algorithms.* Algorithmica **29**, 3-33, 2001.

A Stochastic Fixed Point Equation Related to Weighted Branching

Gerold Alsmeyer
(joint work with Uwe Rösler)

Given a sequence $T = (T_j)_{j\in J}$ of real random variables, we consider the stochastic fixed point equation

$$W \stackrel{d}{=} \sum_{j\geq 1} T_j W_j + C, \tag{1}$$

for i.i.d. real-valued random variables $W, W_1, W_2, ...$ which are independent of (C, T), $T \stackrel{\text{def}}{=} (T_j)_{j\geq 1}$. Here we will focus on the homogeneous case where $C = 0$. The goal is to find all (distributional) fixed points of these equations, that is all distributions of W such that (1) holds true.

Put $N \stackrel{\text{def}}{=} \sum_{j\geq 1} \mathbf{1}_{\{T_j\neq 0\}}$. Our mild standing assumptions are

$$\mathbb{P}(N \geq 2) > 0 \tag{A1}$$

and

$$\mathbb{P}(|T_j| \in \{0, 1\} \text{ for all } j \geq 1) < 1. \tag{A2}$$

Fixed points of (1) turn up in a natural way as limits of recursive equations of the form

$$W^{(n+1)} \stackrel{d}{=} \sum_{j\geq 1} T_j^{(n)} W_j^{(n)} + C^{(n)}, \quad n \geq 0. \tag{2}$$

Here $(T^{(n)}, C^{(n)})$, $W_1^{(n)}, W_2^{(n)}, ...$ are independent random variables for each $n \geq 0$. The joint distribution of $C^{(n)}$ and $T^{(n)} = (T_1^{(n)}, T_2^{(n)}, ...)$ is given and converges to that of (C, T). The $W_j^{(n)}$'s are copies of $W^{(n)}$. The first study of a recursive system of type (2) was given by Rösler (1991) for the sorting algorithm Quicksort which might still be the most prominent example. In this case $W_{(n)}$ denotes the normalization of the number of key comparisons necessary to sort a list of n distinct items. With $U \stackrel{d}{=} \text{Unif}(0, 1)$ and $g(t) \stackrel{\text{def}}{=} 1+2t\log t+2(1-t)\log(1-t)$ for $t \in (0, 1)$ the weak limit of $W^{(n)}$ satisfies the fixed point equation

$$W \stackrel{d}{=} UW_1 + (1 - U)W_2 + g(U) \tag{3}$$

and is in fact the unique solution having zero mean and finite variance.

Given a vector $T = (T_j)_{j\geq 1}$ satisfying (A1), (A2), define the smoothing transform K on the set of probability measures on $\mathbb{R}$ by

$$K(\mu) \stackrel{\text{def}}{=} \mathcal{L}\left(\sum_{j\geq 1} T_j W_j\right) \tag{4}$$

where $W_1, W_2, ...$ are independent random variables with distribution μ and where $\mathcal{L}(X)$ means the distribution of a random variable X. The infinite sum does not need to exist but if it does it is understood in the sense of almost sure convergence. By the independence of T and $(W_j)_{j\geq 1}$, this also implies a.s. convergence of $\sum_{j=1}^n t_j W_j$ for $\mathbb{P}(T \in \cdot)$-almost all $(t_j)_{j\geq 1}$.

Denote by $\mathfrak{D}(K)$ the domain of K, viz the set of all probability distributions μ for which $K(\mu)$ is well defined in the just explained manner, and by

$$\mathfrak{F} \stackrel{\text{def}}{=} \{\mu \in \mathfrak{D}(K) : \mu = K(\mu)\}$$

the set of fixed points of K (= solutions of (1)). Notice that $\mathfrak{F}$ always contains the trivial fixed point δ_0, the point mass at 0. With this setup at hand the first part of the talk provides some basic properties of K, notably the weighted branching process representation of fixed points: Plainly, if K^n denotes the n-fold iteration of K, then

$$\mathfrak{F} \subset \mathfrak{D}^*(K) \stackrel{\text{def}}{=} \cap_{n\geq 1}\mathfrak{D}(K^n).$$

Let $\mathbb{V}$ be the infinite tree with vertex set $\cup_{n\geq 0}\mathbb{N}^n$ where $\mathbb{N}^0 \stackrel{\text{def}}{=} \{\emptyset\}$. Each vertex $v = (v_1, ..., v_n)$, shortly written as $v_1v_2...v_n$, is uniquely connected to the root $\emptyset$ by the path $\emptyset \to v_1 \to v_1v_2 \to ... \to v$ of length $|v| = n$. Given i.i.d. copies $T(v) = (T_j(v))_{j\geq 1}$, $v \in \mathbb{V}$, of T, define $L(\emptyset) \stackrel{\text{def}}{=} 1$ and $L(v) \stackrel{\text{def}}{=} T_{v_1}(\emptyset)T_{v_2}(v_1) \cdot ... \cdot T_{v_n}(v_1...v_{n-1})$ for any $v = v_1...v_n \in \mathbb{V}$ of length $n \geq 1$. Then $L(v)$ gives the total weight of the unique path from the root to v under multiplication. For any further $w \in \mathbb{V}$, put also $L_w(\emptyset) \stackrel{\text{def}}{=} 1$ and $L_w(v) \stackrel{\text{def}}{=} T_{v_1}(w)T_{v_2}(wv_1) \cdot ... \cdot T_{v_n}(wv_1...v_{n-1})$ which is the same as $L(wv)/L(w)$ whenever $L(w) \neq 0$.

Now let $X(v)$, $v \in \mathbb{V}$, be i.i.d. random variables with common distribution μ and define $W_n = \sum_{|v|=n} L(v)X(v), \quad n \geq 0$. $(W_n)_{n\geq 0}$ forms a stochastic sequence called weighted branching process. It satisfies the backward equation $W_{n+1} = \sum_{j\geq 1} T_j W_{n,j}$ where the $W_{n,j} = \sum_{|v|=n} L_j(v)X(jv)$, $j \geq 1$, are i.i.d. copies of W_n. Hence $K^n(\mu) = \mathcal{L}(W_n)$ for each $n \geq 0$. In particular, all W_n have distribution μ if μ is a fixed point of K. In this case it is now easily seen that

$$W \stackrel{d}{=} \sum_{|v|=n} L(v)W(v) \tag{5}$$

for every $n \in \mathbb{N}$ where the $W(v)$, $|v| = n$, are i.i.d. copies of W.

If μ is a fixed point with characteristic function (ch.f.) φ then (5) may be restated as

$$\varphi(t) = \mathbb{E}\left(\prod_{|v|=n} \varphi(L(v)t)\right), \quad t \in \mathbb{R} \tag{6}$$

for every $n \in \mathbb{N}$. Put $L \stackrel{\text{def}}{=} (L_v(w))_{v,w\in\mathbb{V}}$ and $\Theta_u L \stackrel{\text{def}}{=} (L_{uv}(w))_{w\in\mathbb{V}}$ for $u \in \mathbb{V}$. Embarking on the observation that $\Phi_n(L, t) \stackrel{\text{def}}{=} \prod_{|v|=n} \varphi(L(v)t)$ is a martingale for each t we explain that it converges to a limit $\Phi(L, t)$ outside a null set not

depending on t and that $\Phi(L,\cdot)$ is a random ch.f. satisfying $\mathbb{E}\Phi(L,\cdot)=\varphi$ and

$$\Phi(l,t) \;=\; \prod_{|v|=n} \Phi(\Theta_v l, l(v)t) \tag{7}$$

for all $n \geq 1$ and $t \in \mathbb{R}$. As to the map K given by (4), the implications of this result are as follows: To each $\mu \in \mathfrak{F}$ there exists a kernel $P_\mu(l,\cdot)$ with ch.f. $\Phi(l,\cdot)$, characterized through (7). Let $\mathfrak{G}$ be the set of all kernels with this property (identifying those which differ only on $\mathbb{P}(L\in\cdot)$-null sets) and notice that $\mathfrak{G}$ forms a convolution semigroup, i.e. $P_1(l,\cdot), P_2(l,\cdot) \in \mathfrak{G}$ implies $P_1(l,\cdot) * P_2(l,\cdot) \in \mathfrak{G}$. Let further $\mathbb{G}(\mathfrak{F})$ be the semigroup generated by $\mathfrak{F}$ under convolution. With this we arrive at the following disintegration result:

PROPOSITION. *For any $\nu \in \mathbb{G}(\mathfrak{F})$ with ch.f. ψ there exists a $\mathbb{P}(L\in\cdot)$-null set $\mathcal{N}$ such that*

$$\Psi(l,t) \;\stackrel{\text{def}}{=}\; \lim_{n\to\infty} \prod_{|v|=n} \psi(l(v)t)$$

exists for all $t\in\mathbb{R}$ and $l \in \mathcal{N}^c$ and is the ch.f. of an element P_ν in $\mathfrak{G}$. If $\mathcal{K}$ denotes the mapping $\nu \mapsto P_\nu$, then $\mathcal{K}$ is a homomorphism from $(\mathbb{G}(\mathfrak{F}),*)$ to $(\mathfrak{G},*)$, that is

$$\mathcal{K}(\nu_1 * \nu_2) \;=\; \mathcal{K}(\nu_1) * \mathcal{K}(\nu_2) \tag{8}$$

for all $\nu_1,\nu_2 \in \mathbb{G}(\mathfrak{F})$. Furthermore,

$$\overline{\nu} \;\stackrel{\text{def}}{=}\; \mathbb{E}P_\nu(L,\cdot) \;=\; \int P_\nu(l,\cdot)\; Q(dl) \tag{9}$$

is always an element of $\mathfrak{F}$ satisfying $\overline{\nu}=\nu$ for $\nu\in\mathfrak{F}$, and the restriction of $\mathcal{K}$ to $\mathfrak{F}$ a bijection between $\mathfrak{F}$ and $\mathfrak{G}$.

The fact that, by (9), every fixed point μ of K can be written as a Q-mixture of distributions, $Q \stackrel{\text{def}}{=} \mathbb{P}(L\in\cdot)$, namely $\mu = \int P_\mu(l,\cdot)\, Q(dl)$, suggests to take a closer look at this decomposition. By drawing on the theory of independent infinitesimal triangular schemes, Caliebe(2003) showed that $P_\mu(l,\cdot)$ is infinitely divisible for Q-almost all l if T satisfies (A1), (A2) and furthermore

$$L_n^* \;\stackrel{\text{def}}{=}\; \sup_{|v|=n} |L(v)| \;\to\; 0 \quad \text{a.s.} \tag{10}$$

It follows from our disintegration approach that the last condition is automatically satisfied if $\mathfrak{F}$ contains nontrivial elements. Apart from this the first part of the theorem below is a restatement of Caliebe's result while the remaining parts provide further information on the parameters in the Lévy-Khinchine representation of the ch.f. $\Phi(l,t)$ of $P_\mu(l,\cdot)$.

THEOREM. *Let $T=(T_j)_{j\geq 1}$ satisfy (A1), (A2) and μ be a nontrivial fixed point. Then the following assertions hold true:*

(a) For Q-almost all l, $P_\mu(l,\cdot)$ is infinitely divisible, i.e. μ constitutes a mixture of infinitely divisible distributions. Its ch.f. φ has a mixed Lévy-Khinchine

representation via $\varphi = \mathbb{E}\Phi(L,\cdot)$ *and*

$$\log \Phi(L,t) \;=\; i\gamma(L)t - \frac{\sigma^2(L)t^2}{2} + \int \Big(e^{itu} - 1 - it\chi(u)\Big)\, \Gamma(L,du) \quad a.s. \qquad (11)$$

where, for each $l \in \mathbb{T}$, $\Gamma(l,\cdot)$ *is a measure on* $\mathbb{R}^*$ *satisfying* $\int (u^2 \wedge 1)\,\Gamma(l,du) < \infty$, $\gamma(l) \in \mathbb{R}$, $\sigma^2(l) \geq 0$, *and* $\chi(u) \stackrel{\text{def}}{=} u\,\mathbf{1}_{[-1,1]}(u) + \mathbf{1}_{(1,\infty)}(u) - \mathbf{1}_{(-\infty,-1)}(u)$.

(b) The functions $\gamma(l), \sigma^2(l)$ *and the Lévy kernel* $\Gamma(l,\cdot)$ *satisfy the relations*

$$\gamma(L) \;=\; \sum_{|v|=n} \left(L(v)\gamma(\Theta_v L) + \int \Big(\chi(L(v)u) - L(v)\chi(u)\Big)\, \Gamma(\Theta_v L, du)\right) \quad a.s.,$$

$$\sigma^2(L) \;=\; \sum_{|v|=n} L(v)^2 \sigma^2(\Theta_v L) \quad a.s.,$$

$$\Gamma(L,\cdot) \;=\; \sum_{|v|=n} \mathbf{1}_{\{L(v)\neq 0\}} \delta_{L(v)} \star \Gamma(\Theta_v L,\cdot) \quad a.s.$$

for all $n \geq 1$, *where* $\star$ *means multiplicative convolution.*

We finally give an outline of how to proceed further in order to determine the elements of $\mathfrak{F}$. This is work in progress with various technical problems still unsolved. We also give a brief account of the case of deterministic weights $T_1, T_2, ...$ for which a complete answer is given in Alsmeyer and Rösler(2004).

References

[1] Alsmeyer, G. and Rösler, U. (2004). A stochastic fixed point equation related to weighted branching with deterministic weights. *Technical Report.*

[2] Caliebe, A. (2003). Symmetric fixed points of a smoothing transformation. *Adv. Appl. Probab.* **35**, 377-394.

[3] Rösler, U. (1991). A limit theorem for Quicksort. *Inform. Théor. Appl./Theor. Inform. Appl.* **25**, 85-100.

Asymptotics for tails and moments

Svante Janson

(joint work with Jim Fill and probably others)

This talk describes some preliminary results and work in progress on tail estimates for certain random variables that appear as limits in problems coming from random trees or analysis of algorithms.

First, it is pointed out that a general theorem by Kasahara and Davies, see [3] gives the equivalence, under quite general conditions, between asymptotics for tail probabilities of a random variable, asymptotics for moments, and asymptotics for the moment generating function.

Secondly, we presented a very recent result (found during this workshop after yesterday's talk by Uwe Rösler) giving estimates for the moments (and thus tail estimates) for solutions to fixed-point equations of the Max-recursive type, for example for the limit variable for MAX-FIND.

Thirdly, we discussed a family of limit random variables that arises in the study of Catalan trees; most of them can be expressed as functionals of a Brownian excursion (using the methods of Aldous [1, 2]) and then the asymptotics can be found using the well-known large deviation principle for Brownian motion. This involves finding a certain constant as the solution to a non-linear variational problem; in some case we can solve this exactly, but in others we only have upper and lower bounds for this constant (differing by a few percent).

References

[1] D. Aldous, The continuum random tree II: an overview. *Stochastic Analysis (Proc., Durham, 1990)*, 23–70, London Math. Soc. Lecture Note Ser. 167, Cambridge Univ. Press, Cambridge, 1991.

[2] D. Aldous, The continuum random tree III. *Ann. Probab.* **21** (1993), no. 1, 248–289.

[3] Ph. Chassaing & S. Janson, The center of mass of the ISE and the Wiener index of trees. *Electronic Comm. Probab.*, to appear.

Phase changes in random point quadtrees

Hsien-Kuei Hwang

(joint work with H.-H. Chern and M. Fuchs)

Point quadtrees. Point quadtrees were introduced by Bentley and Finkel (1974) and have become useful data structures for spatial data in low dimensions. Henceforth, we will simply say *quadtrees* instead of *point quadtrees*. Given a sequence of points in $\mathbb{R}^d$, we can construct the quadtree (associated to the given sequence) as follows. The first point in the sequence is placed at the root, which splits the space $\mathbb{R}^d$ into 2^d quadrants. The root has at most 2^d subtrees, corresponding to the quadtrees (constructed recursively by the same procedure) of points falling in each quadrant. By construction, such trees allow operations like range search and partial match queries. When $d = 1$, they are nothing but the binary search trees. For more information and applications, see the two books by Samet (1990a, 1990b).

Random quadtrees. To understand the typical shapes and behaviors of quadtrees, the following probability model is always adopted. Assume that the given (sequence of) random variables are independently, identically and uniformly distributed in the hypercube $[0, 1]^d$. Then construct the quadtree; the resulting tree is called a *random quadtree*.

Known probabilistic properties. Many parameters of random quadtrees have been studied in the literature:

- depth (the distance to the root of a randomly chosen node): Devroye and Laforest (1990), Flajolet et al. (1993), Flajolet and Lafforgue (1994), Flajolet et al. (1995), Devroye (1999);
- page usage: Hoshi and Flajolet (1992), Flajolet et al. (1995);

- number of node types: Labelle and Laforest (1995, 1996), Flajolet et al. (1995), Minh and Jacob (2000), Mahmoud (2002);
- height (the number of nodes in the longest path starting from the root): Devroye (1987, 1998);
- partial-match queries: Flajolet et al. (1993), Flajolet et al. (1995), Neininger and Rüschendorf (2001), Martínez et al. (2001), Chern and Hwang (2003);
- total path length (sum of the distances of all nodes to the root): Flajolet et al. (1993), Flajolet et al. (1995), Neininger and Rüschendorf (1999).

In particular, limit distributions were derived in Flajolet and Lafforgue (1994), Devroye (1998) for depth, in Neininger and Rüschendorf for total path length.

Random fragmentation model. This work was partly motivated by the following physical model for quadtrees proposed by Dean and Majumdar (2002). Their model is roughly described as follows. Pick a point at random in $[0, x]^d$ for some $x \gg 1$ (under some law). This point then splits the original space into 2^d quadrants. Continue the same process for all 2^d quadrants until the volumes are less than unity. The limit distribution of the total number of splittings is claimed to be normal for $1 \le d \le 8$ and is not known for $d \ge 9$. The fragmentation process can easily be described in terms of trees with at most 2^d branches for each node, and such trees may be regarded as certain approximation model for random quadtrees.

Main aim. Our purpose of this work is to develop asymptotic tools for proving that the same phase change phenomenon for limit laws holds in random quadtrees (not under the approximate model of Dean and Majumdar). We will also derive finer results reflecting deeper changes of other stochastic behaviors. Since the number of nodes (corresponding to the number of splittings in the model of Dean and Majumdar, 2002) is n in quadtrees with n keys, we consider instead the number of leaves and show that it is asymptotically normally distributed for $1 \le d \le 8$ and that the normalized random variables do not converge to a fixed limit law. Finer results like Berry-Esseen and local limit theorems are also derived for $1 \le d \le 8$. The tools we develop are of some generality and apply for a very wide class of "toll functions" (which is $\delta_{n,1}$ for the number of leaves).

Main results. Let X_n denote the number of leaves in a random quadtree of n nodes. When $d = 1$, the asymptotic normality is known since David and Barton (1962) since there is a bijection between X_n and the number of peaks in random permutation of n elements; see Devroye (1991), Flajolet et al. (1997) for the local limit theorem.

Our results are summarized as follows. Let $X_n^* := (X_n - \mathbb{E}(X_n))/\sqrt{\mathbb{V}(X_n)}$.

– If $1 \le d \le 8$, then $(X_n - \mu n)/(\sigma\sqrt{n}) \xrightarrow{d} N(0,1)$ for some explicitly computable constants μ and σ, where $N(0,1)$ denotes a standard normal distribution; if $d \ge 9$ then the sequence of random variables X_n^* does not converge to a fixed limit law.

– If $1 \le d \le 7$, then the Kolmogorov distance between X_n^* and $N(0,1)$ is of order $n^{-1/2}$; if $d = 8$, then the distance is of order $n^{-3(\frac{3}{2}-\sqrt{2})}$. Both orders are tight up to the implied constants in the O-terms. In each case, we also have a corresponding local limit theorem.

Methods of proof. We apply the *recursive moment-transfer approach* introduced in Hwang (2003), which is a *refined method of moments*. The general asymptotic tools are based on studying the so-called perturbed Cauchy-Euler differential equations of the form

$$P_0(\theta)f(z) = \sum_{1\le j<d} (1-z)^j P_j(\theta)f(z) + g(z),$$

where the P_j's are polynomials of degrees $\le d$ and $\theta := (1-z)(\mathrm{d}/\mathrm{d}z)$. In our case, $P_0(z) = z^d - 2^d$ and the main asymptotic problem is to derive asymptotics of the coefficients $[z^n]f(z)$ when those of $g(z)$ are known to some extent, where $[z^n]f(z)$ denotes the coefficient of z^n in the Taylor expansion of f.

References

[1] H.-H. Chern and H.-K. Hwang (2003). Partial match queries in random quadtrees. *SIAM Journal on Computing* **32** 904–915.

[2] F. N. David and D. E. Barton (1962). *Combinatorial Chance*. Charles Griffin, London.

[3] D. S. Dean and S. N. Majumdar (2002). Phase transition in a random fragmentation problem with applications to computer science. *Journal of Physics A* **35** L501–L507.

[4] L. Devroye (1987). Branching processes in the analysis of the heights of trees. *Acta Informatica* **24** 277–298.

[5] L. Devroye (1999). Universal limit laws for depths in random trees. *SIAM Journal on Computing* **28** 409–432.

[6] L. Devroye and L. Laforest (1990). An analysis of random d-dimensional quad trees. *SIAM Journal on Computing* **19** 821–832.

[7] R. A. Finkel and J. L. Bentley (1974). Quadtrees: A data structure for retrieval on composite keys. *Acta Informatica* **4** 1–9.

[8] P. Flajolet, G. Gonnet, C. Puech and J. M. Robson (1993). Analytic variations on quadtrees. *Algorithmica* **10** 473–500.

[9] P. Flajolet, G. Labelle, L. Laforest and B. Salvy (1995). Hypergeometrics and the cost structure of quadtrees. *Random Structures and Algorithms* **7** 117–114.

[10] P. Flajolet and T. Lafforgue (1994). Search costs in quadtrees and singularity perturbation asymptotics. *Discrete and Computational Geometry* **12** 151–175.

[11] M. Hoshi and P. Flajolet (1992). Page usage in a quadtree index. *BIT* **32** 384–402.

[12] H.-K. Hwang (2003). Second phase changes in random m-ary search trees and generalized quicksort: convergence rates. *Annals of Probability* **31** (2003) 609–629.

[13] G. Labelle and L. Laforest (1995). Combinatorial variations on multidimensional quadtrees. *Journal of Combinatorial Theory, Series A* **69** 1–16.

[14] G. Labelle and L. Laforest (1996). Étude de constantes universelles pour les arborescences hyperquaternaires de recherche. *Discrete Mathematics* **153** 199–211.

[15] H. Mahmoud (2002). The size of random bucket trees via urn models. *Acta Informatica* **38** 813–838.

[16] C. Martínez and A. Panholzer and H. Prodinger (2001). Partial match queries in relaxed multidimensional search trees. *Algorithmica* **29** 181–204.

[17] H. N. Minh and G. Jacob (2000). Symbolic integration of meromorphic differential systems via Dirichlet functions. *Discrete Mathematics* **210** 87–116.

[18] R. Neininger and L. Rüschendorf (1999). On the internal path length of d-dimensional quad trees. *Random Structures and Algorithms* **15** 25–41.
[19] R. Neininger and L. Rüschendorf (2001). Limit laws for partial match queries in quadtrees. *Annals of Applied Probability* **11** 452–469.
[20] H. Samet (1990a). *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, Reading, MA.
[21] H. Samet (1990b). *Applications of Spatial Data Structures: Computer Graphics, Image Processing, and GIS*. Addison-Wesley, Reading, MA.

Transfer theorems and asymptotic distributional results for m-ary search trees

James Allen Fill
(joint work with Nevin Kapur)

This talk is based on two papers [2] [3]. We derive asymptotics of moments and distributions, under the random permutation model (for an introduction, see, for example, [5]) on m-ary search trees (fundamental data structures for searching and sorting), for functionals f that satisfy recurrence relations of a simple additive form. More precisely, we assume that $f(T)$ for a tree T with $n \geq m-1$ keys equals $\sum_{i=1}^{m} f(L_i(T))+t_n$, where (t_n) is a given sequence, usually called the *toll sequence*, and $L_i(T)$ denotes the ith subtree pendant from the root of T. Many important functionals including the space requirement $[t_n = 1]$, internal path length $[t_n = n - (m-1)]$, and the so-called shape functional $[t_n = \ln \binom{n}{m-1}]$ fall under this framework. The approach is based on establishing *transfer theorems* that link the order of growth of the input into a particular (deterministic) recurrence to the order of growth of the output. The transfer theorems are used in conjunction with the method of moments to establish limit laws (see [2] for details, and for historical references). It is shown that (i) for suitably "small" toll sequences we have asymptotic normality if $m \leq 26$ and typically periodic behavior if $m \geq 27$; (ii) for "moderate" toll sequences we have convergence to non-normal distributions if $m \leq m_0$ (where $m_0 \geq 26$) and typically periodic behavior if $m \geq m_0 + 1$; and (iii) for "large" toll sequences we have convergence to non-normal distributions for all values of m.

Very recent research greatly sharpens understanding of periodic cases. Consider for example the space requirement of m-ary search trees under the random permutation model when $m \geq 27$ is fixed. Chauvin and Pouyanne [1] have shown using martingale techniques that the space requirement of an m-ary search tree on n keys equals

$$\frac{1}{H_m - 1}(n+1) + 2\mathrm{Re}[n^{\lambda_2}\Lambda] + \epsilon_n n^{\mathrm{Re}\lambda_2},$$

where H_m is the mth harmonic number $\sum_{i=1}^{m} i^{-1}$, the complex constant λ_2 is the root of the polynomial

$$\phi(z) := (z+1)(z+2)\cdots(z+m-1) - m!$$

having second largest real part and (for definiteness) positive imaginary part, Λ is a complex-valued random variable, and $\epsilon_n \to 0$ a.s. and in L^2 as $n \to \infty$. Extending the elementary but powerful "contraction method," we identify the distribution of Λ. To our knowledge, this is the first application of the contraction method to an oscillatory case.

In future research, we plan to show how our periodic-case contraction-method technique can be extended rather generally to other additive functionals under the random permutation model and, beyond these, to generalized Pólya urn schemes and multi-type branching processes as studied in [4]. In fact, we hope to prove periodic-case results as strong as those in [8], [6], and (perhaps) [7] for the non-periodic case.

References

[1] Chauvin, B. and Pouyanne, N. m-ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. *Random Structures and Algorithms* **24**(2):133–154, 2004.

[2] Fill, J. A. and Kapur, N. Transfer theorems and asymptotic distributional results for m-ary search trees. *Random Structures and Algorithms*, 2004, to appear.

[3] Fill, J. A. and Kapur, N. The space requirement of m-ary search trees: Distributional asymptotics for $m \geq 27$. Invited paper appearing in *Proceedings of the 7th Iranian Statistical Conference*, 2004.

[4] Janson, S. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, **110**(2):177–245, 2004.

[5] Mahmoud, H. M. *Evolution of Random Search Trees.* John Wiley & Sons Inc., New York, 1992. A Wiley–Interscience Publication.

[6] Neininger, R. On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures and Algorithms* **19**(3–4):498–524, 2001.

[7] Neininger, R. and Rüschendorf, L. A general limit theorem for recursive algorithms and combinatorial structures. *Annals of Applied Probability* **14**(1):378–418, 2004.

[8] Roesler, U. On the analysis of stochastic divide and conquer algorithms. *Algorithmica* **29**(1–2):238–261, 2001.

The contraction method in infinite-dimensional spaces

Svante Janson

(joint work with Michael Drmota and Ralph Neininger)

The contraction method has since it was introduced by Rösler [4] found a number of applications, in particular in the analysis of algorithms. It has been extended to random vectors in finite-dimensional spaces [2, 3] and to infinite-dimensional Hilbert spaces [1] (with the ℓ_2-metric).

We discuss the possibility to use the Zolotarev ζ_s metric in a Banach space. The main problem is completeness. We cannot show this in Banach spaces in general, or in natural spaces like $C[0,1]$, but we can do it in Hilbert spaces.

Theorem 1. *Let $s > 0$ and let H be a separable Hilbert space. The Zolotarev metric ζ_s is a complete metric on the set of all H-valued random variables with*

given k:th moments for $k = 0, 1, \dots, \lceil s \rceil - 1$. (The k:th moment lies in the k:th projective tensor power $H^{\hat{\otimes} k}$.)

Using this theorem, the proof in [3] works without changes also for Hilbert space valued variables.

In a tentative application, that is still work in progress, we apply this to the profile of random binary trees (and other trees). The relevant random variables can be represented as continuous functions on an interval $[a, b]$. It is desirable to obtain uniform convergence in $C[a, b]$, but this Banach space is not a Hilbert space and we do not know whether the Zolotarev metric is complete for it. Instead, we observe that the random functions in this case may be extended to analytic functions in a domain in the complex plane; we then use a Bergman space of analytic functions, which is a Hilbert space. We can thus use the contraction method there and obtain convergence in the Bergman space, which implies uniform convergence on compact subsets, and in particular on $[a, b]$.

References

[1] R.M. Burton & U. Rösler, An L_2 convergence theorem for random affine mappings. *J. Appl. Probab.* **32** (1995), 183–192.

[2] R. Neininger, On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms* **19** (2001), 498–524.

[3] R. Neininger & L. Rüschendorf, A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.* **14** (2004), 378–418.

[4] U. Rösler, A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, **25** (1991), 85–100.

The profile of m-ary search trees generated by van der Corput sequences

Wolfgang Steiner

In the last years, the height of binary search trees generated by sequences which are uniformly distributed modulo 1 has been studied. Devroye and Goudjil [2] proved that the height of the tree generated by the first N elements of the Weyl sequence $\{n\alpha\}$ satisfies $H(N) \sim \frac{12}{\pi^2} \log N \log\log N$ for almost all $\alpha \in (0, 1)$. It is therefore larger than the height of a random binary search tree ($c \log N$ with $c \approx 4.31107$ almost surely). They also showed that, for every sequence $(c_N)_{N\ge1}$ which decreases monotonically from 1 to 0, we have some α such that $H(N) \ge c_N N$ infinitely often. On the other hand, Dekking and van der Wal [1] proved $H(N) = o(N)$ for all uniformly distributed sequences modulo 1. Random suffix search trees (which are related to the Lehmer sequences $\{x_0 q^n\}$) were studied by Devroye and Neininger [3].

In this talk, we consider the van der Corput sequence $(x_n)_{n\ge1}$ to the base q for some integer $q \ge 2$, defined by

$$x_n = \sum_{j\ge0} \epsilon_j(n) q^{-j-1}, \text{ where } n = \sum_{j\ge0} \epsilon_j(n) q^j$$

is the (unique) q-ary digital expansion of n with digits $\epsilon_j(n) \in \{0, 1, \ldots, q-1\}$. We generate an m-ary search tree by successively inserting $x_1, x_2, \ldots$ at the appropriate place. Let $d(n)$ be the depth of the node where x_n is inserted. To study the distribution of the depth, we use the generating function

$$B(z,u) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} b_{jk} z^k u^j \text{ with } b_{jk} = |\{n \in \{q^j, \ldots, q^{j+1}-1\} : \ d(n) = k\}|.$$

It is convenient to look also at the m-ary search trees generated only by the elements $x_n \in (i/q, 1)$ for some $i \in \{0, \ldots, q-2\}$. Denote the generating functions of these trees by $B_i(z,u)$ (with $B_0(z,u) = B(z,u)$). Then we have

$$\begin{pmatrix} B_0(z,u) \\ \vdots \\ B_{q-2}(z,u) \end{pmatrix} = zA(u) \begin{pmatrix} B_0(z,u) \\ \vdots \\ B_{q-2}(z,u) \end{pmatrix} + \begin{pmatrix} Q_0(u) \\ \vdots \\ Q_{q-2}(u) \end{pmatrix}$$

for some matrix of polynomials $A(u)$ and some polynomials $Q_i(u)$. This gives

$$B(z,u) = \frac{Q(z,u)}{1 - P(z,u)}$$

for some polynomials $P(z,u), Q(z,u)$.

The following theorems follow almost directly from this representation.

Theorem 1. *Expected value and variance of the depth are given by*

$$E_N = \frac{1}{N} \sum_{n \le N} d(n) = \mu \log_q N + \mathcal{O}(1)$$

$$V_N = \frac{1}{N} \sum_{n \le N} (d(n) - E_N)^2 = \sigma^2 \log_q N + \mathcal{O}(1)$$

with

$$\mu = \frac{q'(1)}{q} = \frac{\frac{\partial P}{\partial z}(1, 1/q)}{\frac{1}{q}\frac{\partial P}{\partial u}(1, 1/q)}, \qquad \sigma^2 = \frac{q''(1)}{q} + \mu - \mu^2,$$

where $q(z)$ denotes the analytic solution of $P(z, 1/q(z)) = 1$ with $q(1) = q$.

Theorem 2. *If $m \neq q^M$ for all integers M, then we have, for every $\delta > 0$,*

$$\frac{1}{N}|\{n \le N : d(n) < E_N + xV_N\}| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt + \mathcal{O}\left((\log N)^{-1/2+\delta}\right)$$

uniformly for all real x as $N \to \infty$ and

$$|\{n \le N : d(n) = k\}| = \frac{N}{\sqrt{2\pi V_N}} \left(\exp\left(-\frac{(k - E_N)^2}{2V_N} \right) + \mathcal{O}\left((\log N)^{-1/2+\delta}\right) \right)$$

uniformly for all nonnegative integers k as $N \to \infty$.

Set $M = \lfloor \log_q m \rfloor$. For $m = q^M$, $m = 2$ (binary search trees) and $q = 2$ (the binary van der Corput sequence), we have simple formulae for μ and σ:

$$(1) \quad m = q^M : \qquad \mu = \frac{1}{M}, \qquad \sigma = 0$$

$$(2) \quad m = 2 : \qquad \mu = (q-1)\left(\frac{1}{2} + \frac{1}{q}\right), \qquad \sigma^2 = \frac{(q-1)(q-2)(q^2+3q-6)}{12q^2}$$

$$(3) \quad q = 2 : \qquad \mu = \frac{1}{M + \frac{m}{2^M} - 1}, \qquad \sigma^2 = \frac{\left(\frac{m}{2^M} - 1\right)\left(2 - \frac{m}{2^M}\right)}{\left(M + \frac{m}{2^M} - 1\right)^3}$$

A close look at the structure of the tree provides asymptotics for the height.

Theorem 3. *The height of the tree is given by*

$$H(N) = \max_{n \le N} d(n) = \frac{1}{M + h_{q,M}(\frac{m}{q^M})} \log_q N + \mathcal{O}(1),$$

where $h_{q,M} : [1, q) \to [0, 1)$ *is a monotonically increasing function.*

For $m < q$ ($M = 0$), $h_{q,0}(m)$ is approximately $1/\left\lceil \frac{q-1}{m-1} \right\rceil$.

References

[1] F. M. Dekking and P. van der Wal, Uniform distribution modulo one and binary search trees, *J. Theor. Nombres Bordx.* **14** (2002), 415–424.

[2] L. Devroye and A. Goudjil, A study of random Weyl trees, *Random Structures and Algorithms* **9** (1998), 271–295.

[3] L. Devroye and R. Neininger, Random suffix search trees, *Random Structures and Algorithms* **23** (2003), 357–396.

About the height of an extended class of random trees

Nicolas Broutin

(joint work with Luc Devroye)

Extreme parameters of random structures are one of the corner stones of analysis of algorithms. A probabilistic approach has been used by Devroye [2, 3, 4] to obtain the height of various kind of random trees, including random binary search trees, uniform recursive trees. We aim at unifying the results that are known and present a unique method that allows to derive the height of a large class of random trees. Our approach permits to obtain some pay off and to get some new applications.

Our work relies on the study of a tree of random variables instead of a tree that is randomly built. More formally, consider an infinite rooted b-ary tree T_∞. Assign independently to each node a random vector $\{(Z_1, E_1), \ldots, (Z_b, E_b)\}$, where $Z_1, \ldots, Z_b$ and $E_1, \ldots, E_b$ are distributed like Z and E respectively. Dependence is allowed between the random variables of a single random vector, but E and Z are always assumed independent. We map the variables of each random vector

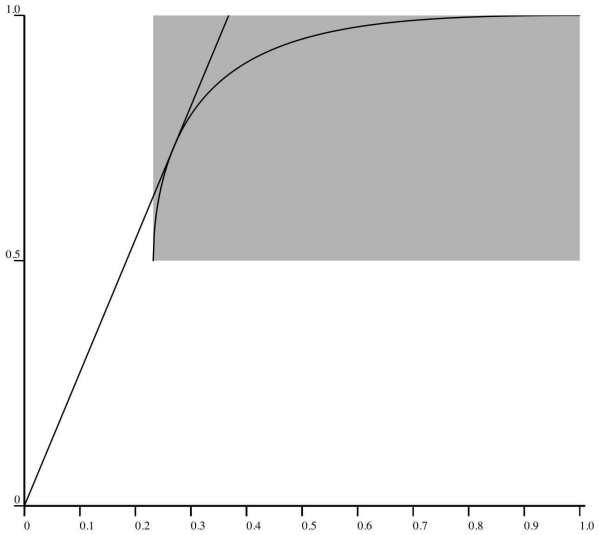

FIGURE 1. The curve $\mathscr{C}_{Z,E}$ (ρ on the horizontal axis, α on the vertical one) for uniform recursive trees, where $b = 2$, $E =$ exponential(1) and $Z =$ binomial(1/2).

to the edges down the node so that each edge e receives a couple (Z_e, E_e). Let $\pi(u)$ be the set of edges from the root to the node u. We introduce two additional random variables, $D_u = \sum_{e\in\pi(u)} Z_e$, that will account for the depth of the nodes, and $G_u = \sum_{e\in\pi(u)} E_e$, that has some meaning for the shape of the tree. The random tree T_n is made by pruning the nodes u for which $G_u \geq n$. We then define the height $H_n = \max\{D_u, u \in T_n\}$. Quite naturally, H_n is related to some large deviation properties. Recalling Crámer's Theorem for large deviations of sums of i.i.d. random variables $X_1, X_2, \ldots$, we introduce the rate function $\Lambda^\star_X$ such that

$$-\log \mathbf{P}\left\{\sum_{i=1}^{n} X_i \geq xn\right\} = \Lambda^\star_X(x)n + o(n),$$

assuming $\mathbf{E}\left\{e^{\lambda X}\right\} < \infty$ for some $\lambda > 0$. Then we can state our main result: under some minor technical conditions,

$$\frac{H_n}{n} \xrightarrow[n\to\infty]{} c$$

in probability, where c is the unique maximum value of α/ρ along the curve $\mathscr{C}_{Z,E}(b)$ and

$$\mathscr{C}_{Z,E}(b) = \{(\rho,\alpha) : \Lambda^\star_Z(\alpha) + \Lambda^\star_E(\rho) = \log b, \rho \leq \mathbf{E}E, \alpha \geq \mathbf{E}Z\}. \tag{1}$$

The curve $\mathscr{C}_{Z,E}$ turns out to be increasing and concave so that there exist a unique maximal value of α/ρ in the domain to be considered. An example of curve is shown on Figure 1.

One can notice that the trees we intend to study have logarithmic height. This is because our cutting scheme produces trees T_n that have an exponential number

of nodes. Using some age dependent branching processes, we are able to fix this minor problem and obtain the correct rate of growth.

Applications include of course random binary search trees ($Z = 1$ and $E =$ exponential(1)), random recursive trees ($Z =$ binomial(1/2), $E =$ exponential(1)). But we may also obtain some new results. Lopsided trees were introduced by Kapoor and Reingold [5], and studied mostly by Golin [1] on the basis of Varn codes [6]. These are codes in which characters have an encoding whose length may change. Therefore, algorithms processing sequences with such encodings are better represented by some trees whose edges have costs that change depending on their positions: from a node u, the edge to the first child costs c_1, c_2 to the second, ..., c_b to the bth one. We define a random version of lopsided trees by the following process: we start from a unique ancestor, and at each step we pick a random vertex and replace it by a node with all of its children. This can be generalized by replacing a uniform random vertex by some fixed deterministic tree. The equations one obtains for the height of these kinds of random trees are implicit and one has to turn towards numerical methods to get the asymptotic constants c involved in Theorem 1. Remember that this is already the case for random binary search trees.

References

[1] V. Choi and M. L. Golin, Lopsided trees I: a combinatorial analysis, *Algorithmica* , 31, (2001) 240–290.

[2] L. Devroye, A note on the height of binary search trees, *Journal of the ACM*, 33, (1986),489–498.

[3] L. Devroye, Branching processes in the analysis of the heights of trees, *Acta Informatica*, 24, (1987), 277–298.

[4] L. Devroye, Applications of the theory of records in the study of random trees, *Acta Informatica*, 26, (1988), 123–130.

[5] S. Kapoor and E. Reingold, Optimum lopsided binary trees, *Journal of the ACM*, 36 (1989) 3, 573–590.

[6] B. F. Varn, Optimal variable length codes (arbitrary symbol costs and equal code word probabilities, *Informat. Contr.*, (1971), 19, 289–301.

Silhouette and profile of binary search trees

Rudolf Grübel

Let T_n be the binary search tree associated with a random permutation of $\{1, 2, \ldots, n\}$. For each $t \in [0, 1]$ let $X_n(t)$ be the level of the external node of T_n along the path through the tree given by the binary expansion of t. This associates a stochastic process $(X_n(t))_{0 \le t \le 1}$ with T_n, the *silhouette* of the tree.

Regarding $t \mapsto X_n(t)$ as a random variable on the standard probability space we obtain a random probability mass function $\Psi(T_n) = (\Psi_k(T_n))_{k=1,2,\ldots}$, where $\Psi_k(T_n)$ is the Lebesgue measure of the subset of the unit interval where X_n takes the value k; $\Psi(T_n)$ can be interpreted as a discounted version of the profile of the tree T_n.

We obtain asymptotic bounds for the expected ℓ_2-distance of the random sequence $\Psi(T_n)$ from the Poisson mass function with parameter $\log n$, as $n \to \infty$. Problems of this type have often been tackled with methods from classical asymptotic analysis (singularity analysis, saddlepoint method). Here we use a more probabilistic approach, in particular Poisson approximation for the convolution of Bernoulli distributions and various conditioning steps that can be interpreted directly in terms of the underlying random mechanism.

On the contraction method with degenerate limit equation

Ralph Neininger

A large number of parameters of recursive combinatorial structures, random trees, and recursive algorithms satisfy recurrences of the divide-and-conquer type

$$Y_n \stackrel{\mathcal{L}}{=} Y_{I_n} + b_n, \quad n \geq n_0$$

where $n_0 \geq 1$, (I_n, b_n), (Y_k) are independent, b_n is random and I_n is a random index in $\{0, \ldots, n\}$ with $P(I_n = n) < 1$ for $n \geq n_0$, and $\stackrel{\mathcal{L}}{=}$ denotes equality in distribution. Typical parameters Y_n range from the depths of random trees, the number of various sub-structures in combinatorial structures, various cost measures of algorithms to parameters of communication models, and many more.

The contraction method is an efficient and quite universal probabilistic tool for the asymptotic analysis of recurrences as stated above. It has been introduced for the analysis of the Quicksort algorithm by Rösler (1991) and further developed independently in Rösler (1992) and Rachev and Rüschendorf (1995). It has been applied since then successfully to a large number of problems.

Recently, a fairly general unifying limit theorem for this type of recurrences has been obtained by the contraction method in Neininger and Rüschendorf (2004a) in the *nondegenerate case*, where the limit distribution of the normalized recurrence is uniquely characterized by a fixed point equation. By this result one in general obtains the limit distribution from the limiting recurrence and asymptotics of moments.

In this talk we extend the contraction method to a general limit theorem for the *degenerate case*. In the degenerate case the characterizing equations for the normalized algorithm degenerate in the limit to the trivial equation $X \stackrel{\mathcal{L}}{=} X$ and thus give no indication on the limit distribution. This case is also quite common in many examples.

To derive a limit in distribution for (Y_n) by the contraction method the first step is to introduce a scaling of Y_n, say $X_n := (Y_n - \mu_n)/\sigma_n$, where $\mu_n = \mathbb{E}\, Y_n$ and $\sigma_n = \sqrt{\mathrm{Var}(Y_n)}$ and to derive a recurrence relation for X_n:

$$X_n \stackrel{\mathcal{L}}{=} \frac{\sigma_{I_n}}{\sigma_n} X_{I_n} + b^{(n)}, \quad \text{where} \quad b^{(n)} = \frac{1}{\sigma_n}\left(b_n - \mu_n + \mu_{I_n}\right)$$

and with independence relations as above.

The next step to prove a limit theorem for X_n is to establish convergence of the random coefficients in the recursive equation:

$$\frac{\sigma_{I_n}}{\sigma_n} \to A, \quad b^{(n)} \to b,$$

thus leading to a limit equation of the form

$$X \stackrel{\mathcal{L}}{=} AX + b.$$

Here, (A, b) and X are independent. Essential for the application of the contraction method is that the limit equation has a unique solution under appropriate constraints. The final step of the method is to establish convergence of the X_n to the solution of the limit equation.

In this talk we discuss a case which appears quite often for parameters X_n with logarithmic orders for the variance. Here, in the limiting equation we are led to the case $A = 1$, $b = 0$, i.e., to the *degenerate limit equation*

$$X \stackrel{\mathcal{L}}{=} X.$$

The degenerate limit equation does not give any hint to a limit of the recursive sequence (X_n) and so the contraction method does not work in this case. However, a number of recurrences leading to a degenerate limit equation can be treated by the following theorem of Neininger and Rüschendorf (2004b):

Theorem: *Assume that $(Y_n)_{n\geq 0}$ satisfies the recursion stated above with $\|Y_n\|_3 < \infty$ for all $n \geq 0$ and*

$$\limsup_{n\to\infty} \mathbb{E} \log\left(\frac{I_n \vee 1}{n}\right) < 0, \quad \sup_{n\geq 1} \left\| \log\left(\frac{I_n \vee 1}{n}\right) \right\|_3 < \infty.$$

Furthermore assume that for real numbers α, λ, κ with $0 \leq \lambda < 2\alpha$ the mean and the variance of Y_n satisfy

$$\|b_n - \mu_n + \mu_{I_n}\|_3 = O\left(\ln^{\kappa} n\right), \quad \sigma_n^2 = C\, \ln^{2\alpha} n + O(\ln^{\lambda} n),$$

with some constant $C > 0$. If

$$\beta := \tfrac{3}{2} \wedge 3(\alpha - \kappa) \wedge 3(\alpha - \lambda/2) \wedge (\alpha - \kappa + 1) > 1,$$

then

$$\frac{Y_n - \mathbb{E}\, Y_n}{\sqrt{C}\, \ln^{\alpha} n} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

and we have the following rate of convergence for the Zolotarev-metric ζ_3:

$$\zeta_3\left(\frac{Y_n - \mathbb{E}\, Y_n}{\sqrt{Var(Y_n)}}, \mathcal{N}(0,1)\right) = O\left(\frac{1}{\ln^{\beta-1} n}\right).$$

References

[1] Neininger, R. and L. Rüschendorf (2004a) A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab. 14*, 378–418.
[2] Neininger, R. and L. Rüschendorf (2004b) On the contraction method with degenerate limit equation. *Ann. Probab. 32*, 2838–2856.
[3] Rösler, U. (1991) A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl. 25*, 85–100.
[4] Rösler, U. (1992) A fixed point theorem for distributions. *Stochastic Process. Appl. 42*, 195–214.
[5] Rachev, S. T. and L. Rüschendorf (1995) Probability metrics and recursive algorithms. *Adv. in Appl. Probab. 27*, 770–799.

Reporter: Gerold Alsmeyer and Holger Kösters

Participants

Prof. Dr. Gerold Alsmeyer

Institut für Mathematische
Statistik
Universität Münster
Einsteinstr. 62
48149 Münster

Nicolas Broutin

School of Computer Science
McGill University
Montreal Quebec H3A 2K6
Canada

Prof. Dr. Luc Devroye

School of Computer Science
McGill University
Montreal Quebec H3A 2K6
Canada

Prof. Dr. James Allen Fill

Department of Applied Mathematics
and Statistics
Johns Hopkins University
34th and Charles Streets
Baltimore, MD 21218-2682
USA

Prof. Dr. Rudolf Grübel

Institut für Mathematische
Stochastik
Universität Hannover
Welfengarten 1
30167 Hannover

Prof. Dr. Hsien-Kuei Hwang

Institute of Statistical Science
Academia Sinica
Taipei, 11529
TAIWAN

Prof. Dr. Svante Janson

Matematiska institutionen
Uppsala Universitet
Box 480
S-751 06 Uppsala

Dr. Holger Kösters

Institut für Mathematische
Statistik
Universität Münster
Einsteinstr. 62
48149 Münster

Dr. Dirk Kuhlbusch

Institut für Mathematische
Statistik
Universität Münster
Einsteinstr. 62
48149 Münster

Erin McLeish

School of Computer Science
McGill University
Montreal Quebec H3A 2K6
Canada

Dr. Ralph Neininger

Institut für Stochastik und
Mathematische Informatik
Johann-Wolfgang-Goethe-Universität
Robert-Mayer-Str. 10
60325 Frankfurt

Prof. Dr. Uwe Rösler

Mathematisches Seminar
Christian-Albrechts-Universität
Kiel
24098 Kiel

Prof. Dr. Ludger Rüschendorf

Institut für Mathematische
Stochastik
Universität Freiburg
Eckerstr. 1
79104 Freiburg

Dr. Wolfgang Steiner

Institut für Diskrete Mathematik
und Geometrie
TU Wien
Wiedner Hauptstr. 8 - 10
A-1040 Wien